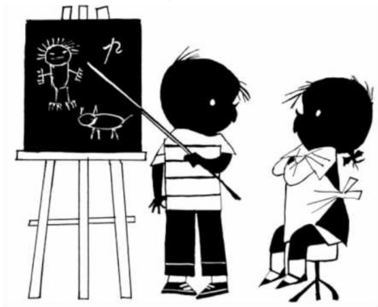
An Introduction to Stochastic Processes in Continuous Time:

the non-Jip-and-Janneke-language approach

Flora Spieksma

adaptation of the text by
Harry van Zanten

to be used at your own expense



May 5, 2016

Contents

Chapter 1

Stochastic Processes

1.1 Introduction

Loosely speaking, a stochastic process is a phenomenon that can be thought of as evolving in time in a random manner. Common examples are the location of a particle in a physical system, the price of stock in a financial market, interest rates, mobile phone networks, internet traffic, etcetc.

A basic example is the erratic movement of pollen grains suspended in water, so-called Brownian motion. This motion was named after the English botanist R. Brown, who first observed it in 1827. The movement of pollen grain is thought to be due to the impacts of water molecules that surround it. Einstein was the first to develop a model for studying the erratic movement of pollen grains in in an article in 1926. We will give a sketch of how this model was derived. It is more heuristically than mathematically sound.

The basic assumptions for this model (in dimension 1) are the following:

1) the motion is continuous.

Moreover, in a time-interval $[t, t+\tau]$, τ small,

2) particle movements in two non-overlapping time intervals of length τ are mutually independent;

3) the relative proportion of particles experiencing a displacement of size between δ and $\delta+d\delta$ is approximately $\phi_\tau(\delta)$ with

- the probability of *some* displacement is 1: $\int_{-\infty}^{\infty} \phi_\tau(\delta) d\delta = 1$;
- the *average* displacement is 0: $\int_{-\infty}^{\infty} \delta \phi_\tau(\delta) d\delta = 0$;
- the variation in displacement is linear in the length of the time interval: $\int_{-\infty}^{\infty} \delta^2 \phi_\tau(\delta) d\delta = D\tau$, where $D \geq 0$ is called the *diffusion coefficient*.

Denote by $f(x,t)$ the density of particles at position x, at time t. Under differentiability assumptions, we get by a first order Taylor expansion that

$$f(x, t+\tau) \approx f(x,t) + \tau \frac{\partial f}{\partial t}(x,t).$$

On the other hand, by a second order expansion

$$\begin{aligned}
f(x,t+\tau) &= \int_{-\infty}^{\infty} f(x-\delta,t)\phi_\tau(\delta)d\delta \\
&\approx \int_{-\infty}^{\infty} [f(x,t) - \delta\frac{\partial f}{\partial x}(x,t) + \tfrac{1}{2}\delta^2\frac{\partial^2 f}{\partial x^2}(x,t)]\phi_\tau(\delta)d\delta \\
&\approx f(x,t) + \tfrac{1}{2}D\tau\frac{\partial^2 f}{\partial x^2}(x,t).
\end{aligned}$$

Equating gives rise to the heat equation in one dimension:

$$\frac{\partial f}{\partial t} = \tfrac{1}{2}D\frac{\partial^2 f}{\partial x^2},$$

which has the solution

$$f(x,t) = \frac{\#\text{particles}}{\sqrt{4\pi Dt}}\cdot e^{-x^2/4Dt}.$$

So $f(x,t)$ is the density of a $\mathcal{N}(0,4Dt)$-distributed random variable multiplied by the number of particles.

Side remark. In section 1.5 we will see that under these assumptions paths of pollen grain through liquid are non-differentiable. However, from physics we know that the velocity of a particle is the derivative (to time) of its location. Hence pollen grain paths must be differentiable. We have a conflict between the properties of the physical model and the mathematical model. What is wrong with the assumptions? Already in 1926 editor R. Fürth doubted the validity of the independence assumption (2). Recent investigation seems to have confirmed this doubt.

Brownian motion will be one of our objects of study during this course. We will now turn to a mathematical definition.

Definition 1.1.1 Let T be a set and $(E,\mathcal{E})$ a measurable space. A *stochastic process* indexed by T, with values in $(E,\mathcal{E})$, is a collection $X=(X_t)_{t\in T}$ of measurable maps from a (joint) probability space $(\Omega,\mathcal{F},\mathsf{P})$ to $(E,\mathcal{E})$. X_t is called a *random element* as a generalisation of the concept of a random variable (where $(E,\mathcal{E})=(\mathbf{R},\mathcal{B})$). The space $(E,\mathcal{E})$ is called the state space of the process.

Review BN §1

The index t is a time parameter, and we view the index set T as the set of all observation instants of the process. In these notes we will usually have $T=\mathbf{Z}_+=\{0,1,\ldots\}$ or $T=\mathbf{R}_+=[0,\infty)$ (or T is a sub-interval of one these sets). In the former case, we say that time is *discrete*, in the latter that time is *continuous*. Clearly a discrete-time process can always be viewed as a continuous-time process that is constant on time-intervals $[n,n+1)$.

The state space $(E,\mathcal{E})$ will generally be a Euclidian space $\mathbf{R}^d$, endowed with its Borel σ-algebra $\mathcal{B}(\mathbf{R}^d)$. If E is the state space of the process, we call the process E*-valued.*

For every fixed observation instant $t\in T$, the stochastic process X gives us an E-valued random element X_t on $(\Omega,\mathcal{F},\mathsf{P})$. We can also fix $\omega\in\Omega$ and consider the map $t\to X_t(\omega)$ on T. These maps are called the *trajectories* or *sample paths* of the process. The sample paths are functions from T to E and so they are elements of the function space E^T. Hence, we can view the process X as an E^T-valued random element.

Quite often, the sample paths belong to a nice subset of this space, e.g. the continuous or right-continuous functions, alternatively called the path space. For instance, a discrete-time process viewed as the continuous-time process described earlier, is a process with right-continuous sample paths.

Clearly we need to put an appropriate σ-algebra on the path space E^T. For consistency purposes it is convenient that the marginal distribution of X_t be a probability measure on the path space. This is achieved by ensuring that the projection $x \to x_t$, where $t \in T$, is measurable. The σ-algebra $\mathcal{E}^T$, described in BN §2, is the minimal σ-algebra with this property.

Review BN §2

We will next introduce the formal requirements for the stochastic processes that are called Brownian motion and Poisson process respectively. First, we introduce processes with independent increments.

Definition 1.1.2 Let E be a separable Banach space, and $\mathcal{E}$ the Borel-σ-algebra of subsets of E. Let $T = [0, \tau] \subset \mathbf{R}^+$. Let $X = \{X_t\}_{t \in T}$ be an $(E, \mathcal{E})$-valued stochastic process, defined on an underlying probability space $(\Omega, \mathcal{F}, \mathsf{P})$.

i) X is called a process with independent increments, if $\sigma(X_t - X_s)$ and $\sigma(X_u, u \leq s)$, are independent for all $s \leq t \leq \tau$.

ii) X is called a process with stationary, independent increments, if, in addition, $X_t - X_s \stackrel{\mathsf{d}}{=} X_{t-s} - X_0$, for $s \leq t \leq \tau$.

The mathematical model of the physical Brownian motion is a stochastic process that is defined as follows.

Definition 1.1.3 The stochastic process $W = (W_t)_{t \geq 0}$ is called a (standard) *Brownian motion* or *Wiener process*, if

i) $W_0 = 0$, a.s.;

ii) W is a stochastic process with stationary, independent increments;

iii) $W_t - W_s \stackrel{\mathsf{d}}{=} \mathcal{N}(0, t - s)$;

iv) almost all sample paths are continuous.

In these notes we will abbreviate 'Brownian motion' as BM. Property (i) tells that standard BM *starts at 0*. A stochastic process with property (iv) is called a *continuous process*. Similarly, a stochastic process is said to be *right-continuous* if almost all of its sample paths are right-continuous functions. Finally, the acronym *cadlag* (continu à droite, limites à gauche) is used for processes with right-continuous sample paths having finite left-hand limits at every time instant.

Simultaneously with Brownian motion we will discuss another fundamental process: the Poisson process.

Definition 1.1.4 A real-valued stochastic process $N = (N_t)_{t\geq 0}$ is called a Poisson process if

i) N is a counting process, i.o.w.

- **a)** N_t takes only values in $(\mathbf{Z}_+, 2^{\mathbf{Z}_+})$, $t \geq 0$;
- **b)** $t \mapsto N_t$ is increasing, i.o.w. $N_s \leq N_t$, $t \geq s$.
- **c)** (no two occurrences can occur simultaneously) $\lim_{s\downarrow t} N_s \leq \lim_{s\uparrow t} N_s + 1$, for all $t \geq 0$.

ii) $N_0 = 0$, a.s.;

iii) N is a stochastic process with stationary, independent increments.

Note: so far we do not know yet whether a BM process and a Poisson process exist at all! The Poisson process can be constructed quite easily and we will do so first before delving into more complex issues.

Construction of the Poisson process The construction of a Poisson process is simpler than the construction of Brownian motion. It is illustrative to do this first.

Let a probability space $(\Omega, \mathcal{F}, \mathsf{P})$ be given. We construct a sequence of i.i.d. $(\mathbf{R}_+, \mathcal{B}(\mathbf{R}_+))$-measurable random variables X_n, $n = 1, \ldots$, on this space, such that $X_n \stackrel{\mathsf{d}}{=} \exp(\lambda)$. This means that

$$\mathsf{P}\{X_n > t\} = e^{-\lambda t}, \quad t \geq 0.$$

Put $S_0 = 0$, and $S_n = \sum_{i=1}^n X_i$. Clearly S_n, $n = 0, \ldots$ are in increasing sequence of random variables. Since X_n are all $\mathcal{F}/\mathcal{B}(\mathbf{R}_+)$-measurable, so are S_n. Next define

$$N_t = \max\{n \,|\, S_n \leq t\}.$$

We will show that this is a Poisson process. First note that N_t can be described alternatively as

$$N_t = \sum_{n=1}^{\infty} \mathbf{1}_{\{S_n \leq t\}}.$$

N_t maybe infinite, but we will show that it is finite with probability 1 for all t. Moreover, no two points S_n and S_{n+1} are equal. Denote by $\mathcal{E}$ the σ-algebra generated by the one-point sets of $\mathbf{Z}_+$.

Lemma 1.1.5 *N_t is $\mathcal{F}/2^{\mathbf{Z}_+}$-measurable. There exists a set $\Omega^* \in \mathcal{F}$ with $\mathsf{P}\{\Omega^*\} = 1$, such that $N_t(\omega) < \infty$ for all $t \geq 0$, $\omega \in \Omega^*$, and $S_n(\omega) < S_{n+1}(\omega)$, $n = 0, \ldots$.*

Proof. From the law of large numbers we find a set Ω', $\mathsf{P}\{\Omega'\} = 1$, such that $N_t(\omega) < \infty$ for all $t \geq 0$, $\omega \in \Omega'$. It easily follows that there exists a subset $\Omega^* \subset \Omega'$, $\mathsf{P}\{\Omega^*\} = 1$, meeting the requirements of the lemma. Measurability follows from the fact that $\mathbf{1}_{\{S_n \leq t\}}$ is measurable. Hence a finite sum of these terms is measurable. The infinite sum is then measurable as well, being the monotone limit of measurable functions. QED

Since $\Omega^* \in \mathcal{F}$, we may restrict to this smaller space without further ado. Denote the restricted probability space again by $(\Omega, \mathcal{F}, \mathsf{P})$.

Theorem 1.1.6 *For the constructed process N on $(\Omega, \mathcal{F}, \mathsf{P})$ the following hold.*

i) *N is a $(\mathbf{Z}_+, \mathcal{E})$-measurable stochastic process that has properties (i,...,iv) from Definition 1.1.4. Moreover, N_t is $\mathcal{F}/\mathcal{E}$-measurable, it has a Poisson distribution with parameter λt and S_n has a Gamma distribution with parameters (n, λ). In particular $\mathsf{E}N_t = \lambda t$, and $\mathsf{E}N_t^2 = \lambda t + (\lambda t)^2$.*

ii) *All paths of N are cadlag.*

Proof. The second statement is true by construction, as well as are properties (i,ii). The fact that N_t has a Poisson distribution with parameter λt, and that S_n has $\Gamma(n, \lambda)$ distribution is standard.

We will prove property (iv). It suffices to show for $t \geq s$ that $N_t - N_s$ has a Poisson $(\lambda(t-s))$ distribution. Clearly

$$\begin{aligned} \mathsf{P}\{N_t - N_s = j\} &= \sum_{i \geq 0} \mathsf{P}\{N_s = i, N_t - N_s = j\} \\ &= \sum_{i \geq 0} \mathsf{P}\{S_i \leq s, S_{i+1} > s, S_{i+j} \leq t, S_{i+j+1} > t\}. \end{aligned} \tag{1.1.1}$$

First let $i, j > 1$. Recall the density $f_{n,\lambda}$ of the $\Gamma(n, \lambda)$ distribution:

$$f_{n,\lambda}(x) = \frac{\lambda^n x^{n-1} e^{-\lambda x}}{\Gamma(n)}, \quad n \neq 1$$

where $\Gamma(n) = (n-1)!$. Then, with a change of variable $u = s_2 - (s - s_1)$ in the third equation,

$$\begin{aligned} &\mathsf{P}\{N_t - N_s = j, N_s = i\} = \mathsf{P}\{S_i \leq s, S_{i+1} > s, S_{i+j} \leq t, S_{i+j+1} > t\} \\ &= \int_0^s \int_{s-s_1}^{t-s_1} \int_0^{t-s_2-s_1} e^{-\lambda(t-s_3-s_2-s_1)} f_{j-1,\lambda}(s_3) ds_3 \lambda e^{-\lambda s_2} ds_2 f_{i,\lambda}(s_1) ds_1 \\ &= \int_0^s \int_0^{t-s} \int_0^{t-s-u} e^{-\lambda(t-s-u-s_3)} f_{j-1,\lambda}(s_3) ds_3 \lambda e^{-\lambda u} du \cdot e^{-\lambda(s-s_1)} f_{i,\lambda}(s_1) ds_1 \\ &= \mathsf{P}\{S_j \leq t-s, S_{j+1} > t-s\} \cdot \mathsf{P}\{S_i \leq s, S_{i+1} > s\} \\ &= \mathsf{P}\{N_{t-s} = j\} \mathsf{P}\{N_s = i\}. \end{aligned} \tag{1.1.2}$$

For $i = 0, 1, j = 1$, we get the same conclusion. (1.1.1) then implies that $\mathsf{P}\{N_t - N_s = j\} = \mathsf{P}\{N_{t-s} = j\}$, for $j > 0$. By summing over $j > 0$ and substracting from 1, we get the relation for $j = 0$ and so we have proved property (iv).

Finally, we will prove property (iii). Let us first consider $\sigma(N_u, u \leq s)$. This is the smallest σ-algebra that makes all maps $\omega \mapsto N_u(\omega)$, $u \leq s$, measurable. Section 2 of BN studies its structure. It follows that (see Exercise 1.1) the collection $\mathcal{I}$, with

$$\mathcal{I} = \Big\{ A \in \mathcal{F} \mid \quad \exists n \in \mathbf{Z}_+, t_0 \leq t_1 < t_2 < \cdots < t_n, t_l \in [0, s], i_l \in \mathbf{Z}_+, l = 0, \ldots, n, \\ \text{such that } A = \{N_{t_l} = i_l, l = 0, \ldots, n\} \Big\}$$

a π-system for this σ-algebra.

To show independence property (iii), it therefore suffices show for each n, for each sequence $0 \leq t_0 < \cdots < t_n$, and $i_0, \ldots, i_n, i$ that

$$\mathsf{P}\{N_{t_l} = i_l, l = 0, \ldots, n, N_t - N_s = i\} = \mathsf{P}\{N_{t_l} = i_l, l = 0, \ldots, n\} \cdot \mathsf{P}\{N_t - N_s = i\}.$$

This is analogous to the proof of (1.1.2). QED

A final observation. We have constructed a mapping $N : \Omega \to \Omega' \subset \mathbf{Z}_+^{[0,\infty)}$, with $\mathbf{Z}_+^{[0,\infty)}$ the space of all integer-valued functions. The space Ω' consists of all integer valued paths ω', that are right-continuous and non-decreasing, and have the property that $\omega'_t \leq \lim_{s\uparrow t} \omega'_s + 1$.

It is desirable to consider Ω' as the underlying space. One can then construct a Poisson process directly on this space, by taking the identity map. This will be called the *canonical process*. The σ-algebra to consider, is then the minimal σ-algebra $\mathcal{F}'$ that makes all maps $\omega' \mapsto \omega'_t$ measurable, $t \geq 0$. It is precisely $\mathcal{F}' = \mathcal{E}^{[0,\infty)} \cap \Omega'$.

Then $\omega \mapsto N(\omega)$ is measurable as a map $\Omega \to \Omega'$. On $(\Omega', \mathcal{F}')$ we now put the induced probability measure P' by $\mathsf{P}'\{A\} = \mathsf{P}\{\omega \mid N(\omega) \in A\}$.

In order to construct BM, we will next discuss a procedure to construct a stochastic process, with given marginal distributions.

1.2 Finite-dimensional distributions

In this section we will recall Kolmogorov's theorem on the existence of stochastic processes with prescribed finite-dimensional distributions. We will use the version that is based on the fact hat T is ordered. It allows to prove the existence of a process with properties (i,ii,iii) of Definition 1.1.3.

Definition 1.2.1 Let $X = (X_t)_{t\in T}$ be a stochastic process. The distributions of the finite-dimensional vectors of the form $(X_{t_1}, X_{t_2}, \ldots, X_{t_n})$, $t_1 < t_2 < \cdots < t_n$, are called the *finite-dimensional distributions* (fdd's) of the process.

It is easily verified that the fdd's of a stochastic process form a consistent system of measures in the sense of the following definition.

Definition 1.2.2 Let $T \subset \mathbf{R}$ and let $(E, \mathcal{E})$ be a measurable space. For all $n \in \mathbf{Z}_+$ and all $t_1 < \cdots < t_n$, $t_i \in T$, $i = 1, \ldots, n$, let $\mu_{t_1,\ldots,t_n}$ be a probability measure on $(E^n, \mathcal{E}^n)$. This collection of measures is called *consistent* if it has the property that

$$\mu_{t_1,\ldots,t_{i-1},t_{i+1},\ldots,t_n}(A_1\times\cdots\times A_{i-1}\times A_{i+1}\times\cdots\times A_n) = \mu_{t_1,\ldots,t_n}(A_1\times\cdots\times A_{i-1}\times E\times A_{i+1}\times\cdots\times A_n),$$

for all $A_1, \ldots, A_{i-1}, A_{i+1}, \ldots, A_n \in \mathcal{E}$.

The Kolmogorov consistency theorem states that, conversely, under mild regularity conditions, every consistent family of measures is in fact the family of fdd's of some stochastic process.

Some assumptions are needed on the state space $(E, \mathcal{E})$. We will assume that E is a *Polish space*. This is a topological space, on which we can define a metric that consistent with the topology, and which makes the space complete and separable. As $\mathcal{E}$ we take the Borel-σ-algebra, i.e. the σ-algebra generated by the open sets. Clearly, the Euclidian spaces $(\mathbf{R}^n, \mathcal{B}(\mathbf{R}^n))$ fit in this framework.

Theorem 1.2.3 (Kolmogorov's consistency theorem) *Suppose that E is a Polish space and $\mathcal{E}$ its Borel-σ-algebra. Let $T \subset \mathbf{R}$ and for all $n \in \mathbf{Z}_+$, $t_1 < \ldots < t_n \in T$, let $\mu_{t_1,\ldots,t_n}$ be a probability measure on $(E^n, \mathcal{E}^n)$. If the measures $\mu_{t_1,\ldots,t_n}$ form a consistent system, then* *there exists a probability measure* P *on* E^T*, such that the* canonical (or co-ordinate variable) process $(X_t)_t$ *on* $(\Omega = E^T, \mathcal{F} = \mathcal{E}^T, \mathsf{P})$*, defined by*

$$X(\omega) = \omega, \quad X_t(\omega) = \omega_t,$$

has fdd*'s* $\mu_{t_1,\ldots,t_n}$*.*

The proof can for instance be found in Billingsley (1995). Before discussing this theorem, we will discuss its implications for the existence of BM.

Review BN §4 on multivariate normal distributions

Corollary 1.2.4 *There exists a probability measure* P *on the space* $(\Omega = \mathbf{R}^{[0,\infty)}, \mathcal{F} = \mathcal{B}(\mathbf{R})^{[0,\infty)})$*, such that the co-ordinate process* $W = (W_t)_{t\geq 0}$ *on* $(\Omega = \mathbf{R}^{[0,\infty)}, \mathcal{F} = \mathcal{B}(\mathbf{R})^{[0,\infty)}, \mathsf{P})$ *has properties (i,ii,iii) of Definition 1.1.3.*

Proof. The proof could contain the following ingredients.
(1) Show that for $0 \leq t_0 < t_1 < \cdots < t_n$, there exist multivariate normal distributions with covariance matrices

$$\mathbf{\Sigma} = \begin{pmatrix} t_0 & 0 & \ldots & \ldots & 0 \\ 0 & t_1 - t_0 & 0 & \ldots & 0 \\ 0 & 0 & t_2 - t_1 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \ldots & t_n - t_{n-1} \end{pmatrix},$$

and

$$\mathbf{\Sigma}_{t_0,\ldots,t_n} = \begin{pmatrix} t_0 & t_0 & \ldots & \ldots & t_0 \\ t_0 & t_1 & t_1 & \ldots & t_1 \\ t_0 & t_1 & t_2 & \ldots & t_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ t_0 & t_1 & t_2 & \ldots & t_n \end{pmatrix}.$$

(2) Show that a stochastic process W has properties (i,ii,iii) if and only if for all $n \in \mathbf{Z}$, $0 \leq t_0 < \ldots < t_n$ the vector $(W_{t_0}, W_{t_1} - W_{t_0}, \ldots, W_{t_n} - W_{t_{n-1}}) \stackrel{\mathsf{d}}{=} \mathsf{N}(0, \mathbf{\Sigma})$.
(3) Show that for a stochastic process W the (a) and (b) below are equivalent:

a) for all $n \in \mathbf{Z}$, $0 \leq t_0 < \ldots < t_n$ the vector $(W_{t_0}, W_{t_1} - W_{t_0}, \ldots, W_{t_n} - W_{t_{n-1}}) \stackrel{\mathsf{d}}{=} \mathsf{N}(0, \mathbf{\Sigma})$;

b) for all $n \in \mathbf{Z}$, $0 \leq t_0 < \ldots < t_n$ the vector $(W_{t_0}, W_{t_1}, \ldots, W_{t_n}) \stackrel{\mathsf{d}}{=} \mathsf{N}(0, \mathbf{\Sigma}_{t_0,\ldots,t_n})$.

QED

The drawback of Kolmogorov's Consistency Theorem is, that in principle all functions on the positive real line are possible sample paths. Our aim is the show that we may restrict to the subset of continuous paths in the Brownian motion case.

However, the set of continuous paths is not even a measurable subset of $\mathcal{B}(\mathbf{R})^{[0,\infty)}$, and so the probability that the process W has continuous paths is not well defined. The next section discussed how to get around the problem concerning continuous paths.

1.3 Kolmogorov's continuity criterion

Why do we really insist on Brownian motion to have continuous paths? First of all, the connection with the physical process. Secondly, without regularity properies like continuity, or weaker right-continuity, events of interest are not ensured to measurable sets. An example is: $\{\sup_{t\geq 0} W_t \leq x\}$, $\inf\{t \geq 0 \,|\, W_t = x\}$.

The idea to address this problem, is to try to modify the constructed process W in such a way that the resulting process, $\tilde{W}$ say, has continuous paths and satisfies properties (i,ii,iii), in other words, it has the same fdd's as W. To make this idea precise, we need the following notions.

Definition 1.3.1 Let X and Y be two stochastic processes, indexed by the same set T and with the same state space $(E, \mathcal{E})$, defined on probability spaces $(\Omega, \mathcal{F}, \mathsf{P})$ and $(\Omega', \mathcal{F}', \mathsf{P}')$ respectively. The processes are called *versions* of each other, if they have the same fdd's. In other words, if for all $n \in \mathbf{Z}_+$, $t_1, \ldots, t_n \in T$ and $B_1, \ldots, B_n \in \mathcal{E}$

$$\mathsf{P}\{X_{t_1} \in B_1, X_{t_2} \in B_2, \ldots, X_{t_n} \in B_n\} = \mathsf{P}'\{Y_{t_1} \in B_1, Y_{t_2} \in B_2, \ldots, Y_{t_n} \in B_n\}.$$

X and Y are both $(E, \mathcal{E})$-valued stochastic processes. They can be viewed as random elements with values in the measurable path space $(E^T, \mathcal{E}^T)$. X induces a probability measure P_X on the path space with $\mathsf{P}_X\{A\} = \mathsf{P}\{X^{-1}(A)\}$. In the same way Y induces a probability P_Y on the path space. Since X and Y have the same fdds, it follows for each $n \in \mathbf{Z}_+$ and $t_1 < \cdots < t_n$, $t_1, \ldots, t_n \in T$, and $A_1, \ldots, A_n \in \mathcal{E}$ that

$$\mathsf{P}_X\{A\} = \mathsf{P}_Y\{A\},$$

for $A = \{x \in E^T \,|\, x_{t_i} \in A_i,\, i = 1, \ldots, n\}$. The collection of sets B of this form are a π-system generating $\mathcal{E}^T$ (cf. remark after BN Lemma 2.1), hence $\mathsf{P}_X = \mathsf{P}_Y$ on $(E^T, \mathcal{E}^T)$ by virtue of BN Lemma 1.2(i).

Definition 1.3.2 Let X and Y be two stochastic processes, indexed by the same set T and with the same state space $(E, \mathcal{E})$, defined on the same probability space $(\Omega, \mathcal{F}, \mathsf{P})$.

i) The processes are called *modifications* of each other, if for every $t \in T$

$$X_t = Y_t, \quad \text{a.s.}$$

ii) The processes are called *indistinguishable*, if there exists a set $\Omega^* \in \mathcal{F}$, with $\mathsf{P}\{\Omega^*\} = 1$, such that for every $\omega \in \Omega^*$ the paths $t \to X_t(\omega)$ and $t \to Y_t(\omega)$ are equal.

The third notion is stronger than the second notion, which in turn is clearly stronger than the first one: if processes are indistinguishable, then they are modifications of each other. If they are modifications of each other, then they certainly are versions of each other. The reverse is not true in general (cf. Exercises 1.3, 1.6). The following theorem gives a sufficient condition for a process to have a continuous modification. This condition (1.3.1) is known as *Kolmogorov's continuity condition*.

Denote by $C^d[0, T]$ the collection of $\mathbf{R}^d$-valued continuous functions on $[0, T]$.

Theorem 1.3.3 (Kolmogorov's continuity criterion) *Let $X = (X_t)_{t\in[0,T]}$ be an $(\mathbf{R}^d, \mathcal{B}^d)$-valued stochastic process on a probability space $(\Omega, \mathcal{F}, \mathsf{P})$. Suppose that there exist constants $\alpha, \beta, K > 0$ such that*

$$\mathsf{E}\|X_t - X_s\|^\alpha \leq K|t-s|^{1+\beta}, \tag{1.3.1}$$

for all $s, t \in [0,T]$. Then there exists a (everywhere!) continuous modification $\hat{X}$ of X, i.o.w. $\hat{X}(\omega)$ is a continuous function on $[0,T]$ for each $\omega \in \Omega$. Thus the map $\hat{X} : (\Omega, \mathcal{F}, \mathsf{P}) \to (C^d[0,T], C^d[0,T] \cap \mathcal{B}(\mathbf{R}^d)^{[0,T]})$ is an $\mathcal{F}/C^d[0,T] \cap \mathcal{B}(\mathbf{R}^d)^{[0,T]}$-measurable map.

Note: $\beta > 0$ is needed for the continuous modification to exist. See Exercise 1.5.

Proof. The proof consists of the following steps:

1 (1.3.1) implies that X_t is continuous in probability on $[0,T]$;

2 X_t is a.s. uniformly continuous on a countable, dense subset $D \subset [0,T]$;

3 'Extend' X to a continuous process Y on all of $[0,T]$.

4 Show that Y is a well-defined stochastic process, and a continuous modification of X.

Without loss of generality we may assume that $T = 1$.

Step 1 Apply Chebychev's inequality to the r.v. $Z = \|X_t - X_s\|$ and the function $\phi : \mathbf{R} \to \mathbf{R}^+$ given by

$$\phi(x) = \begin{cases} 0, & x \leq 0 \\ x^\alpha, & x > 0. \end{cases}$$

Since ϕ is non-decreasing, non-negative and $\mathsf{E}\phi(Z) < \infty$, it follows for every $\epsilon > 0$ that

$$\mathsf{P}\{\|X_t - X_s\| > \epsilon\} \leq \frac{\mathsf{E}\|X_t - X_s\|^\alpha}{\epsilon^\alpha} \leq \frac{K|t-s|^{1+\beta}}{\epsilon^\alpha}. \tag{1.3.2}$$

Let $t, t_1, \ldots \in [0,1]$ with $t_n \to t$ as $n \to \infty$. By the above,

$$\lim_{n\to\infty} \mathsf{P}\{\|X_t - X_{t_n}\| > \epsilon\} = 0,$$

for any $\epsilon > 0$. Hence $X_{t_n} \xrightarrow{P} X_t$, $n \to \infty$. In other words, X_t is continuous in probability.

Step 2 As the set D we choose the dyadic rationals. Let $D_n = \{k/2^n \mid k = 0, \ldots, 2^n\}$. Then D_n is an increasing sequence of sets. Put $D = \cup_n D_n = \lim_{n\to\infty} D_n$. Clearly $\bar{D} = [0,1]$, i.e. D is dense in $[0,1]$.

Fix $\gamma \in (0, \beta/\alpha)$. Apply Chebychev's inequality (1.3.2) to obtain

$$\mathsf{P}\{\|X_{k/2^n} - X_{(k-1)/2^n}\| > 2^{-\gamma n}\} \leq \frac{K 2^{-n(1+\beta)}}{2^{-\gamma n \alpha}} = K 2^{-n(1+\beta-\alpha\gamma)}.$$

It follows that

$$\begin{aligned} \mathsf{P}\{\max_{1\leq k\leq 2^n} \|X_{k/2^n} - X_{(k-1)/2^n}\| > 2^{-\gamma n}\} &\leq \sum_{k=1}^{2^n} \mathsf{P}\{\|X_{k/2^n} - X_{(k-1)/2^n} > 2^{-\gamma n}\|\} \\ &\leq 2^n K 2^{-n(1+\beta-\alpha\gamma)} = K 2^{-n(\beta-\alpha\gamma)}. \end{aligned}$$

Define the set $A_n = \{\max_{1\le k\le 2^n} \|X_{k/2^n} - X_{(k-1)/2^n}\| > 2^{-\gamma n}\}$. Then

$$\sum_n \mathsf{P}\{A_n\} \le \sum_n K 2^{-n(\beta-\alpha\gamma)} = K\frac{1}{1-2^{-(\beta-\alpha\gamma}} < \infty,$$

since $\beta-\alpha\gamma > 0$. By virtue of the first Borel-Cantelli Lemma this implies that $\mathsf{P}\{\limsup_m A_m\} = \mathsf{P}\{\cap_{m=1}^\infty \cup_{n=m}^\infty A_n\} = 0$. Hence there exists a set $\Omega^* \subset \Omega$, $\Omega^* \in \mathcal{F}$, with $\mathsf{P}\{\Omega^*\} = 1$, such that for each $\omega \in \Omega^*$ there exists N_ω, for which $\omega \notin \cup_{n\ge N_\omega} A_n$, in other words

$$\max_{1\le k\le 2^n} \|X_{k/2^n}(\omega) - X_{(k-1)/2^n}(\omega)\| \le 2^{-\gamma n}, \quad n \ge N_\omega. \tag{1.3.3}$$

Fix $\omega \in \Omega^*$. We will show the existence of a constant K', such that

$$\|X_t(\omega) - X_s(\omega)\| \le K'|t-s|^\gamma, \quad \forall s,t \in D, 0 < t-s < 2^{-N_\omega}. \tag{1.3.4}$$

Indeed, this implies uniform continuity of $X_t(\omega)$ for $t \in D$, for $\omega \in \Omega^*$. Step 2 will then be proved.

Let s,t satisfy $0 < t-s < 2^{-N_\omega}$. Hence, there exists $n \ge N_\omega$, such that $2^{-(n+1)} \le t-s < 2^{-n}$.

Fix $n \ge N_\omega$. For the moment, we restrict to the set of $s,t \in \cup_{m\ge n+1} D_m$, with $0 < t-s < 2^{-n}$. By induction to $m \ge n+1$ we will first show that

$$\|X_t(\omega) - X_s(\omega)\| \le 2\sum_{k=n+1}^m 2^{-\gamma k}, \tag{1.3.5}$$

if $s,t \in D_m$.

Suppose that $s,t \in D_{n+1}$. Then $t-s = 2^{-(n+1)}$. Thus s,t are neighbouring points in D_{n+1}, i.e. there exists $k \in \{0,\ldots,2^{n+1}-1\}$, such that $t = k/2^{n+1}$ and $s = (k+1)/2^{n+1}$. (1.3.5) with $m = n+1$ follows directly from (1.3.3). Assume that the claim holds true upto $m \ge n+1$. We will show its validity for $m+1$.

Put $s' = \min\{u \in D_m \,|\, u \ge s\}$ and $t' = \max\{u \in D_m \,|\, u \le t\}$. By construction $s \le s' \le t' \le t$, and $s'-s, t-t' \le 2^{-(m+1)}$. Then $0 < t'-s' \le t-s < 2^{-n}$. Since $s', t' \in D_m$, they satisfy the induction hypothesis. We may now apply the triangle inequality, (1.3.3) and the induction hypothesis to obtain

$$\begin{aligned}\|X_t(\omega) - X_s(\omega)\| &\le \|X_t(\omega) - X_{t'}(\omega)\| + \|X_{t'}(\omega) - X_{s'}(\omega)\| + \|X_{s'}(\omega) - X_s(\omega)\| \\ &\le 2^{-\gamma(m+1)} + 2\sum_{k=n+1}^m 2^{-\gamma k} + 2^{-\gamma(m+1)} = 2\sum_{k=n+1}^{m+1} 2^{-\gamma k}.\end{aligned}$$

This shows the validity of (1.3.5). We prove (1.3.4). To this end, let $s,t \in D$ with $0 < t-s < 2^{-N_\omega}$. As noted before, there exists $n > N_\omega$, such that $2^{-(n+1)} \le t-s < 2^{-n}$. Then there exists $m \ge n+1$ such that $t,s \in D_m$. Apply (1.3.5) to obtain

$$\|X_t(\omega) - X_s(\omega)\| \le 2\sum_{k=n+1}^m 2^{-\gamma k} \le \frac{2}{1-2^{-\gamma}} 2^{-\gamma(n+1)} \le \frac{2}{1-2^{-\gamma}}|t-s|^\gamma.$$

Consequently (1.3.4) holds with constant $K' = 2/(1-2^{-\gamma})$.
Step 3 Define a new stochastic process $Y = (Y_t)_{t\in[0,1]}$ on $(\Omega, \mathcal{F}, \mathsf{P})$ as follows: for $\omega \notin \Omega^*$, we put $Y_t = 0$ for all $t \in [0,1]$; for $\omega \in \Omega^*$ we define

$$Y_t(\omega) = \begin{cases} X_t(\omega), & \text{if } t \in D, \\ \lim\limits_{\substack{t_n \to t \\ t_n \in D}} X_{t_n}(\omega), & \text{if } t \notin D. \end{cases}$$

For each $\omega \in \Omega^*$, $t \to X_t(\omega)$ is uniformly continuous on the dense subset D of $[0,1]$. It is a theorem from Analysis, that $t \to X_t(\omega)$ can be uniquely extended as a continuous function on $[0,1]$. This is the function $t \to Y_t(\omega)$, $t \in [0,1]$.
Step 4 Uniform continuity of X implies that Y is a well-defined stochastic process. Since X is continuous in probability, it follows that Y is a modification of X (Exercise 1.4). See BN §5 for a useful characterisation of convergence in probability. QED

> The fact that Kolmogorov's continuity criterion requires $K|t-s|^{1+\beta}$ for some $\beta > 0$, guarantees uniform continuity of a.a. paths $X(\omega)$ when restricted to the dyadic rationals, whilst it does not so for $\beta = 0$ (see Exercise 1.5). This uniform continuity property is precisely the basis of the proof of the Criterion.

Corollary 1.3.4 *Brownian motion exists.*

Proof. By Corollary 1.2.4 there exists a process $W = (W_t)_{t\geq 0}$ that has properties (i,ii,iii) of Definition 1.1.3. By property (iii) the increment $W_t - W_s$ has a $\mathsf{N}(0, t-s)$-distribution for all $s \leq t$. This implies that $\mathsf{E}(W_t - W_s)^4 = (t-s)^2 \mathsf{E} Z^4$, with Z a standard normally distributed random variable. This means the Kolmogorov's continuity condition (1.3.1) is satisfied with $\alpha = 4$ and $\beta = 1$. So for every $T \geq 0$, there exists a continuous modification $W^T = (W_t^T)_{t\in[0,T]}$ of the process $(W_t)_{t\in[0,T]}$. Now define the process $X = (X_t)_{t\geq 0}$ by

$$X_t = \sum_{n=1}^{\infty} W_t^n \mathbf{1}_{\{[n-1,n)\}}(t).$$

In Exercise 1.7 you are asked to show that X is a Brownian motion process. QED

Remarks on the canonical process Lemma 1.3.5 below allows us to restrict to continuous paths. There are two possibilities now to define Brownian motion as a canonical stochastic process with everywhere continuous paths.

The first one is to 'kick out' the discontinuous paths from the underlying space. This is allowed by means of the outer measure.

Let $(\Omega, \mathcal{F}, \mathsf{P})$ be a probability space. Define the outer measure

$$\mathsf{P}^*\{A\} = \inf_{B \in \mathcal{F}, B \supset A} \mathsf{P}\{B\}.$$

Lemma 1.3.5 *Suppose that A is a subset of Ω with $\mathsf{P}^*\{A\} = 1$. Then for any $F \in \mathcal{F}$, one has $\mathsf{P}^*\{F\} = \mathsf{P}\{F\}$. Moreover, $(A, \mathcal{A}, \mathsf{P}^*)$ is a probability space, where $\mathcal{A} = \{A \cap F \mid F \in \mathcal{F}\}$.*

Kolmogorov's continuity criterion applied to canonical BM implies that the outer measure of the set $\mathcal{C}[0, \infty)$ of continuous paths equals 1. The BM process after modification is the canonical process on the restricted space $(\mathbf{R}^{[0,\infty)} \cap \mathcal{C}[0, \infty), \mathcal{B}^{[0,\infty)} \cap \mathcal{C}[0, \infty), \mathsf{P}^*)$, with P^* the outer measure associated with P.

The second possibility is the construction mentioned at the end of section 1.1, described in more generality below. Given any $(E, \mathcal{E})$-valued stochastic process X on an underlying probability space $(\Omega, \mathcal{F}, \mathsf{P})$. Then $X : (\Omega, \mathcal{F}) \to (E^T, \mathcal{E}^T)$ is a measurable map inducing a probability measure P_X on the path space $(E^T, \mathcal{E}^T)$. The canonical map on $(E^T, \mathcal{E}^T, \mathsf{P}_X)$ now has the same distribution as X by construction. Hence, we can always associate a canonical stochastic process with a given stochastic process.

Suppose now that there exists a subset $\Gamma \subset E^T$, such that $X : \Omega \to \Gamma \cap E^T$. That is, the paths of X have a certain structure. Then X is $\mathcal{F}/\Gamma \cap \mathcal{E}^T$-measurable, and induces a probability measure P_X on $(\Gamma, \Gamma \cap \mathcal{E}^T)$. Again, we may consider the canonical process on this restricted probability space $(\Gamma, \Gamma \cap \mathcal{E}^T, \mathsf{P}_X)$.

1.4 Gaussian processes

Brownian motion is an example of a so-called Gaussian process. The general definition is as follows.

Definition 1.4.1 A real-valued stochastic process is called *Gaussian* of all its fdd's are Gaussian, in other words, if they are multivariate normal distributions.

Let X be a Gaussian process indexed by the set T. Then $m(t) = \mathsf{E}X_t$, $t \in T$, is the *mean function* of the process. The function $r(s,t) = \mathsf{cov}(X_s, X_t)$, $(s,t) \in T \times T$, is the *covariance function*. By virtue of the following uniqueness lemma, fdd's of Gaussian processes are determined by their mean and covariance functions.

Lemma 1.4.2 *Two Gaussian processes with the same mean and covariance functions are versions of each other.*

Proof. See Exercise 1.8. QED

Brownian motion is a special case of a Gaussian process. In particular it has $m(t) = 0$ for all $t \geq 0$ and $r(s,t) = s \wedge t$, for all $s \leq t$. Any other Gaussian process with the same mean and covariance function has the same fdd's as BM itself. Hence, it has properties (i,ii,iii) of Definition 1.1.3. We have the following result.

Lemma 1.4.3 *A continuous or a.s. continuous Gaussian process* $X = (X_t)_{t\geq 0}$ *is a* BM *process if and only if it has the same mean function* $m(t) = \mathsf{E}X_t = 0$ *and covariance function* $r(s,t) = \mathsf{E}X_sX_t = s \wedge t$.

The lemma looks almost trivial, but provides us with a number extremely useful scaling and symmetry properties of BM!

Remark that a.s. continuity means that the collection of discontinuous paths is contained in a null-set. By continuity we mean that *all paths are continuous.*

Theorem 1.4.4 *Let* W *be a* BM *process on* $(\Omega, \mathcal{F}, \mathsf{P})$. *Then the following are* BM *processes as well:*

i) time-homogeneity *for every* $s \geq 0$ *the shifted process* $W(s) = (W_{t+s} - W_s)_{t\geq 0}$*;*

ii) symmetry *the process* $-W = (-W_t)_{t\geq 0}$*;*

iii) scaling *for every* $a > 0$, *the process* W^a *defined by* $W_t^a = a^{-1/2}W_{at}$*;*

iv) time inversion *the process* $X = (X_t)_{t\geq 0}$ *with* $X_0 = 0$ *and* $X_t = tW_{1/t}$, $t > 0$.

If W *has (a.s.) continuous paths then* $W(s)$, $-W$ *and* W^a *have (a.s.) continuous paths and* X *has a.s. continuous paths. There exists a set* $\Omega^* \in \mathcal{F}$, *such that* X *has continuous paths on* $(\Omega^*, \mathcal{F} \cap \Omega^*, \mathsf{P})$.

Proof. We would like to apply Lemma 1.4.3. To this end we have to check that (i) the defined processes are Gaussian; (ii) that (almost all) sample paths are continuous and (iii) that they have the same mean and covariance functions as BM. In Exercise 1.9 you are asked to show this for the processes in (i,ii,iii). We will give an outline of the proof (iv).

The most interesting step is to show that almost all sample paths of X are continuous. The remainder is analogous to the proofs of (i,ii,iii).

So let us show that almost all sample paths of X are continuous. By time inversion, it is immediate that $(X_t)_{t>0}$ (a.s.) has continuous sample paths if W has. We only need show a.s. continuity at $t = 0$, that is, we need to show that $\lim_{t\downarrow 0} X_t = 0$, a.s.

Let $\Omega^* = \{\omega \in \Omega \,|\, (W_t(\omega))_{t\geq 0} \text{ continuous}, W_0(\omega) = 0\}$. By assumption $\Omega \setminus \Omega^*$ is contained in a P-null set. Further, $(X_t)_{t>0}$ has continuous paths on Ω^*.

Then $\lim_{t\downarrow 0} X_t(\omega) = 0$ iff for all $\epsilon > 0$ there exists $\delta_\omega > 0$ such that $|X_t(\omega)| < \epsilon$ for all $t \leq \delta_\omega$. This is true if and only if for all integers $m \geq 1$, there exists an integer n_ω, such that $|X_q(\omega)| < 1/m$ for all $q \in \mathsf{Q}$ with $q < 1/n_\omega$, because of continuity of $X_t(\omega)$, $t > 0$. Check that this implies

$$\{\omega : \lim_{t\downarrow 0} X_t(\omega) = 0\} \cap \Omega^* = \bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{q\in(0,1/n]\cap\mathsf{Q}} \{\omega : |X_q(\omega)| < 1/m\} \cap \Omega^*.$$

The fdd's of X and W are equal. Hence (cf. Exercise 1.10)

$$\mathsf{P}\{\bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{q\in(0,1/n]\cap\mathsf{Q}} \{\omega : |X_q(\omega)| < 1/m\}\} = \mathsf{P}\{\bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{q\in(0,1/n]\cap\mathsf{Q}} \{\omega : |W_q(\omega)| < 1/m\}\}.$$

It follows that (cf. Exercise 1.10) the probability of the latter equals 1. As a consequence $\mathsf{P}\{\omega : \lim_{t\downarrow 0} X_t(\omega) = 0\} = 1$.

QED

These scaling and symmetry properties can be used to show a number of properties of Brownian motion. The first is that Brownian motion sample paths oscillate between $+\infty$ and $-\infty$.

Corollary 1.4.5 *Let W be a* BM *with the property that all paths are continuous. Then*

$$\mathsf{P}\{\sup_{t\geq 0} W_t = \infty, \inf_{t\geq 0} W_t = -\infty\} = 1.$$

Proof. It is sufficient to show that

$$\mathsf{P}\{\sup_{t\geq 0} W_t = \infty\} = 1. \tag{1.4.1}$$

Indeed, the symmetry property implies

$$\sup_{t\geq 0} W_t \stackrel{\mathsf{d}}{=} \sup_{t\geq 0}(-W_t) = -\inf_{t\geq 0} W_t.$$

Hence (1.4.1) implies that $\mathsf{P}\{\inf_{t\geq 0} W_t = -\infty\} = 1$. As a consequence, the probability of the intersection equals 1 (why?).

First of all, notice that $\sup_t W_t$ is well-defined. We need to show that $\sup_t W_t$ is a measurable function. This is true (cf. BN Lemma 1.3) if $\{\sup_t W_t \leq x\}$ is measurable for all $x \in \mathbf{R}$. (Q is sufficient of course).

This follows from

$$\{\sup_t W_t \leq x\} = \bigcap_{q\in\mathsf{Q}} \{W_q \leq x\}.$$

Here we use that all paths are continuous. We cannot make any assertions on measurability of $\{W_q \leq x\}$ restricted to the set of discontinuous paths, unless $\mathcal{F}$ is P-complete.

By the scaling property we have for all $a > 0$

$$\sup_t W_t \stackrel{\mathsf{d}}{=} \sup_t \frac{1}{\sqrt{a}} W_{at} = \frac{1}{\sqrt{a}} \sup_t W_t.$$

It follows for $n \in \mathbf{Z}_+$ that

$$\mathsf{P}\{\sup_t W_t \leq n\} = \mathsf{P}\{n^2 \sup_t W_t \leq n\} = \mathsf{P}\{\sup_t W_t \leq 1/n\}.$$

By letting n tend to infinity, we see that

$$\mathsf{P}\{\sup_t W_t < \infty\} = \mathsf{P}\{\sup_t W_t \leq 0\}.$$

Thus, for (1.4.1) it is sufficient to show that $\mathsf{P}\{\sup_t W_t \leq 0\} = 0$. We have

$$\begin{aligned}
\mathsf{P}\{\sup_t W_t \leq 0\} &\leq \mathsf{P}\{W_1 \leq 0, \sup_{t\geq 1} W_t \leq 0\} \\
&\leq \mathsf{P}\{W_1 \leq 0, \sup_{t\geq 1} W_t - W_1 < \infty\} \\
&= \mathsf{P}\{W_1 \leq 0\}\mathsf{P}\{\sup_{t\geq 1} W_t - W_1 < \infty\},
\end{aligned}$$

by the independence of Brownian motion increments. By the time-homogeneity of BM, the latter probability equals the probability that the supremum of BM is finite. We have just showed that this equals $\mathsf{P}\{\sup W_t \leq 0\}$. And so we find

$$\mathsf{P}\{\sup_t W_t \leq 0\} \leq \tfrac{1}{2}\mathsf{P}\{\sup_t W_t \leq 0\}.$$

This shows that $\mathsf{P}\{\sup_t W_t \leq 0\} = 0$ and so we have shown (1.4.1). QED

Since BM has a.s. continuous sample paths, this implies that almost every path visits every point of $\mathbf{R}$. This property is called *recurrence*. With probability 1 it even visits every point infinitely often. However, we will not further pursue this at the moment and merely mention the following statement.

Corollary 1.4.6 BM *is recurrent.*

An interesting consequence of the time inversion property is the following strong law of large numbers for BM.

Corollary 1.4.7 *Let W be a* BM. *Then*

$$\frac{W_t}{t} \stackrel{\text{a.s.}}{\to} 0, \quad t \to \infty.$$

Proof. Let X be as in part (iv) of Theorem 1.4.4. Then

$$\mathsf{P}\{\frac{W_t}{t} \to 0, \quad t \to \infty\} = \mathsf{P}\{X_{1/t} \to 0, \quad t \to \infty\} = 1.$$

QED

1.5 Non-differentiability of the Brownian sample paths

We have already seen that the sample paths of W are continuous functions that oscillate between $+\infty$ and $-\infty$. Figure 1.1 suggests that the sample paths are very rough. The following theorem shows that this is indeed the case.

Theorem 1.5.1 *Let W be a* BM *defined on the space $(\Omega, \mathcal{F}, \mathsf{P})$. There is a set $\Omega*$ with $\mathsf{P}\{\Omega^*\} = 1$, such that the sample path $t \to W_(\omega)$ is nowhere differentiable, for any $\omega \in \Omega^*$.*

Proof. Let W be a BM. Consider the upper and lower right-hand derivatives

$$\begin{aligned} D^W(t,\omega) &= \limsup_{h \downarrow 0} \frac{W_{t+h}(\omega) - W_t(\omega)}{h} \\ D_W(t,\omega) &= \liminf_{h \downarrow 0} \frac{W_{t+h}(\omega) - W_t(\omega)}{h}. \end{aligned}$$

Let

$$A = \{\omega \mid \text{ there exists } t \geq 0 \text{ such that } D^W(t,\omega) \text{ and } D_W(t,\omega) \text{ are finite } \}.$$

Note that A is not necessarily a measurable set. We will therefore show that A is contained in a measurable set B with $\mathsf{P}\{B\} = 0$. In other words, A has *outer measure* 0.

To define the set B, first consider for $k, n \in \mathbf{Z}_+$ the random variable

$$X_{n,k} = \max\left\{|W_{(k+1)/2^n} - W_{k/2^n}|, |W_{(k+2)/2^n} - W_{(k+1)/2^n}|, |W_{(k+3)/2^n} - W_{(k+2)/2^n}|\right\}.$$

Define for $n \in \mathbf{Z}_+$

$$Y_n = \min_{k \le n2^n} X_{n,k}.$$

A the set B we choose

$$B = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} \{Y_k \le k2^{-k}\}.$$

We claim that $A \subseteq B$ and $\mathsf{P}\{B\} = 0$.

To prove the inclusion, let $\omega \in A$. Then there exists $t = t_\omega$, such that $D_W(t,\omega)$, $D^W(t,\omega)$ are finite. Hence, there exists $K = K_\omega$, such that

$$-K < D_W(t,\omega) \le D^W(t,\omega) < K.$$

As a consequence, there exists $\delta = \delta_\omega$, such that

$$|W_s(\omega) - W_t(\omega)| \le K \cdot |s - t|, \quad s \in [t, t+\delta]. \tag{1.5.1}$$

Now take $n = n_\omega \in \mathbf{Z}_+$ so large that

$$\frac{4}{2^n} < \delta, \quad 8K < n, \quad t < n. \tag{1.5.2}$$

Next choose $k \in \mathbf{Z}_+$, such that

$$\frac{k-1}{2^n} \le t < \frac{k}{2^n}. \tag{1.5.3}$$

By the first relation in (1.5.2) we have that

$$\left|\frac{k+3}{2^n} - t\right| \le \left|\frac{k+3}{2^n} - \frac{k-1}{2^n}\right| \le \frac{4}{2^n} < \delta,$$

so that $k/2^n$, $(k+1)/2^n$, $(k+2)/2^n$, $(k+3)/2^n \in [t, t+\delta]$. By (1.5.1) and the second relation in (1.5.2) we have our choice of n and k that

$$\begin{aligned} X_{n,k}(\omega) &\le \max\big\{|W_{(k+1)/2^n} - W_t| + |W_t - W_{k/2^n}|, |W_{(k+2)/2^n} - W_t| + |W_t - W_{(k+1)/2^n}|, \\ &\qquad |W_{(k+3)/2^n} - W_t| + |W_t - W_{(k+2)/2^n}|\big\} \\ &\le 2K\frac{4}{2^n} < \frac{n}{2^n}. \end{aligned}$$

The third relation in (1.5.2) and (1.5.3) it holds that $k - 1 \le t2^n < n2^n$. This implies $k \le n2^n$ and so $Y_n(\omega) \le X_{n,k}(\omega) \le n/2^n$, for our choice of n.

Summarising, $\omega \in A$ implies that $Y_n(\omega) \le n/2^n$ for all sufficiently large n. This implies $\omega \in B$. We have proved that $A \subseteq B$.

In order to complete the proof, we have to show that $\mathsf{P}\{B\} = 0$. Note that $|W_{(k+1)/2^n} - W_{k/2^n}|$, $|W_{(k+2)/2^n} - W_{(k+1)/2^n}|$ and $|W_{(k+3)/2^n} - W_{(k+2)/2^n}|$ are i.i.d. random variables. We have for any $\epsilon > 0$ and $k = 0, \ldots, n2^n$ that

$$\begin{aligned} \mathsf{P}\{X_{n,k} \le \epsilon\} &\le \mathsf{P}\{|W_{(k+i)/2^n} - W_{(k+i-1)/2^n}| \le \epsilon, i = 1,2,3\} \\ &\le (\mathsf{P}\{|W_{(k+1)/2^n} - W_{(k+2)/2^n}| \le \epsilon\})^3 = (\mathsf{P}\{|W_{1/2^n}| \le \epsilon\})^3 \\ &= (\mathsf{P}\{|W_1| \le 2^{n/2}\epsilon\})^3 \le (2 \cdot 2^{n/2}\epsilon)^3 = 2^{3n/2+1}\epsilon^3. \end{aligned}$$

We have used time-homogeneity in the third step, the time-scaling property in the fourth and the fact that the density of a standard normal random variable is bounded by 1 in the last equality. Next,

$$\begin{aligned} \mathsf{P}\{Y_n \le \epsilon\} &= \mathsf{P}\{\cup_{k=1}^{n2^n}\{X_{n,l} > \epsilon, l = 0, \ldots, k-1, X_{n,k} \le \epsilon\}\} \\ &\le \sum_{k=1}^{n2^n} \mathsf{P}\{X_{n,k} \le \epsilon\} \le n2^n \cdot 2^{3n/2+1}\epsilon^3 = n2^{5n/2+1}\epsilon^3. \end{aligned}$$

Choose $\epsilon = n/2^n$, we see that $\mathsf{P}\{Y_n \le n/2^n\} \to 0$, as $n \to \infty$. This implies that $\mathsf{P}\{B\} = \mathsf{P}\{\liminf_{n\to\infty}\{Y_n \le n/2^n\}\} \le \liminf_{n\to\infty} \mathsf{P}\{Y_n \le n/2^n\} = 0$. We have used Fatou's lemma in the last inequality. QED

1.6 Filtrations and stopping times

If W is a BM, the increment $W_{t+h} - W_t$ is independent of 'what happened up to time t'. In this section we introduce the concept of a *filtration* to formalise the notion of 'information that we have up to time t'. The probability space $(\Omega, \mathcal{F}, \mathsf{P})$ is fixed again and we suppose that T is a subinterval of $\mathbf{Z}_+$ or $\mathbf{R}_+$.

Definition 1.6.1 A collection $(\mathcal{F}_t)_{t\in T}$ of sub-σ-algebras is called a *filtration* if $\mathcal{F}_s \subset \mathcal{F}_t$ for all $s \le t$. A stochastic process X defined on $(\Omega, \mathcal{F}, \mathsf{P})$ and indexed by T is called *adapted* to the filtration if for every $t \in T$, the random variable X_t is $\mathcal{F}_t$-measurable. Then $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in T}, \mathsf{P})$ is a *filtered probability space.*

We can think of a filtration as a flow of information. The σ-algebra $\mathcal{F}_t$ contains the events that can happen 'upto time t'. An adapted process is a process that 'does not look into the future'. If X is a stochastic process, then we can consider the filtration $(\mathcal{F}_t^X)_{t\in T}$ generated by X:

$$\mathcal{F}_t^X = \sigma(X_s, s \le t).$$

We call this *the filtration generated by* X, or the *natural filtration of* X. It is the 'smallest' filtration, to which X is adapted. Intuitively, the natural filtration of a process keeps track of the 'history' of the process. A stochastic process is always adapted to its natural filtration.

Canonical process and filtration If X is a canonical process on $(\Gamma, \Gamma \cap \mathcal{E}^T)$ with $\Gamma \subset E^T$, then $\mathcal{F}_t^X = \Gamma \cap \mathcal{E}^{[0,t]}$.

As has been pointed out in Section 1.3, with a stochastic process X one can associate a canonical process with the same distribution.

Indeed, suppose that $X : \Omega \to \Gamma \cap E^T$. The canonical process on $\Gamma \cap E^T$ is adapted to the filtration $(\mathcal{E}^{[0,t]\cap T} \cap \Gamma)_t$.

Review BN §2, the paragraph on σ-algebra generated by a random variable or a stochastic process.

If $(\mathcal{F}_t)_{t\in T}$ is a filtration, then for $t \in T$ we may define the σ-algebra

$$\mathcal{F}_{t+} = \bigcap_{n=1}^{\infty} \mathcal{F}_{t+1/n}.$$

This is the σ-algebra $\mathcal{F}_t$, augmented with the events that 'happen immediately after time t'. The collection $(\mathcal{F}_{t+})_{t\in T}$ is again a filtration (see Exercise 1.16). Cases in which it coincides with the original filtration are of special interest.

Definition 1.6.2 We call a filtration $(\mathcal{F}_t)_{t\in T}$ *right-continuous* if $\mathcal{F}_{t+} = \mathcal{F}_t$ for all $t \in T$.

Intuitively, right-continuity of a filtration means that 'nothing can happen in an infinitesimal small time-interval' after the observed time instant. Note that for every filtration $(\mathcal{F}_t)$, the corresponding filtration $(\mathcal{F}_{t+})$ is always right-continuous.

In addition to right-continuity it is often assumed that $\mathcal{F}_0$ contains all events in $\mathcal{F}_\infty$ that have probability 0, where

$$\mathcal{F}_\infty = \sigma(\mathcal{F}_t, t \geq 0).$$

As a consequence, every $\mathcal{F}_t$ then also contains these events.

Definition 1.6.3 A filtration $(\mathcal{F}_t)_{\in T}$ on a probability space $(\Omega, \mathcal{F}, \mathsf{P})$ is said to satisfy the *usual conditions* if it is right-continuous and $\mathcal{F}_0$ contains all P-negligible events of $\mathcal{F}_\infty$.

Stopping times We now introduce a very important class of 'random times' that can be associated with a filtration.

Definition 1.6.4 An $[0, \infty]$-valued random variable τ is called a *stopping time* with respect to the filtration $(\mathcal{F}_t)$ if for every $t \in T$ it holds that the event $\{\tau \leq t\}$ is $\mathcal{F}_t$-measurable. If $\tau < \infty$, we call τ a *finite* stopping time, and if $\mathsf{P}\{\tau < \infty\} = 1$, then we call τ a.s. finite. Similarly, if there exists a constant K such that $\tau(\omega) \leq K$, then τ is said to be *bounded*, and if $\mathsf{P}\{\tau \leq K\} = 1$ τ is *a.s. bounded*.

Loosely speaking, τ is a stopping time if for every $t \in T$ we can determine whether τ has occurred before time t on basis of the information that we have upto time t. Note that τ is $\mathcal{F}/\mathcal{B}([0, \infty])$-measurable.

With a stopping time τ we can associate the the σ-algebra σ^τ generated by τ. However, this σ-algebra only contains the information about *when* τ occurred. If τ is associated with an adapted process X, then σ^τ contains no further information on the history of the process upto the stopping time. For this reason we associate with τ the (generally) larger σ-algebra $\mathcal{F}_\tau$ defined by

$$\mathcal{F}_\tau = \{A \in \mathcal{F}_\infty : A \cap \{\tau \leq t\} \in \mathcal{F}_t \text{ for all } t \in T\}.$$

(see Exercise 1.17). This should be viewed as the collection of all events that happen prior to the stopping time τ. Note that the notation is unambiguous, since a deterministic time $t \in T$ is clearly a stopping time and its associated σ-algebra is simply the σ-algebra $\mathcal{F}_t$.

Our loose description of stopping times and stopped σ-algebra can made more rigorous, when we consider the canonical process with natural filtration.

Lemma 1.6.5 *Let $\Omega \subset E^T$ have the property that for each $t \in T$ and $\omega \in \Omega$ there exists $\overline{\omega} \in \Omega$ such that $\overline{\omega}_s = \omega_{s \wedge t}$ for all $s \in T$. Let $\mathcal{F} = \mathcal{E}^T \cap \Omega$. Let X be the canonical process on $(\Omega, \mathcal{F})$. Define $\mathcal{F}_\infty^X = \sigma(X_t, t \in T)$. The following assertions are true.*

a) *Let $A \subset \Omega$. Then $A \in \mathcal{F}_t^X$ if and only if (i) $A \in \mathcal{F}_\infty^X$ and (ii) $\omega \in A$ and $X_s(\omega) = X_s(\omega')$ for all $s \in T$ with $s \leq t$ imply that $\omega' \in A$.*

b) Galmarino test *τ is an $\mathcal{F}_t^X$-stopping time if and only τ is $\mathcal{F}_\infty^X$-measurable and $\tau(\omega) \leq t$, $X_s(\omega) = X_s(\omega')$ for $s \in T$, $s \leq t$ implies $\tau(\omega') \leq t$, for all $t \in T$, and $\omega, \omega' \in \Omega$.*

c) *Let τ be an $\mathcal{F}_t^X$-stopping time. Let $A \subset \Omega$. Then $A \in \mathcal{F}_\tau^X$ if and only if (i) $A \in \mathcal{F}_\infty^X$ and (ii) $\omega \in A$ and $X_s(\omega) = X_s(\omega')$ for all $s \leq \tau(\omega)$ imply that $\omega' \in A$.*

Proof. See Exercise 1.24. QED

If the filtration $(\mathcal{F}_t)$ is right-continuous, then τ is a stopping time if and only if $\{\tau < t\} \in \mathcal{F}_t$ for all $t \in T$ (see Exercise 1.23). The latter defines another type of random time called optional time.

Definition 1.6.6 A $[0, \infty]$-valued random variable τ is called an *optional time* with respect to the filtration $(\mathcal{F}_t)$ if for every $t \in T$ it holds that $\{\tau < t\} \in \mathcal{F}_t$. If $\tau < \infty$ almost surely, we call the optional time *finite*.

Check that τ is an optional time if and only if $\tau + t$ is a stopping time for each $t \in T$, $t > 0$.

Lemma 1.6.7 *τ is an $(\mathcal{F}_t)_t$-optional time if and only if it is a $(\mathcal{F}_{t+})_t$-stopping time with respect to $(\mathcal{F}_{t+})$. Every $(\mathcal{F}_t)_t$-stopping time is a $(\mathcal{F}_t)_t$-optional time.*

Proof. See Exercise 1.25. QED

The associated σ-algebra $\mathcal{F}_{\tau+}$ is defined to be

$$\mathcal{F}_{\tau+} = \{A \in \mathcal{F}_\infty : A \cap \{\tau \leq t\} \in \mathcal{F}_{t+}, t \geq 0\}.$$

For an $\{\mathcal{F}_t\}_t$-optional time τ it holds (cf. Exercise 1.34) that $\mathcal{F}_{\tau+}$ is a σ-algebra, with respect to which τ is measurable. Moreover $\mathcal{F}_{\tau+} = \{A \in \mathcal{F}_\infty : A \cap \{\tau < t\} \in \mathcal{F}_t, t \geq 0\}$.

Special stopping and optional times The so-called *hitting* and *first entrance times* form an important class of stopping times and optional times. They are related to the first time that the process visits a set B.

Definition 1.6.8 Let $X = (X_t)_{t \geq 0)}$ an $(E, \mathcal{E})$-valued stochastic process defined on the underlying probability space $(\Omega, \mathcal{F}, \mathit{Prob})$. Let $B \in \mathcal{E}$. The *first entrance time of B* is defined by

$$\sigma_B = \inf\{t \geq 0 \mid X_t \in B\}.$$

The *first hitting time* of B is defined by

$$\tau_B = \inf\{t > 0 \mid X_t \in B\}.$$

Lemma 1.6.9 *Let (E, d) be a metric space and let $\mathcal{B}(E)$ be the Borel-σ-algebra of open sets compatible with the metric d. Suppose that $X = (X_t)_{t \geq 0}$ is a continuous, $(E, \mathcal{B}(E))$-valued stochastic process and that B is closed in E. Then σ_B is an $(\mathcal{F}_t^X)$-stopping time.*[1]

[1] As is usual, we define $\inf \emptyset = \infty$.

Proof. Denote the distance of a point $x \in E$ to the set B by $d(x, B)$. In other words

$$d(x, B) = \inf\{d(x, y) \mid y \in B\}.$$

First note that $x \to d(x, B)$ is a continuous function. Hence it is $\mathcal{B}(E)$-measurable. It follows that $Y_t = d(X_t, B)$ is $(\mathcal{F}_t^X)$-measurable as a composition of measurable maps. Since X_t is continuous, the real-valued process $(Y_t)_t$ is continuous as well. Moreover, since B is closed, it holds that $X_t \in B$ if and only if $Y_t = 0$. By continuity of Y_t, it follows that $\sigma_B > t$ if and only if $Y_s > 0$ for all $s \le t$. This means that

$$\{\sigma_B > t\} = \{Y_s > 0, 0 \le s \le t\} = \bigcup_{n=1}^{\infty} \bigcap_{q \in \mathsf{Q}_t} \{Y_q > \tfrac{1}{n}\} = \bigcup_{n=1}^{\infty} \bigcap_{q \in \mathsf{Q}_t} \{d(X_q, B) > \tfrac{1}{n}\} \in \mathcal{F}_t^X,$$

where $\mathsf{Q}_t = \{tq \mid q \in \mathsf{Q} \cap [0, 1]\}$. QED

Lemma 1.6.10 *Let (E, d) be a metric space and let $\mathcal{B}(E)$ be the Borel-σ-algebra of open sets compatible with the metric d. Suppose that $X = (X_t)_{t \ge 0}$ is a right-continuous, $(E, \mathcal{B}(E))$-valued stochastic process and that B is an open set in E. Then, τ_B is an $(\mathcal{F}_t^X)$-optional time.*

Proof. By right-continuity of X and the fact that B is open, $\tau_B(\omega) < t$ if and only if there exists a rational number $0 < q_\omega < t$ such that $X_{q_\omega}(\omega) \in B$. Hence

$$\{\tau_B < t\} = \cup_{q \in (0, t \cap \mathsf{Q}} \{X_q \in B\}.$$

The latter set is $\mathcal{F}_t^X$-measurable, and so is the first. QED

Example 1.6.1 Let W be a BM with continuous paths and, for $x > 0$, consider the random variable

$$\tau_x = \inf\{t > 0 \mid W_t = x\}.$$

Since $x > 0$, W is continuous and $W_0 = 0$ a.s., τ_x can a.s. be written as

$$\tau_x = \inf\{t \ge 0 \mid W_t = x\}.$$

By Lemma 1.6.9 this is an $(\mathcal{F}_t^W)$-stopping time. Next we will show that $\mathsf{P}\{\tau_x < \infty\} = 1$.

Note that $\{\tau_x < \infty\} = \cup_{n=1}^{\infty} \{\tau_x \le n\}$ is a measurable set. Consider $A = \{\omega : \sup_{t \ge 0} W_t = \infty, \inf_{t \ge 0} W_t = -\infty\}$. By Corollary 1.4.5 this set has probability 1.

Let $T > |x|$. For each $\omega \in A$, there exist T_ω, T'_ω, such that $W_{T_\omega} \ge T$, $W_{T'_\omega} \le -T$. By continuity of paths, there exists $t_\omega \in (T_\omega \wedge T'_\omega, T_\omega \vee T'_\omega)$, such that $W_{t_\omega} = x$. It follows that $A \subset \{\tau_x < \infty\}$. Hence $\mathsf{P}\{\tau_x < \infty\} = 1$. QED

An important question is whether the first entrance time of a closed set is a stopping time for more general stochastic processes than the continuous ones. The answer in general is that this is not true unless the filtration is suitably augmented with null sets (cf. BN §10). Without augmentation we can derive the two following results. Define $X_{t_-} = \liminf_{s \uparrow t} X_s$.

Lemma 1.6.11 *Let (E,d) be a metric space and let $\mathcal{B}(E)$ be the Borel-σ-algebra of open sets compatible with the metric d. Suppose that $X=(X_t)_{t\geq 0}$ is a (everywhere) cadlag, $(E,\mathcal{B}(E))$-valued stochastic process and that B is a closed set in E. Then*

$$\gamma_B=\inf\{t>0\,|\,X_t\in B,\ \text{or } X_{t-}\in B\},$$

is an $(\mathcal{F}_t^X)$-stopping time.

Lemma 1.6.12 *Let (E,d) be a metric space and let $\mathcal{B}(E)$ be the Borel-σ-algebra of open sets compatible with the metric d. Suppose that $X=(X_t)_{t\geq 0}$ is a right-continuous, $(E,\mathcal{B}(E))$-valued stochastic process and that B is closed in $\mathcal{B}(E)$. Let X be defined on the underlying probability space $(\Omega,\mathcal{F},\mathsf{P})$.*

Suppose further that there exist $\mathcal{F}$-measurable random times $0<\tau_1<\tau_2<\cdots$, such that the discontinuities of X are contained in the set $\{\tau_1,\tau_2,\ldots\}$. Then

a) *$\tau_1,\tau_2,\ldots$ are $(\mathcal{F}_t^X)$-stopping times;*

b) *σ_B is an $(\mathcal{F}_t^X)$-stopping time.*

Proof. We prove (a) and check that $\{\tau_1\leq t\}\in\mathcal{F}_t^X$. Define for $Q_t=\{qt\,|\,q\in[0,1]\cap\mathsf{Q}\}$

$$G=\bigcup_{m\geq 1}\bigcap_{n\geq 1}\bigcup_{u,s\in\mathsf{Q}_t}\{|u-s|\leq\frac{1}{n},d(X_u,X_s)>\frac{1}{m}\}.$$

Claim: $G=\{\tau_1\leq t\}$.
Let $\omega\in G$. Then there exists m_ω such that for each n there exists a pair $(u_{n,\omega},s_{n,\omega})$ with $|u_{n,\omega}-s_{n,\omega}|<1/n$ for which $d(X_{u_{n,\omega}},X_{s_{n,\omega}})>1/m_\omega$.

If $\omega\notin\{\tau_1\leq t\}$, then $\tau_1(\omega)>t$, and $s\to X_s(\omega)$ would be continuous on $[0,t]$, hence uniformly continuous. As a consequence, for each for m_ω there exists n_ω for which $d(X_s,X_u)<1/m_\omega$ for $|u-s|<1/n_\omega$. This contradicts the above, and hence $\omega\in\{\tau_1\leq t\}$.

To prove the converse, assume that $\omega\in\{\tau_1\leq t\}$, i.o.w. $s=\tau_1(\omega)\leq t$. By right-continuity this implies that there exists a sequence $t_l\uparrow s$, along which $X_{t_l}(\omega)\not\to X_s(\omega)$. Hence there exists m, such that for each n there exists $t_{l(n)}$ with $|s-t_{l(n)}|<1/n$, for which $d(X_{t_{l(n)}}(\omega),X_s(\omega))>1/m$.

By right-continuity, for each n one can find $q_n\geq t_{l(n)}$ and $q\geq s$, $q_n,q\in\mathsf{Q}_t$, such that $|q-q_n|<1/n$ and $d(X_{q(n)}(\omega),X_q(\omega))>1/2m$. It follows that

$$\omega\in\bigcap_{n\geq 1}\cup_{q,q(n)\in\mathsf{Q}_t}\{|q-q(n)|<1/n,d(X_{q(n)},X_q)>1/2m\}.$$

Hence $\omega\in G$.

To show that τ_2 is a stopping time, we add the requirement $u,s\geq\tau_1$ in the definition of the analogon of the set G, etc.

Next we prove (b). We will consider only the case that $\{\tau_k(\omega)\}_k$ does not have an accumulation point in $[0,t]$ for any $\omega\in\Omega$. Figure out yourself what to do in the case of an accumulation point! We have to cut out small intervals to the left of jumps. On the remainder we can separate the path $X_s(\omega)$, $s\leq t$ and the set B, if $\sigma_B(\omega)>t$. To this end define

$$I_{k,n}=\{u\,|\,\tau_k-\frac{1}{n}\leq u<\tau_k\leq t\}.$$

For each $\omega \in \Omega$ this is a subset of $[0,t]$. Now, given $u \in [0,t]$,

$$\{\omega \mid I_{k,n}(\omega) \ni u\} = \{\omega \mid u < \tau_k(\omega) \le t \wedge (u + \frac{1}{n})\} \in \mathcal{F}_t^X.$$

Check that

$$\{\sigma_B > t\} = \bigcap_{n\ge1} \bigcup_{m\ge1} \bigcap_{q\in\mathbf{Q}_t} [\bigcap_k \{q \notin I_{k,n}\} \cap \{d(X_q, B) > \frac{1}{m}\}].$$

QED

Measurability of X_τ for τ an adapted stopping time We often would like to consider the stochastic process X evaluated at a finite stopping time τ. However, it is not a priori clear that the map $\omega \to X_{\tau(\omega)}(\omega)$ is measurable. In other words, that X_τ is a random variable. We need measurability of X in both parameters t and ω. This motivates the following definition.

Definition 1.6.13 An $(E, \mathcal{E})$-valued stochastic process is called *progressively measurable* with respect to the filtration $(\mathcal{F}_t)$ if for every $t \in T$ the map $(s, \omega) \to X_s(\omega)$ is measurable as a map from $([0,t] \times \Omega, \mathcal{B}([0,t]) \times \mathcal{F}_t)$ to $(E, \mathcal{E})$.

Lemma 1.6.14 *Let (E,d) be a metric space and $\mathcal{B}(E)$ the Borel-σ-algebra of open sets compatible with d. Every adapted right-continuous, $(E, \mathcal{B}(E))$-valued stochastic process X is progressively measurable.*

Proof. Fix $t \ge 0$. We have to check that

$$\{(s,\omega) \mid X_s(\omega) \in A, s \le t\} \in \mathcal{B}([0,t]) \times \mathcal{F}_t, \quad \forall A \in \mathcal{B}(E).$$

For $n \in \mathbf{Z}_+$ define the process

$$X_s^n = \sum_{k=0}^{n-1} X_{(k+1)t/n} \mathbf{1}_{\{(kt/n,(k+1)t/n]\}}(s) + X_0 \mathbf{1}_{\{0\}}(s).$$

This is a measurable process, since

$$\{(s,\omega) \mid X_s^n(\omega) \in A, s \le t\} = \bigcup_{k=0}^{n-1} (\{s \in (kt/n, (k+1)t/n]\} \times \{\omega \mid X_{(k+1)t/n}(\omega) \in A\}) \bigcup (\{0\} \times \{\omega \mid X_0(\omega) \in A\}).$$

Clearly, $X_s^n(\omega) \to X_s(\omega)$, $n \to \infty$, for all $(s,\omega) \in [0,t] \times \Omega$, pointwise. By BN Lemma 6.1, the limit is measurable. QED

> Review BN §6 containing an example of a non-progressively measurable stochastic process and a stopping time τ with X_τ not $\mathcal{F}_\tau$-measurable.

Lemma 1.6.15 *Suppose that X is a progressively measurable process. Let τ be a stopping time. Then $\mathbf{1}_{\{\tau<\infty\}} X_\tau$ is an $\mathcal{F}_\tau$-measurable random element.*

Proof. We have to show that $\{\mathbf{1}_{\{\tau<\infty\}} X_\tau \in B\} \cap \{\tau \leq t\} = \{X_\tau \in B\} \cap \{\tau \leq t\} \in \mathcal{F}_t$, for every $B \in \mathcal{E}$ and every $t \geq 0$. Now note that

$$\{X_{\tau\wedge t} \in B\} = \{X_\tau \in B, \tau \leq t\} \cup \{X_t \in B, \tau > t\}.$$

Clearly $\{X_t \in B, \tau > t\} \in \mathcal{F}_t$. If we can show that $\{X_{\tau\wedge t} \in B\} \in \mathcal{F}_t$, it easily follows that $\{X_\tau \in B, \tau \leq t\} \in \mathcal{F}_t$. Hence, it suffices to show that the map $\omega \to X_{\tau(\omega)\wedge t}(\omega)$ is $\mathcal{F}_t$-measurable.

To this end consider the map $\phi : ([0,t]\times\Omega, \mathcal{B}([0,t])\times\mathcal{F}_t)) \to ([0,t]\times\Omega, \mathcal{B}([0,t])\times\mathcal{F}_t))$ given by $\phi(s,\omega) = (\tau(\omega)\wedge s, \omega)$ is measurable (this is almost trivial, see Exercise 1.26). Using that X is progressively measurable, it follows that the composition map $(s,\omega) \to X(\phi(s,\omega)) = X_{\tau(\omega)\wedge s}(\omega)$ is measurable.

By Fubini's theorem the section map $\omega \to X_{\tau(\omega)\wedge t}(\omega)$ is $\mathcal{F}_t$-measurable. QED

Very often problems of interest consider a stochastic process upto a given stopping time τ. To this end we define the *stopped process* X^τ by

$$X_t^\tau = X_{\tau\wedge t} = \begin{cases} X_t, & t < \tau, \\ X_\tau, & t \geq \tau. \end{cases}$$

Using Lemma 1.6.15, and the arguments in the proof, as well as Exercises 1.18 and 1.20, we have the following result.

Lemma 1.6.16 *Let X be progressively measurable with respect to $(\mathcal{F}_t)$ and τ an $(\mathcal{F}_t)$-stopping time. The following are true.*

1. *The stopped pocess X^τ is progressively measurable with respect to the filtrations $(\mathcal{F}_{\tau\wedge t})_t$ and $(\mathcal{F}_t)_t$.*

2. *$(\mathcal{F}_{\tau+t})_t$ is a filtration, and the shifted process $\{X_{\tau+t}\}_t$ is progressively measurable with respect to this shifted filtration.*

The proof is left as an exercise. The Lemma is important lateron, when we consider processes shifted by a stopping time.

In the subsequent chapters we repeatedly need the following technical lemma. It states that every stopping time is the decreasing limit of a sequence of stopping times that take only countably many values.

Lemma 1.6.17 *Let τ be a stopping time. Then there exist stopping times τ_n that only take either finitely or countably many values and such $\tau_n \downarrow \tau$.*

Proof. Define

$$\tau_n = \sum_{k=1}^{n2^n-1} \frac{k}{2^n} \mathbf{1}_{\{\tau\in[(k-1)/2^n, k/2^n)\}} + \infty \mathbf{1}_{\{\tau>n\}},$$

to obtain an approximating stopping time taking only finitely many values. Or

$$\tau_n = \sum_{k=1}^{\infty} \frac{k}{2^n} \mathbf{1}_{\{[(k-1)/2^n, k/2^n)\}} + \infty \cdot \mathbf{1}_{\{\tau=\infty\}},$$

for an approximating stopping time taking countable many values. Then τ_n is a stopping time and $\tau_n \downarrow \tau$ (see Exercise 1.27). QED

Optional times Similarly, for a progressively measurable process X, and τ an adapted optional time, it holds that X_τ is $\mathcal{F}_{\tau+}$-measurable. This follows directly from Lemma 1.6.7. We can further approximate τ by a non-decreasing sequence of stopping times: take the sequence from the preceding lemma. Then $\tau_n \geq \tau_{n+1} \geq \tau$ for $n \geq 1$. By virtue of the preceding lemma τ_n is an $(\mathcal{F}_t)_t$-stopping time for all n with $\lim_{n\to\infty} \tau_n = \tau$. Moreover, for all $A \in \mathcal{F}_{\tau+}$ it holds that $A \cap \{\tau_n = k/2^n\} \in \mathcal{F}_{k/2^n}$, $n, k \geq 1$. Furthermore, if σ is a $(\mathcal{F}_t)_t$-stopping time with $\tau \leq \sigma$ and $\tau < \sigma$ on $\{\tau < \infty\}$, then $\mathcal{F}_{\tau+} \subset \mathcal{F}_\sigma$ (cf. Exercise 1.35).

Finally... Using the notion of filtrations, we can extend the definition of BM as follows.

Definition 1.6.18 Suppose that on a probabillity space $(\Omega, \mathcal{F}, \mathsf{P})$ we have a filtration $(\mathcal{F}_t)_{t\geq 0}$ and an adapted stochastic process $W = (W_t)_{t\geq 0}$. Then W is called a *(standard) Brownian motion* (or a *Wiener process*) *with respect to the filtration* $(\mathcal{F}_t)_t$ if

i) $W_0 = 0$;

ii) (independence of increments) $W_t - W_s$ is independent of $\mathcal{F}_s$ for all $s \leq t$;

iii) (stationarity of increments) $W_t - W_s \stackrel{\mathsf{d}}{=} \mathcal{N}(0, t-s)$ distribution;

iv) all sample paths of W are continuous.

Clearly, process W that is a BM in the sense of the 'old' Definition 1.1.2 is a BM with respect to its natural filtration. If in the sequel we do not mention the filtration of a BM explicitly, we mean the natural filtration. However, we will see that it is sometimes necessary to consider Brownian motions with larger filtrations as well.

1.7 Exercises

Exercise 1.1 Show the claim in the proof of Theorem 1.1.6 that the system $\mathcal{I}$ described there is a π-system for the σ-algebra $\sigma(N_u, u \le s)$.

Exercise 1.2 Complete the proof of Corollary 1.2.4. Give full details.

Exercise 1.3 Give an example of two processes that are versions of each other, but not modifications.

Exercise 1.4 Prove that the process Y defined in the proof of Theorem 1.3.3 is indeed a modification of the process X. See remark in **Step 4** of the proof of this theorem.

Exercise 1.5 An example of a right-continuous but not continuous stochastic process X is the following. Let $\Omega = [0,1]$, $\mathcal{F} = \mathcal{B}([0,1])$ and $\mathsf{P} = \lambda$ is the Lebesgue measure on $[0,1]$. Let Y be the identity map on Ω, i.e. $Y(\omega) = \omega$. Define a stochastic process $X = (X_t)_{t\in[0,1]}$ by $X_t = \mathbf{1}_{\{Y\le t\}}$. Hence, $X_t(\omega) = \mathbf{1}_{\{Y(\omega)\le t\}} = \mathbf{1}_{\{[0,t]\}}(\omega)$.

The process X does not satisfy the conditions of Kolmogorov's Continuity Criterion, but it does satisfy the condition

$$\mathsf{E}|X_t - X_s|^\alpha \le K|t-s|,$$

for any $\alpha > 0$ and $K = 1$. Show this.

Prove that X has no continuous modification. Hint: suppose that X has a continuous modification, X' say. Enumerate the elements of $\mathsf{Q} \cap [0,1]$ by $q_1, q_2, \ldots$. Define $\Omega_n = \{\omega : X_{q_n}(\omega) = X'_{q_n}(\omega)\}$. Let $\Omega^* = \cap_{n\ge1}\Omega_n$. Show that $\mathsf{P}\{\Omega^*\} = 1$. Then conclude that a continuous modification cannot exist.

Exercise 1.6 Suppose that X and Y are modifications of each other with values in a Polish space $(E, \mathcal{E})$, and for both X and Y all sample paths are either left or right continuous. Let T be an interval in $\mathbf{R}$. Show that

$$\mathsf{P}\{X_t = Y_t, \text{ for all } t \in T\} = 1.$$

Exercise 1.7 Prove that the process X in the proof of Corollary 1.3.4 is a BM process. To this end, you have to show that X has the correct fdd's, and that X has a.s. continuous sample paths.

Exercise 1.8 Prove Lemma 1.4.2.

Exercise 1.9 Prove parts (i,ii,iii) of Theorem 1.4.4.

Exercise 1.10 Consider the proof of the time-inversion property of Theorem 1.4.4. Prove that

$$\mathsf{P}\{\bigcap_{m=1}^\infty \bigcup_{n=1}^\infty \bigcap_{q\in(0,1/n]\cap\mathsf{Q}} \{\omega : |X_q(\omega)| < 1/m\}\} = \mathsf{P}\{\bigcap_{m=1}^\infty \bigcup_{n=1}^\infty \bigcap_{q\in(0,1/n]\cap\mathsf{Q}} \{\omega : |W_q(\omega)| < 1/m\}\} = 1.$$

Exercise 1.11 Let W be a BM and define $X_t = W_{1-t} - W_1$ for $t \in [0,1]$. Show that $(X_t)_{t\in[0,1]}$ is a BM as well.

Exercise 1.12 Let W be a BM and fix $t > 0$. Define the process B by

$$B_s = W_{s\wedge t} - (W_s - W_{s\wedge t}) = \begin{cases} W_s, & s \le t \\ 2W_t - W_s, & s > t. \end{cases}$$

Draw a picture of the processes W and B and show that B is again a BM. We will see another version of this so-called *reflection principle* in Chapter 3.

Exercise 1.13 **i)** Let W be a BM and define the process $X_t = W_t - tW_1$, $t \in [0,1]$. Determine the mean and covariance functions of X.

ii) The process X of part (i) is called the (standard) *Brownian bridge* on $[0,1]$, and so is every other continuous Gaussian process indexed by the interval $[0,1]$ that has the same mean and covariance function. Show that the processes Y and Z defined by $Y_t = (1-t)W_{t/(1-t)}$, $t \in [0,1)$, and $Y_1 = 0$ and $Z_0 = 0$, $Z_t = tW_{(1/t)-1}$, $t \in (0,1]$ are standard Brownian bridges.

Exercise 1.14 Let $H \in (0,1)$ be given. A continuous, zero-mean Gaussian process X with covariance function $2\mathsf{E}X_sX_t = (t^{2H} + s^{2H} - |t-s|^{2H})$ is called a *fractional Brownian motion (f*BM*) with Hurst index H*. Show that the fBM with Hurst index 1/2 is simply the BM. Show that if X is a fBM with Hurst index H, then for all $a > 0$ the process $a^{-H}X_{at}$ is a fBM with Hurst index H as well.

Exercise 1.15 Let W be a Brownian motion and fix $t > 0$. For $n \in \mathbf{Z}_+$, let π_n be a partition of $[0,t]$ given by $0 = t_0^n < t_1^n < \cdots < t_{k_n}^n = t$ and suppose that the mesh $\|\pi_n\| = \max_k |t_k^n - t_{k-1}^n$ tends to zero as $n \to \infty$. Show that

$$\sum_k (W_{t_k^n} - W_{t_{k-1}^n})^2 \xrightarrow{L^2} t,$$

as $n \to \infty$. Hint: show that the expectation of the sum tends to t and the variance to 0.

Exercise 1.16 Show that if $(\mathcal{F}_t)$ is a filtration, then $(\mathcal{F}_{t+})$ is a filtration as well.

Exercise 1.17 Prove that the collection $\mathcal{F}_\tau$ associated with a stopping time τ is a σ-algebra.

Exercise 1.18 Show that if σ, τ are stopping times with $\sigma \le \tau$, then $\mathcal{F}_\sigma \subseteq \mathcal{F}_\tau$.

Exercise 1.19 Let σ and τ be two $(\mathcal{F}_t)$-stopping times. Show that $\{\sigma \le \tau\} \subset \mathcal{F}_\sigma \cap \mathcal{F}_\tau$.

Exercise 1.20 If σ and τ are stopping times w.r.t. the filtration $(\mathcal{F}_t)$, show that $\sigma \wedge \tau$ and $\sigma \vee \tau$ are stopping times as well. Determine the associated σ-algebras. **Hint:** show that $A \in \mathcal{F}_{\sigma\vee\tau}$ implies $A \cap \{\sigma \le \tau\} \in \mathcal{F}_\tau$.

Exercise 1.21 If σ and τ are stopping times w.r.t. the filtration $(\mathcal{F}_t)$, show that $\sigma + \tau$ is a stopping time as well. Hint: for $t > 0$ write

$$\{\sigma + \tau > t\} = \{\tau = 0, \sigma > t\} \cup \{0 < \tau < t, \sigma + \tau > t\} \cup \{\tau > t, \sigma = 0\} \cup \{\tau \ge t, \sigma > 0\}.$$

Only for the second event on the right-hand side it is non-trivial to prove that it belongs to $\mathcal{F}_t$. Now observe that if $\tau > 0$, then $\sigma + \tau > t$ if and only if there exists a positive $q \in \mathsf{Q}$, such that $q < \tau$ and $\sigma + q > t$.

Exercise 1.22 Show that if σ and τ are stopping times w.r.t. the filtration $(\mathcal{F}_t)$ and X is an integrable random variable, then $\mathbf{1}_{\{\tau=\sigma\}}\mathsf{E}(X \mid \mathcal{F}_\tau) \overset{\text{a.s.}}{=} \mathbf{1}_{\{\tau=\sigma\}}\mathsf{E}(X \mid \mathcal{F}_\sigma)$. Hint: show that $\mathbf{1}_{\{\tau=\sigma\}}\mathsf{E}(X \mid \mathcal{F}_\tau) = \mathbf{1}_{\{\tau=\sigma\}}\mathsf{E}(X \mid \mathcal{F}_\tau \cap \mathcal{F}_\sigma)$.

Exercise 1.23 Show that if the filtration $(\mathcal{F}_t)$ is right-continuous, then τ is an $(\mathcal{F}_t)$-stopping time if and only if $\{\tau < t\} \in \mathcal{F}_t$ for all $t \in T$.

Exercise 1.24 Prove Lemma 1.6.5. Hint for (a): use BN Lemma 3.10.

Exercise 1.25 Prove Lemma 1.6.7.

Exercise 1.26 Show that the map $\omega \to (\tau(\omega) \wedge t, \omega)$ in the proof of Lemma 1.6.15 is measurable as a map from $(\Omega, \mathcal{F}_t)$ to $([0,t] \times \Omega, \mathcal{B}([0,t]) \times \mathcal{F}_t))$.

Exercise 1.27 Show that τ_n in the proof of Lemma 1.6.17 are indeed stopping times and that they converge to τ.

Exercise 1.28 Translate the definitions of §1.6 to the special case that time is discrete, i.e. $T = \mathbf{Z}_+$.

Exercise 1.29 Let W be a BM and let $Z = \{t \geq 0 \mid W_t = 0\}$ be its zero set. Show that with probability 1 the set Z has Lebesgue measure 0, is closed and unbounded.

Exercise 1.30 We define the last exit time of x:

$$L_x = \sup\{t > 0 \,:\, W_t = x\},$$

where $\sup\{\emptyset\} = 0$.

i) Show that τ_0 is measurable (τ_0 is defined in Definition 1.6.8).

ii) Show that L_x is measurable for all x. Derive first that $\{L_x < t\} = \cap_{n>t}\{|W_s - x| > 0, t \leq s \leq n\}$. I had mistakenly and in a hurry changed it, but it was correct as it was!!!.

iii) Show that $L_x = \infty$ a.s. for all x, by considering the set $\{\sup_{t\geq 0} W_t = \infty, \inf_{t\geq 0} W_t = -\infty\}$ as in the proof of Example 1.6.1.

iv) Show that for almost all $\omega \in \Omega$ there exists a strictly decreasing sequence $\{t_n(\omega)\}_n$, $\lim_n t_n(\omega) = 0$, such that $W(t_n)(\omega) = 0$ for all n. Hint: time-inversion + (iii). Hence $t = 0$ is a.s. an accumulation point of zeroes of W and so $\tau_0 = 0$ a.s.

Exercise 1.31 Consider Brownian motion W with continuous paths, defined on a probability space $(\Omega, \mathcal{F}, \mathsf{P})$. Let $Z(\omega) = \{t \geq 0, W_t(\omega) = 0\}$ be its zero set. In Problem 1.30 you have been asked to show that $t = 0$ is an accumulation point of $Z(\omega)$ for almost all $\omega \in \Omega$.

Let λ denote the Lebesgue measure on $[0, \infty)$. Show (by interchanging the order of integration) that

$$\int_\Omega \lambda(Z(\omega))d\mathsf{P}(\omega) = 0,$$

and argue from this that Z a.s. has Lebesgue measure 0, i.e. $\lambda(Z(\omega)) = 0$ for a.a. $\omega \in \Omega$.

Exercise 1.32 Let W be a BM (with continuous paths) with respect to its natural filtration $(\mathcal{F}_t^W)_t$. Define for $a > 0$

$$S_a = \inf\{t \geq 0 : W_t > a\}.$$

i) Is S_a an optional time? Justify your answer.

Let now $\sigma_a = \inf\{t \geq 0 : W_t = a\}$ be the first entrance time of a and let

$$M_a = \sup\{t \geq 0 : W_t = at\},$$

be the last time that W_t equals at.

ii) Is M_a a stopping time? Justify your answer. Show that $M_a < \infty$ with probability 1 (you could use time-inversion for BM).

iii) Show that M_a has the same distribution as $1/\sigma_a$.

Exercise 1.33 Let $X = (X_t)_{t \geq 0}$ be a Gaussian, zero-mean stochastic process starting from 0, i.e. $X_0 = 0$. Moreover, assume that the process has *stationary increments*, meaning that for all $t_1 \geq s_1, t_2 \geq s_2, \ldots, t_n \geq s_n$, the distribution of the vector $(X_{t_1} - X_{s_1}, \ldots, X_{t_n} - X_{s_n})$ only depends on the time points through the differences $t_1 - s_1, \ldots, t_n - s_n$.

a) Show that for all $s, t \geq 0$

$$\mathsf{E} X_s X_t = \tfrac{1}{2}(v(s) + v(t) - v(|t - s|)),$$

where the function v is given by $v(t) = \mathsf{E} X_t^2$.

In addition to stationarity of the increments we now assume that X is H*-self similar* for some parameter $H > 0$. Recall that this means that for every $a > 0$, the process $(X_{at})_t$ has the same finite dimensional distributions as $(a^H X_t)_t$.

b) Show that the variance function $v(t) = \mathsf{E} X_t^2$ must be of the form $v(t) = Ct^{2H}$ for some constant $C \geq 0$.

In view of the (a,b) we now assume that X is a zero-mean Gaussian process with covariance function

$$\mathsf{E} X_s X_t = \tfrac{1}{2}(s^{2H} + t^{2H} - |t - s|^{2H}),$$

or some $H > 0$.

c) Show that we must have $H \leq 1$. (Hint: you may use that by Cauchy-Schwartz, the (semi-)metric $d(s,t) = \sqrt{\mathsf{E}(X_s - X_t)^2}$ on $[0, \infty)$ satisfies the triangle inequality).

d) Show that for $H = 1$, we have $X_t = tZ$ a.s., for a standard normal random variable Z not depending on t.

e) Show that for every value of the parameter $H \in (0, 1]$, the process X has a continuous modification.

Exercise 1.34 Let τ be an $\{\mathcal{F}_t\}_t$-optional time. Show that $\mathcal{F}_{\tau^+}$ is a σ-algebra, with respect to which τ is measurable. Show $\mathcal{F}_{\tau^+} = \{A \in \mathcal{F} : A \cap \{\tau < t\} \in \mathcal{F}_t, t \geq 0\}$. Let σ, τ be $\{\mathcal{F}\}_t$-optional times. Prove that $\mathcal{F}_{(\sigma \wedge \tau)^+} = \mathcal{F}_{\sigma^+} \cap \mathcal{F}_{\tau^+}$. If additionally $\sigma \leq \tau$ everywhere, then finally show that $\mathcal{F}_{\sigma^+} \subset \mathcal{F}_{\tau^+}$.

Exercise 1.35 Show the validity of the assertion in the paragraph on optional times below Lemma 1.6.17.

Chapter 2

Martingales

2.1 Definition and examples

In this chapter we introduce and study a very important class of stochastic processes: the so-called martingales. Martingales arise naturally in many branches of the theory of stochastic processes. In particular, they are very helpful tools in the study of BM. In this section, the index set T is an arbitrary interval of $\mathbf{Z}_+$ and $\mathbf{R}_+$.

Definition 2.1.1 An $(\mathcal{F}_t)$-adapted, real-valued process M is called a *martingale* (with respect to the filtration $(\mathcal{F}_t)$) if

i) $\mathsf{E}|M_t| < \infty$ for all $t \in T$;

ii) $\mathsf{E}(M_t \,|\, \mathcal{F}_s) \stackrel{\text{a.s.}}{=} M_s$ for all $s \le t$.

If property (ii) holds with '$\ge$' (resp. '$\le$') instead of '=', then M is called a *submartingale* (resp. *supermartingale*).

Intuitively, a martingale is a process that is 'constant on average'. Given all information up to time s, the best guess for the value of the process at time $t \ge s$ is smply the current value M_s. In particular, property (ii) implies that $\mathsf{E}M_t = \mathsf{E}M_0$ for all $t \in T$. Likewise, a submartingale is a process that increases on average, and a supermartingale decreases on average. Clearly, M is a submartingale if and only if $-M$ is a supermartingale and M is a martingale if it is both a submartingale and a supermartingale. The basic properties of conditional expectations give us the following result and examples.

> Review BN §7 Conditional expectations.
> **N.B.** Let $T = \mathbf{Z}_+$. The tower property implies that (sub-, super-)martingale property (ii) is implied by (ii') $\mathsf{E}\{M_{n+1} \,|\, \mathcal{F}_n\} = M_n(\ge, \le)$ a.s. for $n \in \mathbf{Z}_+$.

Example 2.1.1 Let X_n, $n = 1, \ldots$, be a sequence of i.i.d. real-valued integrable random variables. Take e.g. the filtration $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$. Then $M_n = \sum_{k=1}^n X_k$ is a martingale if $\mathsf{E}X_1 = 0$, a submartingale if $\mathsf{E}X_1 > 0$ and a supermartingale if $\mathsf{E}X_1 < 0$. The process $M = (M_n)_n$ can be viewed as a random walk on the real line.

If $\mathsf{E}X_1 = 0$, but X_1 is square integrable, $M'_n = M_n^2 - n\mathsf{E}X_1^2$ is a martingale.

Example 2.1.2 (Doob martingale) Suppose that X is an integrable random variable and $(\mathcal{F}_t)_{t\in T}$ a filtration. For $t \in T$, define $M_t = \mathsf{E}(X \mid \mathcal{F}_t)$, or, more precisely, let M_t be a version of $\mathsf{E}(X \mid \mathcal{F}_t)$. Then $M = (M_t)_{t\in T}$ is an $(\mathcal{F}_t)$-martingale and M is uniformly integrable (see Exercise 2.1).

Review BN §8 Uniform integrability.

Example 2.1.3 Suppose that M is a martingale and that ϕ is a convex function such that $\mathsf{E}|\phi(M_t)| < \infty$ for $t \in T$. Then the process $\phi(M)$ is a submartingale. The same is true if M is a submartingale and ϕ is an increasing, convex function (see Exercise 2.2).

BM generates many examples of martingales. The most important ones are given in the following example.

Example 2.1.4 Let W be a BM. Then the following processes are martingales with respect to the same filtration:

i) W itself;

ii) $W_t^2 - t$;

iii) for every $a \in \mathbf{R}$ the process $\exp\{aW_t - a^2t/2\}$;

You are asked to prove this in Exercise 2.3.

Example 2.1.5 Let N be a Poisson process with rate λ. Then $\{N(t) - \lambda t\}_t$ is a martingale.

In the next section we first develop the theory for discrete-time martingales. The generalisation to continuous time is discussed in section 2.3. In section 2.4 we continue our study of BM.

2.2 Discrete-time martingales

In this section we restrict ourselves to martingales (and filtrations) that are indexed by (a subinterval of) $\mathbf{Z}_+$. We will assume the underlying filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_n)_n, \mathsf{P})$ to be fixed. Note that as a consequence, it only makes sense to consider $\overline{\mathbf{Z}}_+$-valued stopping times. In discrete time, τ is a stopping time with respect to the filtration $(\mathcal{F}_n)_{n\in\mathbf{Z}_+}$, if $\{\tau \leq n\} \in \mathcal{F}_n$ for all $n \in \mathbf{Z}_+$.

2.2.1 Martingale transforms

If the value of a process at time n is already known at time $n-1$, we call a process predictable. The precise definition is as follows.

Definition 2.2.1 We call a discrete-time process X *predictable* with respect to the filtration $(\mathcal{F}_n)_n$ if X_n is $\mathcal{F}_{n-1}$-measurable for every n.

In the following definition we introduce discrete-time 'integrals'. This is a useful tool in martingale theory.

Definition 2.2.2 Let M and X be two discrete-time processes. We define the process $X \cdot M$ by $(X \cdot M)_0 = 0$ and for $n \geq 1$

$$(X \cdot M)_n = \sum_{k=1}^{n} X_k(M_k - M_{k-1}).$$

We call $X \cdot M$ the discrete integral of X with respect to M. If M is a (sub-, super-)martingale, it is often called the *martingale transform of* M *by* X.

One can view martingale transforms as a discrete version of the Ito integral. The predictability plays a crucial role in the construction of the Ito integral.

The following lemma explains why these 'integrals' are so useful: the integral of a predictable process with respect to a martingale is again a martingale.

Lemma 2.2.3 *Let X be a predictable process, such that for all n there exists a constant K_n such that $|X_1|, \ldots, |X_n| \leq K_n$. If M is an martingale, then $X \cdot M$ is a martingale. If M be a submartingale (resp. a supermartingale) and X is non-negative then $X \cdot M$ is a submartingale (resp. supermartingale) as well.*

Proof. Put $Y = X \cdot M$. Clearly Y is adapted. Since X is bounded, say $|X_n| \leq K$ a.s., for all n, we have $\mathsf{E}|Y_n| \leq 2K_n \sum_{k \leq n} \mathsf{E}|M_k| < \infty$. Now suppose first that M is a submartingale and X is non-negative. Then a.s.

$$\begin{aligned}
\mathsf{E}(Y_n \,|\, \mathcal{F}_{n-1}) &= \mathsf{E}(Y_{n-1} + X_n(M_n - M_{n-1}) \,|\, \mathcal{F}_{n-1}) \\
&= Y_{n-1} + X_n \mathsf{E}(M_n - M_{n-1} \,|\, \mathcal{F}_{n-1}) \geq Y_{n-1}, \text{ a.s.}
\end{aligned}$$

Consequently, Y is a submartingale. If M is a martingale, the last inequality is an equality, irrespective of the sign of X_n. This implies that then Y is a martingale as well. QED

Using this lemma, it is easy to see that a stopped (sub-, super-)martingale is again a (sub-, super-)martingale.

Theorem 2.2.4 *Let M be a $(\mathcal{F}_n)_n$ (sub-, super-)martingale and τ an $(\mathcal{F}_n)_n$-stopping time. Then the stopped process M^τ is an $(\mathcal{F}_n)_n$ (sub-, super-)martingale as well.*

Proof. Define the process X by $X_n = \mathbf{1}_{\{\tau \geq n\}}$. Verify that $M^\tau = M_0 + X \cdot M$. Since τ is a stopping time, we have that $\{\tau \geq n\} = \{\tau \leq n-1\}^c \in \mathcal{F}_{n-1}$. Hence the process X is predictable. It is also a bounded process, and so the statement follows from the preceding lemma.

We will also give a direct proof. First note that $\mathsf{E}|M_t^\tau| = \mathsf{E}|M_{t\wedge\tau}| \leq \sum_{n=0}^{t} \mathsf{E}|M_n| < \infty$ for $t \in T$. Write

$$\begin{aligned}
M_t^\tau &= M_{t\wedge\tau} = \Big(\sum_{n=0}^{t-1} \mathbf{1}_{\{\tau=n\}} + \mathbf{1}_{\{\tau \geq t\}} \Big) M_{t\wedge\tau} \\
&= \sum_{n=0}^{t-1} M_n \mathbf{1}_{\{\tau=n\}} + M_t \mathbf{1}_{\{\tau \geq t\}}.
\end{aligned}$$

Taking conditional expectations yields

$$\mathsf{E}(M_t^\tau \mid \mathcal{F}_{t-1}) = \sum_{n=0}^{t-1} M_n \mathbf{1}_{\{\tau=n\}} + \mathbf{1}_{\{\tau\geq t\}} \mathsf{E}(M_t \mid \mathcal{F}_{t-1}),$$

since $\{\tau \geq t\} \in \mathcal{F}_{t-1}$. The rest follows immediately. QED

The following result can be viewed as a first version of the so-called *optional sampling theorem*. The general version will be discussed in section 2.2.5.

Theorem 2.2.5 *Let M be a (sub)martingale and let σ, τ be two stopping times such that $\sigma \leq \tau \leq K$, for some constant $K > 0$. Then*

$$\mathsf{E}(M_\tau \mid \mathcal{F}_\sigma)(\geq) = M_\sigma, \quad \text{a.s.} \tag{2.2.1}$$

An adapted integrable process M is a martingale if and only if

$$\mathsf{E}M_\tau = \mathsf{E}M_\sigma,$$

for any pairs of bounded stopping times $\sigma \leq \tau$.

Proof. Suppose first that M is a martingale. Define the predictable process $X_n = \mathbf{1}_{\{\tau\geq n\}} - \mathbf{1}_{\{\sigma\geq n\}}$. Note that $X_n \geq 0$ a.s.! Hence, $X \cdot M = M^\tau - M^\sigma$. By Lemma 2.2.3 the process $X \cdot M$ is a martingale, hence $\mathsf{E}(M_n^\tau - M_n^\sigma) = \mathsf{E}(X \cdot M)_n = 0$ for all n. Since $\sigma \leq \tau \leq K$ a.s., it follows that

$$\mathsf{E}M_\tau = \mathsf{E}M_K^\tau = \mathsf{E}M_K^\sigma = \mathsf{E}M_\sigma.$$

Now we take $A \in \mathcal{F}_\sigma$ and we define the 'truncated' random times

$$\sigma^A = \sigma \mathbf{1}_{\{A\}} + K\mathbf{1}_{\{A^c\}}, \quad \tau^A = \tau \mathbf{1}_{\{A\}} + K\mathbf{1}_{\{A^c\}}. \tag{2.2.2}$$

By definition of $\mathcal{F}_\sigma$ it holds for every n that

$$\{\sigma^A \leq n\} = (A \cap \{\sigma \leq n\}) \cup (A^c \cap \{K \leq n\}) \in \mathcal{F}_n,$$

and so σ^A is a stopping time. Similarly, τ^A is a stopping time and clearly $\sigma^A \leq \tau^A \leq K$ a.s. By the first part of the proof, it follows that $\mathsf{E}M_{\sigma^A} = \mathsf{E}M_{\tau^A}$, in other words

$$\int_A M_\sigma d\mathsf{P} + \int_{A^c} M_K d\mathsf{P} = \int_A M_\tau d\mathsf{P} + \int_{A^c} M_K d\mathsf{P}, \tag{2.2.3}$$

by which $\int_A M_\sigma d\mathsf{P} = \int_A M_\tau d\mathsf{P}$. Since $A \in \mathcal{F}_\sigma$ is arbitrary, $\mathsf{E}(M_\tau \mid \mathcal{F}_\sigma) = M_\sigma$ a.s. (recall that M_σ is $\mathcal{F}_\sigma$-measurable, cf. Lemma 1.6.15).

Let M be an adapted process with $\mathsf{E}M_\sigma = \mathsf{E}M_\tau$ for each bounded pair $\sigma \leq \tau$ of stopping times. Take $\sigma = n-1$ and $\tau = n$ in the preceding and use truncated stopping times σ^A and τ^A as in (2.2.2) for $A \in \mathcal{F}_{n-1}$. Then (2.2.3) for $A \in \mathcal{F}_{n-1}$ and stopping times σ^A and τ^A implies that $\mathsf{E}(M_n \mid \mathcal{F}_{n-1}) = M_{n-1}$ a.s. In other words, M is a martingale.

If M is a submartingale, the same reasoning applies, but with inequalities instead of equalities.

As in the previous lemma, we will also give a direct proof of (2.2.1). First note that

$$\mathsf{E}(M_K \mid \mathcal{F}_n) \geq M_n \text{ a.s.} \quad \Longleftrightarrow \quad \mathsf{E}M_K\mathbf{1}_{\{F\}} \geq \mathsf{E}M_n\mathbf{1}_{\{F\}}, \quad \forall F \in \mathcal{F}_n, \tag{2.2.4}$$

(this follows from Exercise 2.7 (a)). We will first show that

$$\mathsf{E}(M_K \,|\, \mathcal{F}_\sigma) \geq M_\sigma \text{ a.s.} \tag{2.2.5}$$

Similarly to (2.2.4) it is sufficient to show that $\mathsf{E}\mathbf{1}_{\{F\}}M_\sigma \leq \mathsf{E}\mathbf{1}_{\{F\}}M_K$ for all $F \in \mathcal{F}_\sigma$. Now,

$$\begin{aligned}
\mathsf{E}\mathbf{1}_{\{F\}}M_\sigma &= \mathsf{E}\mathbf{1}_{\{F\}}\Big(\sum_{n=0}^{K}\mathbf{1}_{\{\sigma=n\}} + \mathbf{1}_{\{\sigma>K\}}\Big)M_\sigma \\
&= \sum_{n=0}^{K}\mathsf{E}\mathbf{1}_{\{F\cap\{\sigma=n\}\}}M_n \\
&\leq \sum_{n=0}^{K}\mathsf{E}\mathbf{1}_{\{F\cap\{\sigma=n\}\}}M_K \\
&= \mathsf{E}\mathbf{1}_{\{F\}}\Big(\sum_{n=0}^{K}\mathbf{1}_{\{\sigma=n\}} + \mathbf{1}_{\{\sigma>K\}}\Big)M_K = \mathsf{E}\mathbf{1}_{\{F\}}M_K.
\end{aligned}$$

In the second and fourth equalities we have used that $\mathsf{E}\mathbf{1}_{\{\sigma>K\}}M_\sigma = \mathsf{E}\mathbf{1}_{\{\sigma>K\}}M_K = 0$, since $\mathsf{P}\{\sigma > K\} = 0$. In the third inequality, we have used (2.2.4) and the fact that $F \cap \{\sigma = n\} \in \mathcal{F}_n$ (why?). This shows the validity of (2.2.5).

Apply (2.2.5) to the stopped process M^τ. This yields

$$\mathsf{E}(M_K^\tau \,|\, \mathcal{F}_\sigma) \geq M_\sigma^\tau.$$

Now, note that $M_K^\tau = M_\tau$ a.s. and $M_\sigma^\tau = M_\sigma$ a.s. (why?). This shows (2.2.1). QED

Note that we may in fact allow that $\sigma \leq \tau \leq K$ a.s. Lateron we need $\sigma \leq \tau$ everywhere.

2.2.2 Inequalities

Markov's inequality implies that if M is a discrete time process, then

$$\lambda\mathsf{P}\{M_n \geq \lambda\} \leq \mathsf{E}|M_n|$$

for all $n \in \mathbf{Z}_+$ and $\lambda > 0$. Doob's classical submartingale inequality states that for submartingales we have a much stronger result.

Theorem 2.2.6 (Doob's submartingale inequality) *Let M be a submartingale. For all $\lambda > 0$ and $n \in \mathbf{N}$*

$$\lambda\mathsf{P}\{\max_{k\leq n} M_k \geq \lambda\} \leq \mathsf{E}M_n\mathbf{1}_{\{\max_{k\leq n} M_k \geq \lambda\}} \leq \mathsf{E}|M_n|.$$

Proof. Define $\tau = n \wedge \inf\{k \,|\, M_k \geq \lambda\}$. This is a stopping time (see Lemma 1.6.9) with $\tau \leq n$. By Theorem 2.2.5, we have $\mathsf{E}M_n \geq \mathsf{E}M_\tau$. It follows that

$$\begin{aligned}
\mathsf{E}M_n &\geq \mathsf{E}M_\tau\mathbf{1}_{\{\max_{k\leq n} M_k \geq \lambda\}} + \mathsf{E}M_\tau\mathbf{1}_{\{\max_{k\leq n} M_k < \lambda\}} \\
&\geq \lambda\mathsf{P}\{\max_{k\leq n} M_k \geq \lambda\} + \mathsf{E}M_n\mathbf{1}_{\{\max_{k\leq n} M_k < \lambda\}}.
\end{aligned}$$

This yields the first inequality. The second one is obvious. QED

Theorem 2.2.7 (Doob's L^p inequality) *If M is a martingale or a non-negative submartingale and $p > 1$, then for all $n \in \mathbf{N}$*

$$\mathsf{E}\Big(\max_{k \le n} |M_n|^p\Big) \le \Big(\frac{p}{1-p}\Big)^p \mathsf{E}|M_n|^p,$$

provided M is in L^p.

Proof. Define $M^* = \max_{k \le n} |M_k|$. Assume that M is defined on an underlying probability space $(\Omega, \mathcal{F}, \mathsf{P})$. We have for any $m \in \mathbf{N}$

$$\begin{aligned} \mathsf{E}(M^* \wedge m)^p &= \int_\omega (M^*(\omega) \wedge m)^p d\mathsf{P}(\omega) \\ &= \int_\omega \int_0^{M^*(\omega)\wedge m} px^{p-1} dx d\mathsf{P}(\omega) \\ &= \int_\omega \int_0^m px^{p-1} \mathbf{1}_{\{M^*(\omega) \ge x\}} dx d\mathsf{P}(\omega) \\ &= \int_0^m px^{p-1} \mathsf{P}\{M^* \ge x\} dx, \end{aligned} \tag{2.2.6}$$

where we have used Fubini's theorem in the last equality (non-negative integrand!). By conditional Jensen's inequality, $|M|$ is a submartingale, and so we can apply Doob's submartingale inequality to estimate $\mathsf{P}\{M^* \ge x\}$. Thus

$$\mathsf{P}\{M^* \ge x\} \le \frac{\mathsf{E}(|M_n| \mathbf{1}_{\{M^* \ge x\}})}{x}.$$

Insert this in (2.2.6), then

$$\begin{aligned} \mathsf{E}(M^* \wedge m)^p &\le \int_0^m px^{p-2} \mathsf{E}(|M_n| \mathbf{1}_{\{M^* \ge x\}}) dx \\ &= \int_0^m px^{p-2} \int_{\omega: M^*(\omega) \ge x} |M_n(\omega)| d\mathsf{P}(\omega) dx \\ &= p \int_\omega |M_n(\omega)| \int_0^{M^*(\omega)\wedge m} x^{p-2} dx\, d\mathsf{P}(\omega) \\ &= \frac{p}{p-1} \mathsf{E}(|M_n| (M^* \wedge m)^{p-1}). \end{aligned}$$

By Hölder's inequality, it follows that with $p^{-1} + q^{-1} = 1$

$$\mathsf{E}|M^* \wedge m|^p \le \frac{p}{p-1} (\mathsf{E}|M_n|^p)^{1/p} (\mathsf{E}|M^* \wedge m|^{(p-1)q})^{1/q}.$$

Since $p > 1$ we have $q = p/(p-1)$, so that

$$\mathsf{E}|M^* \wedge m|^p \le \frac{p}{p-1} (\mathsf{E}|M_n|^p)^{1/p} (\mathsf{E}|M^* \wedge m|^p)^{(p-1)/p}.$$

Now take pth power of both sides and cancel common factors. Then

$$\mathsf{E}|M^* \wedge m|^p \le \Big(\frac{p}{p-1}\Big)^p \mathsf{E}|M_n|^p.$$

The proof is completed by letting m tend to infinity. QED

2.2.3 Doob decomposition

An adapted, integrable process X can always be written as a sum of a martingale and a predictable process. This is called the *Doob decomposition* of the process X.

Theorem 2.2.8 *Let X be an adapted, integrable process. There exists a martingale M and a predictable process A, such that $A_0 = M_0 = 0$ and $X = X_0 + M + A$. The processes M and A are a.s. unique. The process X is a submartingale if and only if A is a.s. increasing (i.e. $\mathsf{P}\{A_n \leq A_{n+1}\} = 1$).*

Proof. Suppose first that there exist a martingale M and a predictable process A such that $A_0 = M_0 = 0$ and $X = X_0 + M + A$. The martingale property of M and predictability of A show that a.s.

$$\mathsf{E}(X_n - X_{n-1} \mid \mathcal{F}_{n-1}) = A_n - A_{n-1} \quad \text{a.s.} \tag{2.2.7}$$

Since $A_0 = 0$ it follows that

$$A_n = \sum_{k=1}^{n} \mathsf{E}(X_k - X_{k-1} \mid \mathcal{F}_{k-1}), \tag{2.2.8}$$

for $n \geq 1$ and hence $M_n = X_n - A_n - X_0$. This shows that M and A are a.s. unique.

Conversely, given a process X, (2.2.8) defines a predictable process A. It is easily seen that the process M defined by $M = X - A - X_0$ is a martingale. This proves the existence of the decomposition.

Equation (2.2.7) shows that X is a submartingale if and only if A is increasing. QED

An important application of the Doob decomposition is the following.

Corollary 2.2.9 *Let M be a martingale with $\mathsf{E}M_n^2 < \infty$ for all n. Then there exists an a.s. unique predictable, increasing process A with $A_0 = 0$ such that $M^2 - A$ is a martingale. Moreover the random variable $A_{n+1} - A_n$ is a version of the conditional variance of M_n given $\mathcal{F}_{n-1}$, i.e.*

$$A_n - A_{n-1} = \mathsf{E}\Big((M_n - \mathsf{E}(M_n \mid \mathcal{F}_{n-1}))^2 \mid \mathcal{F}_{n-1}\Big) = \mathsf{E}\Big((M_n - M_{n-1})^2 \mid \mathcal{F}_{n-1}\Big) \quad \text{a.s.}$$

It follows that Pythagoras' theorem holds for square integrable martingales

$$\mathsf{E}M_n^2 = \mathsf{E}M_0^2 + \sum_{k=1}^{n} \mathsf{E}(M_k - M_{k-1})^2.$$

The process A is called the *predictable quadratic variation process of M* and is often denoted by $\langle M \rangle$.

Proof. By conditional Jensen, it follows that M^2 is a submartingale. Hence Theorem 2.2.8 applies. The only thing left to prove is the statement about conditional variance. Since M is a martingale, we have a.s.

$$\begin{aligned} \mathsf{E}((M_n - M_{n-1})^2 \mid \mathcal{F}_{n-1}) &= \mathsf{E}(M_n^2 - 2M_n M_{n-1} + M_{n-1}^2 \mid \mathcal{F}_{n-1}) \\ &= \mathsf{E}(M_n^2 \mid \mathcal{F}_{n-1}) - 2M_{n-1}\mathsf{E}(M_n \mid \mathcal{F}_{n-1}) + M_{n-1}^2 \\ &= \mathsf{E}(M_n^2 \mid \mathcal{F}_{n-1}) - M_{n-1}^2 \\ &= \mathsf{E}(M_n^2 - M_{n-1}^2 \mid \mathcal{F}_{n-1}) = A_n - A_{n-1}. \end{aligned}$$

QED

Using the Doob decomposition in combination with the submartingale inequality yields the following result.

Theorem 2.2.10 *Let X be a sub- or supermartingale. For all $\lambda > 0$ and $n \in \mathbf{Z}_+$*

$$\lambda \mathsf{P}\{\max_{k\le n} |X_k| \ge 3\lambda\} \le 4\mathsf{E}|X_0| + 3\mathsf{E}|X_n|.$$

Proof. Suppose that X is a submartingale. By the Doob decomposition theorem there exist a martingale M and an increasing, predictable process A such that $M_0 = A_0 = 0$ and $X = X_0 + M + A$. By the triangle inequality and the fact that A is increasing

$$\mathsf{P}\{\max_{k\le n} |X_k| \ge 3\lambda\} \le \mathsf{P}\{|X_0| \ge \lambda\} + \mathsf{P}\{\max_{k\le n} |M_k| \ge \lambda\} + \mathsf{P}\{A_n \ge \lambda\}.$$

Hence, by Markov's inequality and the submartingale inequality ($|M_n|$ is a submartingale!)

$$\lambda \mathsf{P}\{\max_{k\le n} |X_k| \ge 3\lambda\} \le \mathsf{E}|X_0| + \mathsf{E}|M_n| + \mathsf{E}A_n.$$

Since $M_n = X_n - X_0 - A_n$, the right-hand side is bounded by $2\mathsf{E}|X_0| + \mathsf{E}|X_n| + 2\mathsf{E}A_n$. We know that A_n is given by (2.2.7). Taking expectations in the latter expression shows that $\mathsf{E}A_n = \mathsf{E}X_n - \mathsf{E}X_0 \le \mathsf{E}|X_n| + \mathsf{E}|X_0|$. This completes the proof. QED

2.2.4 Convergence theorems

Let M be a supermartingale and consider a compact interval $[a, b] \subset \mathbf{R}$. The number of *upcrossings* of $[a, b]$ that the process makes upto time n is the number of time that the process passes from a level below a to a level above b. The precise definition is as follows.

Definition 2.2.11 The number $U_n[a, b]$ is the largest value $k \in \mathbf{Z}_+$, such that there exis $0 \le s_1 < t_1 < s_2 < \cdots < s_k < t_k \le n$ with $M_{s_i} < a$ and $M_{t_i} > b$, $i = 1, \ldots, k$.

First we define the "limit σ-algebra

$$\mathcal{F}_\infty = \sigma\Big(\bigcup_n \mathcal{F}_n\Big).$$

Lemma 2.2.12 (Doob's upcrossing lemma) *Let M be a supermartingale. Then for all $a < b$, the number of upcrossings $U_n[a, b]$ of the interval $[a, b]$ by M upto time n is an $\mathcal{F}_n$-measurable random variable and satisfies*

$$(b - a)\mathsf{E}U_n[a, b] \le \mathsf{E}(M_n - a)^-.$$

The total number of upcrossings $U_\infty[a, b]$ is $\mathcal{F}_\infty$-measurable.

Proof. Check yourself that $U_n[a, b]$ is $\mathcal{F}_n$-measurable and that $U_\infty[a, b]$ is $\mathcal{F}_\infty$-measurable. Consider the bounded, predictable process X given by $X_0 = \mathbf{1}_{\{M_0<a\}}$ and

$$X_n = \mathbf{1}_{\{X_{n-1}=1\}}\mathbf{1}_{\{M_{n-1}\le b\}} + \mathbf{1}_{\{X_{n-1}=0\}}\mathbf{1}_{\{M_{n-1}<a\}}, \quad n \in \mathbf{Z}_+.$$

Define $Y = X \cdot M$. The process X equals 0, until M drops below level a, then stays until M gets above b etc. So every completed upcrossing of $[a, b]$ increases the value of Y by at least $b - a$. If the last upcrossing has not yet been completed at time n, then this may reduce Y by at most $(M_n - a)^-$. Hence

$$Y_n \geq (b-a)U_n[a,b] - (M_n - a)^-. \qquad (2.2.9)$$

By Lemma 2.2.3, the process $Y = X \cdot M$ is a supermartingale. In particular $\mathsf{E}Y_n \leq \mathsf{E}Y_0 = 0$. The proof is completed by taking expectations in both sides of (2.23). QED

Observe that the upcrossing lemma implies for a supermartingale M that is bounded in L^1 (i.e. $\sup_n \mathsf{E}|M_n| < \infty$) that $\mathsf{E}U_\infty[a,b] < \infty$ for all $a \leq b$. In particular, the total number $U_\infty[a,b]$ of upcrossings of the interval $[a,b]$ is almost surely finite. The proof of the classical martingale convergence theorem is now straightforward.

Theorem 2.2.13 (Doob's martingale convergence theorem) *If M is a supermartingale that is bounded in L^1, then M_n converges a.s. to a finite $\mathcal{F}_\infty$-measurable limit M_∞ as $n \to \infty$, with $\mathsf{E}|M_\infty| < \infty$.*

Proof. Assume that M is defined on the underlying probability space $(\Omega, \mathcal{F}, \mathsf{P})$. Suppose that $M(\omega)$ does not converge to a limit in $[-\infty, \infty]$. Then there exist two rationals $a < b$ such that $\liminf M_n(\omega) < a < b < \limsup M_n(\omega)$. In particular, we must have $U_\infty[a,b](\omega) = \infty$. By Doob's upcrossing lemma $\mathsf{P}\{U_\infty[a,b] = \infty\} = 0$. Now note that

$$A := \{\omega \mid M(\omega) \text{ does not converge to a limit in } [-\infty,\infty]\} \subset \bigcup_{\substack{a,b \in \mathbf{Q}: \\ a<b}} \{\omega \mid U_\infty[a,b](\omega) = \infty\}.$$

Hence $\mathsf{P}\{A\} \leq \sum_{\substack{a,b\in\mathbf{Q}: \\ a<b}} \mathsf{P}\{U_\infty[a,b] = \infty\} = 0$. This implies that M_n a.s. converges to a limit M_∞ in $[-\infty, \infty]$. Moreover, in view of Fatou's lemma

$$\mathsf{E}|M_\infty| = \mathsf{E}(\liminf |M_n|) \leq \liminf \mathsf{E}|M_n| \leq \sup \mathsf{E}|M_n| < \infty.$$

It follows that M_∞ is a.s. finite and it is integrable. Note that M_n is $\mathcal{F}_n$-measurable, hence it is $\mathcal{F}_\infty$-measurable. Since $M_\infty = \lim_{n\to\infty} M_n$ is the limit of $\mathcal{F}_\infty$-measurable maps, it is $\mathcal{F}_\infty$-measurable as well (see MTP-lecture notes). QED

If the supermartingale M is not only bounded in L^1 but also uniformly integrable, the in addition to a.s. convergence we have convergence in L^1. Moreover, in the case, the whole sequence $M_1, \ldots, M_\infty$ is a supermartingale.

Theorem 2.2.14 *Let M be a supermartingale that is bounded in L^1. Then $M_n \stackrel{\mathsf{L}^1}{\to} M_\infty$, $n \to \infty$, if and only if $\{M_n \mid n \in \mathbf{Z}_+\}$ is uniformly integrable, where M_∞ is integrable and $\mathcal{F}_\infty$-measurable. In that case*

$$\mathsf{E}(M_\infty \mid \mathcal{F}_n) \leq M_n, \quad a.s. \qquad (2.2.10)$$

If in addition M is a martingale, then there is equality in (2.2.10), in other words, M is a Doob martingale.

Proof. By virtue of Theorem 2.2.13 $M_n \to M_\infty$ a.s., for a finite random variable M_∞. BN Theorem 8.5 implies the first statement. To prove the second statement, suppose that $M_n \xrightarrow{\mathsf{L}^1} M_\infty$. Since M is a supermartingale, we have

$$\mathsf{E}\mathbf{1}_{\{A\}}M_m \le \mathsf{E}\mathbf{1}_{\{A\}}M_n, \quad A \in \mathcal{F}_n, m \ge n. \tag{2.2.11}$$

Since $|\mathbf{1}_{\{A\}}M_m - \mathbf{1}_{\{A\}}M_\infty| \le \mathbf{1}_{\{A\}}|M_m - M_\infty| \le |M_m - M_\infty|$, it follows directly that $\mathbf{1}_{\{A\}}M_m \xrightarrow{\mathsf{L}^1} \mathbf{1}_{\{A\}}M_\infty$. Taking the limit $m \to \infty$ in (2.2.11) yields

$$\mathsf{E}\mathbf{1}_{\{A\}}M_\infty \le \mathsf{E}\mathbf{1}_{\{A\}}M_n, \quad A \in \mathcal{F}_n.$$

This implies (see BN Exercise 2.7(a)) that $\mathsf{E}(M_\infty \mid \mathcal{F}_n) \le M_n$ a.s. QED

Hence uniformly integrable martingales that are bounded in L^1, are Doob martingales. On the other hand, let X be an $\mathcal{F}$-measurable, integrable random variable and let $(\mathcal{F}_n)_n$ be a filtration. Then (Example 2.1.2) $\mathsf{E}(X \mid \mathcal{F}_n)$ is a uniformly integrable Doob martingale. By uniform integrability, it is bounded in L^1. For Doob martingales, we can identify the limit explicitly in terms of the limit σ-algebra $\mathcal{F}_\infty$.

Theorem 2.2.15 (Lévy's upward theorem) *Let X be an integrable random variable, defined on a probability space $(\Omega, \mathcal{F}, \mathsf{P})$, and let $(\mathcal{F}_n)_n$ be a filtration, $\mathcal{F}_n \subset \mathcal{F}$, for all n. Then as $n \to \infty$*

$$\mathsf{E}(X \mid \mathcal{F}_n) \to \mathsf{E}(X \mid \mathcal{F}_\infty),$$

a.s. and in L^1.

Proof. The process $M_n = \mathsf{E}(X \mid \mathcal{F}_n)$ is uniformly integrable (see Example 2.1.2), hence bounded in L^1 (explain!). By Theorem 2.2.14 $M_n \to M_\infty$ a.s. and in L^1, as $n \to \infty$ with M_∞ integrable and $\mathcal{F}_\infty$-measurable. It remains to show that $M_\infty = \mathsf{E}(X \mid \mathcal{F}_\infty)$ a.s. Note that

$$\mathsf{E}\mathbf{1}_{\{A\}}M_\infty = \mathsf{E}\mathbf{1}_{\{A\}}M_n = \mathsf{E}\mathbf{1}_{\{A\}}X, \quad A \in \mathcal{F}_n, \tag{2.2.12}$$

where we have used Theorem 2.2.14 for the first equality and the definition of M_n for the second. First assume that $X \ge 0$, then $M_n = \mathsf{E}(X \mid \mathcal{F}_n) \ge 0$ a.s. (see BN Lemma 7.2 (iv)), hence $M_\infty \ge 0$ a.s.

As in the construction of the conditional expectation we will associate measures with X and M_∞ and show that they agree on a π-system for $\mathcal{F}_\infty$. Define measures Q_1 and Q_2 on $(\Omega, \mathcal{F}_\infty)$ by

$$Q_1(A) = \mathsf{E}\mathbf{1}_{\{A\}}X, \quad Q_2(A) = \mathsf{E}\mathbf{1}_{\{A\}}M_\infty.$$

(Check that these are indeed measures). By virtue of (2.2.12) Q_1 and Q_2 agree on the π-system (algebra) $\cup_n \mathcal{F}_n$. Moreover, $Q_1(\Omega) = Q_2(\Omega)(= \mathsf{E}X)$ since $\Omega \in \mathcal{F}_n$. By virtue of BN Lemma 1.1, Q_1 and Q_2 agree on $\sigma(\cup_n \mathcal{F}_n)$. This implies by definition of conditional expectation that $M_\infty = \mathsf{E}(X \mid \mathcal{F}_\infty)$ a.s.

Finally we consider the case of general $\mathcal{F}$-measurable X. Then $X = X^+ - X^-$, is the difference of two non-negative $\mathcal{F}$-measurable functions X^+ and X^-. Use the linearity of conditional expectation. QED

The message here is that one cannot know more than what one can observe. We will also need the corresponding result for decreasing families of σ-algebras. If we have a filtration of the form $(\mathcal{F}_n)_{n\in-\mathbf{Z}_+}$, i.e. a collection of σ-algebras such that $\mathcal{F}_{-(n+1)} \subseteq \mathcal{F}_{-n}$, then we define

$$\mathcal{F}_{-\infty} = \bigcap_n \mathcal{F}_{-n}.$$

Theorem 2.2.16 (Lévy-Doob downward theorem) *Let $(\mathcal{F}_{-n} \,|\, n \in \mathbf{Z}_+)$ be a collection of σ-algebras, such that $\mathcal{F}_{-(n+1)} \subseteq \mathcal{F}_{-n}$ for every n, and let $M = (\cdots, M_{-2}, M_{-1})$ be a supermartingale, i.e.*

$$\mathsf{E}(M_{-m} \,|\, \mathcal{F}_{-n}) \leq M_{-n} \quad a.s., \qquad \text{for all } -n \leq -m \leq -1.$$

If $\sup \mathsf{E}M_{-n} < \infty$, then the process M is uniformly integrable and the limit

$$M_{-\infty} = \lim_{n\to\infty} M_{-n}$$

exists a.s. and in L^1. Moreover,

$$\mathsf{E}(M_{-n} \,|\, \mathcal{F}_{-\infty}) \leq M_{-\infty} \quad a.s. \tag{2.2.13}$$

If M is a martingale, we have equality in (2.2.13) and in particular $M_{-\infty} = \mathsf{E}(M_{-1} \,|\, \mathcal{F}_{-\infty})$.

Proof. For every $n \in \mathbf{Z}_+$ the upcrossing inequality applied to the supermartingale

$$(M_{-n}, M_{-(n-1)}, \ldots, M_{-1})$$

yields $(b-a)\mathsf{E}U_n[a,b] \leq \mathsf{E}(M_{-1}-a)^-$ for every $a < b$. By a similar reasoning as in the proof of Theorem 2.2.13, we see that the limit $M_{-\infty} = \lim_{n\to-\infty} M_{-n}$ exists and is finite almost surely.

Next, we would like to show uniform integrability. For all $K > 0$ and $n \in -\mathbf{Z}_+$ we have

$$\int_{|M_{-n}|>K} |M_{-n}| d\mathsf{P} = \mathsf{E}M_{-n} - \int_{M_{-n}\leq K} M_{-n} d\mathsf{P} - \int_{M_{-n}<-K} M_{-n} d\mathsf{P}.$$

The sequence $\mathsf{E}M_{-n}$ is non-decreasing in $n \to -\infty$, and bounded. Hence the limit $\lim_{n\to\infty} \mathsf{E}M_{-n}$ exists (as a finite number). For arbitrary $\epsilon > 0$, there exists $m \in \mathbf{Z}_+$, such that $\mathsf{E}M_{-n} \leq \mathsf{E}M_{-m} + \epsilon$, $n \geq m$. Together with the supermartingale property this implies for all $n \geq m$

$$\begin{aligned}\int_{|M_{-n}|>K} |M_{-n}| d\mathsf{P} &\leq \mathsf{E}M_{-m} + \epsilon - \int_{M_{-n}\leq K} M_{-m} d\mathsf{P} - \int_{M_{-n}<-K} M_{-m} d\mathsf{P} \\ &\leq \int_{|M_{-n}|>K} |M_{-m}| d\mathsf{P} + \epsilon.\end{aligned}$$

Hence to prove uniform integrability, in view of BN Lemma 8.1 it is sufficient to show that we can make $\mathsf{P}\{|M_{-n}| > K\}$ arbitrarily small for all n simultaneously. By Chebychev's inequality, it suffices to show that $\sup_n \mathsf{E}|M_{-n}| < \infty$.

To this end, consider the process $M^- = \max\{-M, 0\}$. With $g : \mathbf{R} \to \mathbf{R}$ given by $g(x) = \max\{x, 0\}$, one has $M^- = g(-M)$. The function g is a non-decreasing, convex function.

Since $-M$ is a submartingale, it follows that M^- is a submartingale (see Example 2.1.3). In particular, $\mathsf{E}M^-_{-n} \le \mathsf{E}M^-_{-1}$ for all $n \in \mathbf{Z}_+$. It follows that

$$\mathsf{E}|M_{-n}| = \mathsf{E}M_{-n} + 2\mathsf{E}M^-_{-n} \le \sup \mathsf{E}M_{-n} + 2\mathsf{E}|M_{-1}|.$$

Consequently

$$\mathsf{P}\{|M_{-n}| > K\} \le \frac{1}{K}(\sup_n \mathsf{E}M_{-n} + 2\mathsf{E}|M_{-1}|).$$

Indeed, M is uniformly integrable. The limit $M_{-\infty}$ therefore exists in L^1 as well.

Suppose that M is a martingale. Then $M_{-n} = \mathsf{E}(M_{-1} \,|\, \mathcal{F}_{-n})$ a.s. The rest follows in a similar manner as the proof of the Lévy upward theorem. QED

Note that the downward theorem includes the "downward version" of Theorem 2.2.15 as a special case. Indeed, if X is an integrable random variable and $\mathcal{F}_1 \supseteq \mathcal{F}_2 \supseteq \cdots \supseteq \cap_n \mathcal{F}_n = \mathcal{F}_\infty$ is a decreasing sequence of σ-algebras, then

$$\mathsf{E}(X \,|\, \mathcal{F}_n) \to \mathsf{E}(X \,|\, \mathcal{F}_\infty), \quad n \to \infty$$

a.s. and in L^1. This is generalised in the following corollary to Theorems 2.2.15 and 2.2.16. It will be useful in the sequel.

Corollary 2.2.17 *Suppose that $X_n \to X$ a.s., and that $|X_n| \le Y$ a.s. for all n, where Y is an integrable random variable. Moreover, suppose that $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \cdots$ (resp. $\mathcal{F}_1 \supseteq \mathcal{F}_2 \supseteq \cdots$) is an increasing (resp. decreasing) sequence of σ-algebras. Then $\mathsf{E}(X_n \,|\, \mathcal{F}_n) \to \mathsf{E}(X \,|\, \mathcal{F}_\infty)$ a.s., where $\mathcal{F}_\infty = \sigma(\cup_n \mathcal{F}_n)$ (resp. $\mathcal{F}_\infty = \cap_n \mathcal{F}_n$).*

In case of an increasing sequence of σ-algebras, the corollary is known as *Hunt's lemma*.

Proof. For $m \in \mathbf{Z}_+$, put $U_m \inf_{n \ge m} X_n$ and $V_m = \sup_{n \ge m} X_n$. Since $X_m \to X$ a.s., necessarily $V_m - U_m \to 0$ a.s., as $m \to \infty$. Furthermore $|V_m - U_m| \le 2Y$. Dominated convergence then implies that $\mathsf{E}(V_m - U_m) \to 0$, as $m \to \infty$. Fix $\epsilon > 0$ and choose m so large that $\mathsf{E}(V_m - U_m) < \epsilon$. For $n \ge m$ we have

$$U_m \le X_n \le V_m \quad \text{a.s.} \tag{2.2.14}$$

Consequently $\mathsf{E}(U_m \,|\, \mathcal{F}_n) \le \mathsf{E}(X_n \,|\, \mathcal{F}_n) \le \mathsf{E}(V_m \,|\, \mathcal{F}_n)$ a.s. The processes on the left and right are martingales that satisfy the conditions of the upward (resp. downward) theorem. Letting n tend to ∞ we obtain

$$\mathsf{E}(U_m \,|\, \mathcal{F}_\infty) \le \liminf \mathsf{E}(X_n \,|\, \mathcal{F}_n) \le \limsup \mathsf{E}(X_n \,|\, \mathcal{F}_n) \le \mathsf{E}(V_m \,|\, \mathcal{F}_\infty) \quad \text{a.s.} \tag{2.2.15}$$

It follows that

$$\begin{aligned} 0 \le \mathsf{E}\Big(\limsup \mathsf{E}(X_n \,|\, \mathcal{F}_n) - \liminf \mathsf{E}(X_n \,|\, \mathcal{F}_n)\Big) &\le \mathsf{E}\Big(\mathsf{E}(V_m \,|\, \mathcal{F}_\infty) - \mathsf{E}(U_m \,|\, \mathcal{F}_\infty)\Big) \\ &\le \mathsf{E}(V_m - U_m) < \epsilon. \end{aligned}$$

Letting $\epsilon \downarrow 0$ yields that $\limsup \mathsf{E}(X_n \,|\, \mathcal{F}_n) = \liminf \mathsf{E}(X_n \,|\, \mathcal{F}_n)$ a.s. and so $\mathsf{E}(X_n \,|\, \mathcal{F}_n)$ converges a.s. We wish to identify the limit. Let $n \to \infty$ in (2.2.14). Then $U_m \le X \le V_m$ a.s. Hence

$$\mathsf{E}(U_m \,|\, \mathcal{F}_\infty) \le \mathsf{E}(X \,|\, \mathcal{F}_\infty) \le \mathsf{E}(V_m \,|\, \mathcal{F}_\infty) \quad \text{a.s.} \tag{2.2.16}$$

Equations (2.2.15) and (2.2.16) impy that both $\lim \mathsf{E}(X_n \mid \mathcal{F}_n)$ and $\mathsf{E}(X \mid \mathcal{F}_\infty)$ are a.s.between V_m and U_m. Consequently

$$\mathsf{E}|\lim \mathsf{E}(X_n \mid \mathcal{F}_n) - \mathsf{E}(X \mid \mathcal{F}_\infty)| \le \mathsf{E}(V_m - U_m) < \epsilon.$$

By letting $\epsilon \downarrow 0$ we obtain that $\lim_n \mathsf{E}(X_n \mid \mathcal{F}_n) = \mathsf{E}(X \mid \mathcal{F}_\infty)$ a.s. QED

2.2.5 Optional sampling theorems

Theorem 2.2.5 implies for a martingale M and two *bounded* stopping times $\sigma \le \tau$ that $\mathsf{E}(M_\tau \mid \mathcal{F}_\sigma) = M_\sigma$. The following theorem extends this result.

Theorem 2.2.18 (Optional sampling theorem) *Let M be a uniformly integrable (super)martingale. Then the family of random variables $\{M_\tau \mid \tau$ is a finite stopping time$\}$ is uniformly integrable and for all stopping times $\sigma \le \tau$ we have*

$$\mathsf{E}(M_\tau \mid \mathcal{F}_\sigma) = (\le) M_\sigma \quad \text{a.s.}$$

Proof. We will only prove the martingale statement. For the proof in case of a supermartingale see Exercise 2.14.

By Theorem 2.2.14, $M_\infty = \lim_{n\to\infty} M_n$ exists a.s. and in L^1 and $\mathsf{E}(M_\infty \mid \mathcal{F}_n) = M_n$ a.s. Now let τ be an arbitrary stopping time and $n \in \mathbf{Z}_+$. Since $\tau \wedge n \le n$, $\mathcal{F}_{\tau\wedge n} \subseteq \mathcal{F}_n$. By the tower property, it follows for every n that

$$\mathsf{E}(M_\infty \mid \mathcal{F}_{\tau\wedge n}) = \mathsf{E}(\mathsf{E}(M_\infty \mid \mathcal{F}_n) \mid \mathcal{F}_{\tau\wedge n}) = \mathsf{E}(M_n \mid \mathcal{F}_{\tau\wedge n}) \quad \text{a.s.}$$

By Theorem 2.2.5 we a.s. have

$$\mathsf{E}(M_\infty \mid \mathcal{F}_{\tau\wedge n}) = M_{\tau\wedge n}.$$

Now let n tend to infinity. Then the right-hand side converges a.s. to M_τ. By the Levy upward convergence theorem, the left-hand side converges a.s. and in L^1 to $\mathsf{E}(M_\infty \mid \mathcal{G})$, where

$$\mathcal{G} = \sigma\Big(\bigcup_n \mathcal{F}_{\tau\wedge n}\Big).$$

Therefore

$$\mathsf{E}(M_\infty \mid \mathcal{G}) = M_\tau \quad \text{a.s.} \tag{2.2.17}$$

We have to show that $\mathcal{G}$ can be replaced by $\mathcal{F}_\tau$. Take $A \in \mathcal{F}_\tau$. Then

$$\mathsf{E}\mathbf{1}_{\{A\}} M_\infty = \mathsf{E}\mathbf{1}_{\{A\cap\{\tau<\infty\}\}} M_\infty + \mathsf{E}\mathbf{1}_{\{A\cap\{\tau=\infty\}\}} M_\infty.$$

By virtue of Exercise 2.6, relation (2.2.17) implies that

$$\mathsf{E}\mathbf{1}_{\{A\cap\{\tau<\infty\}\}} M_\infty = \mathsf{E}\mathbf{1}_{\{A\cap\{\tau<\infty\}\}} M_\tau.$$

Trivially

$$\mathsf{E}\mathbf{1}_{\{A\cap\{\tau=\infty\}\}} M_\infty = \mathsf{E}\mathbf{1}_{\{A\cap\{\tau=\infty\}\}} M_\tau.$$

Combination yields
$$\mathsf{E}\mathbf{1}_{\{A\}}M_\infty = \mathsf{E}\mathbf{1}_{\{A\}}M_\tau, \quad A \in \mathcal{F}_\tau.$$
We conclude hat $\mathsf{E}(M_\infty \,|\, \mathcal{F}_\tau) = M_\tau$ a.s. The first statement of the theorem follows from BN Lemma 8.4. The second statement follows from the tower property and the fact that $\mathcal{F}_\sigma \subseteq \mathcal{F}_\tau$. QED

For the equality $\mathsf{E}(M_\tau \,|\, \mathcal{F}_\sigma) = M_\sigma$ a.s. in the preceding theorem to hold, it is necessary that M is uniformly integrable. There exist (positive) martingales that are bounded in L^1 but not uniformly integrable, for which the equality fails in general (see Exercise 2.27)! For nonnegative supermartingales without additional integrability properties we only have an inequality.

Theorem 2.2.19 *Let M be a nonnegative supermartingale and let $\sigma \leq \tau$ be stopping times. Then*
$$\mathsf{E}(M_\tau \,|\, \mathcal{F}_\sigma) \leq M_\sigma \quad \text{a.s.}$$

Proof. First note that M is bounded in L^1 and so it converges a.s. Fix $n \in \mathbf{Z}_+$. The stopped supermartingale $M^{\tau\wedge n}$ is a supermartingale again (cf. Theorem 2.2.4). Check that it is uniformly integrable. Precisely as in the proof of the preceding theorem we find that
$$\mathsf{E}(M_{\tau\wedge n} \,|\, \mathcal{F}_\sigma) = \mathsf{E}(M_\infty^{\tau\wedge n} \,|\, \mathcal{F}_\sigma) \leq M_\sigma^{\tau\wedge n} = M_{\sigma\wedge n} \quad \text{a.s.}$$
Since the limit exists, we have $M_\tau \mathbf{1}_{\{\tau=\infty\}} = M_\infty \mathbf{1}_{\{\tau=\infty\}}$. By conditional Fatou
$$\begin{aligned}\mathsf{E}(M_\tau \,|\, \mathcal{F}_\sigma) &\leq \mathsf{E}(\liminf M_{\tau\wedge n} \,|\, \mathcal{F}_\sigma)\\ &\leq \liminf \mathsf{E}(M_{\tau\wedge n} \,|\, \mathcal{F}_\sigma)\\ &\leq \liminf M_{\sigma\wedge n} = M_\sigma, \quad \text{a.s.}\end{aligned}$$
This proves the result. QED

2.2.6 Law of Large numbers

We have already pointed out that martingales are a generalisation of sums of i.i.d. random variables with zero expectation. For such sums, we can derive the Law of Large Numbers, Central Limit Theorem, and law of Iterated Logarithm. The question is then: do these Laws also apply to martingales? If yes, what sort of conditions do we need to require.

Here we discuss a simplest version. To this end we need some preparations.

Theorem 2.2.20 *Let $S = (S_n = \sum_{k=1}^n X_k)_{n=1,2,\ldots}$ be a martingale with respect to the filtration $(\mathcal{F}_n)_{n=1,2,\ldots}$. Assume that $\mathsf{E}S_n = 0$ and that $\mathsf{E}S_n^2 < \infty$ for all n. Then S_n converges a.s. on the set $\{\sum_{k=1}^\infty \mathsf{E}(X_k^2 \,|\, \mathcal{F}_{k-1}) < \infty\}$.*

Proof. Let $\mathcal{F}_0$ be the trivial σ-algebra. Fix $K > 0$ and let $\tau = \min\{n \,|\, \sum_{k=1}^{n+1} \mathsf{E}(X_k^2 \,|\, \mathcal{F}_{k-1}) > K\}$, if such n exists. Otherwise let $\tau = \infty$. Clearly τ is a stopping time (check yourself). Then S^τ is a martingale.

Note that $S_n^\tau = S_{\tau\wedge n} = \sum_{k=1}^n \mathbf{1}_{\{\tau\geq k\}} X_k$. Using the martingale property and the fact that $\{\tau \geq k\} \in \mathcal{F}_{k-1}$, we obtain

$$\begin{aligned} \mathsf{E} S_{\tau\wedge n}^2 &= \mathsf{E}\Big(\sum_{k=1}^n \mathbf{1}_{\{\tau\geq k\}} X_k^2\Big) \\ &= \mathsf{E}\Big(\sum_{k=1}^n \mathsf{E}(\mathbf{1}_{\{\tau\geq k\}} X_k^2 \,|\, \mathcal{F}_{k-1})\Big) \\ &= \mathsf{E}\Big(\sum_{k=1}^n \mathbf{1}_{\{\tau\geq k\}} \mathsf{E}(X_k^2 \,|\, \mathcal{F}_{k-1})\Big) \\ &= \mathsf{E}\Big(\sum_{k=1}^{\tau\wedge n} \mathsf{E}(X_k^2 \,|\, \mathcal{F}_{k-1}\Big) \leq K. \end{aligned}$$

By Jensen's inequality $(\mathsf{E}|S_n^\tau|)^2 \leq \mathsf{E}(S_n^\tau)^2 \leq K^2$. As a consequence S^τ is a martingale that is bounded in L^1. By the martingale convergence theorem, it converges a.s. to an integrable limit S_∞, say. Thus S_n converges a.s. on the event $\tau = \infty$, in other words, on the event $\sum_{k=1}^\infty \mathsf{E}(X_k^2 \,|\, \mathcal{F}_{k-1}) \leq K\}$. Let $K \uparrow \infty$. QED

Without proof we will now recall a simple but effective lemma.

Lemma 2.2.21 (Kronecker Lemma) *Let $\{x_n\}_{n\geq 1}$ be a sequence of real numbers such that $\sum_n x_n$ converges. Le $\{b_n\}_n$ be a non-decreasing sequence of positive constants with $b_n \uparrow \infty$ as $n \to \infty$. Then $b_n^{-1} \sum_{k=1}^n b_k x_k \to 0$, as $n \to \infty$.*

This allows to formulate the following version of a martingale Law of Large Numbers.

Theorem 2.2.22 *Let $(S_n = \sum_{k=1}^n X_k)_{n\geq 1}$ be a martingale with respect to the filtration $(\mathcal{F}_n)_{n\geq 1}$. Then $\sum_{k=1}^n X_k/k$ converges a.s. on the set $\{\sum_{k=1}^\infty k^{-2} \mathsf{E}(X_k^2 \,|\, \mathcal{F}_{k-1}) < \infty\}$. Hence $S_n/n \to 0$ a.s. on the set $\{\sum_{k=1}^\infty k^{-2} \mathsf{E}(X_k^2 \,|\, \mathcal{F}_{k-1}) < \infty\}$.*

Proof. Combine Theorem 2.2.6 and Lemma 2.2.21.

2.3 Continuous-time martingales

In this section we consider general martingales indexed by a subset T of $\mathbf{R}_+$. If the martingale $M = (M_t)_{t\geq 0}$ has 'nice' sample paths, for instance they are right-continuous, then M can be 'approximated' accurately by a discrete-time martingale. Simply choose a countable dense subset $\{t_n\}$ of the index set T and compare the continuous-times martingale M with the discrete-time martingale $(M_{t_n})_n$. This simple idea allows to transfer many of the discrete-time results to the continuous-time setting.

2.3.1 Upcrossings in continuous time

For a continuous-time process X we define the number of upcrossings of the interval $[a, b]$ in the *bounded* set of time points $T \subset \mathbf{R}_+$ as follows. For a finite set $F = \{t_1, \ldots, t_n\} \subseteq T$ we define $U_F[a, b]$ as the number of upcrossings of $[a, b]$ of the discrete-time process $(X_{t_i})_{i=1,\ldots,n}$ (see Definition 2.2.11). We put

$$U_T[a, b] = \sup\{U_F[a, b] \,|\, F \subseteq T, F \text{ finite }\}.$$

Doob's upcrossing lemma has the following extension.

Lemma 2.3.1 *Let M be a supermartingale and let $T \subseteq \mathbf{R}_+$ be a countable, bounded set. Then for all $a < b$, the number of upcrossings $U_T[a,b]$ of the interval $[a,b]$ by M satisfies*

$$(b-a)\mathsf{E}U_T[a,b] \leq \sup_{t\in T} \mathsf{E}(M_t - a)^-.$$

Proof. Let T_n be a nested sequence of finite sets, such that $U_T[a,b] = \lim_{n\to\infty} U_{T_n}[a,b]$. For every n, the discrete-time upcrossing inequality states that

$$(b-a)\mathsf{E}U_{T_n}[a,b] \leq \mathsf{E}(M_{t_n} - a)^-,$$

where t_n is the largest element of T_n. By the conditional version of Jensen's inequality the process $(M-a)^-$ is a submartingale (see Example 2.1.3). In particular, the function $t \to \mathsf{E}(M_t - a)^-$ is increasing, so

$$\mathsf{E}(M_{t_n} - a)^- = \sup_{t\in T_n} \mathsf{E}(M_t - a)^-.$$

So, for every n we have the inequality

$$(b-a)\mathsf{E}U_{T_n}[a,b] \leq \sup_{t\in T_n} \mathsf{E}(M_t - a)^-.$$

The proof is completed by letting n tend to infinity. QED

Hence $U_T[a,b]$ is a.s. finite! By Doob's upcrossing Lemma 2.2.12 $U_T[a,b]$ is $\mathcal{F}_t$-measurable if $t = \sup\{s \mid s \in T\}$.

2.3.2 Regularisation

We always consider the processes under consideration to be defined on an underlying filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F})t)_{t\in T}, \mathsf{P})$. Remind that $\mathcal{F}_\infty = \sigma(\mathcal{F}_t, t \in T)$. We will assume that $T = \mathbf{R}_+$. For shorthand notation, if we write $\lim_{q\downarrow(\uparrow)t}$, we mean the limit along non-increasing (non-decreasing) rational sequences converging to t. The same holds for lim sup / lim inf.

Theorem 2.3.2 *Let M be a supermartingale. Then there exists a set $\Omega_s^* \in \mathcal{F}_s$ of probability 1, such that for all $\omega \in \Omega_s^*$ the limits*

$$\lim_{q\uparrow t} M_q(\omega) \quad \text{and} \quad \lim_{q\downarrow t} M_q(\omega)$$

exist and are finite for every $t \in (0,s]$ and $t \in [0,s)$ respectively, for any $s \leq \infty$.

Is the set of discontinuities of each path $M(\omega) = (M_t(\omega))_{t\in\mathrm{R}_+}$, $\omega \in \Omega^*$, at most countable? This is not (yet) clear.

Proof. We give the proof for $s = \infty$. Fix $n \in \mathbf{Z}_+$. Let $a < b$, $a, b \in \mathsf{Q}$. By virtue of Lemma 2.3.1 there exists a set $\Omega_{n,a,b} \in \mathcal{F}_n$, of probability 1, such that

$$U_{[0,n]\cap\mathsf{Q}}[a,b](\omega) < \infty, \qquad \text{for all } \omega \in \Omega_{n,a,b}.$$

Put

$$\Omega_n = \bigcap_{a<b,\, a,b\in\mathsf{Q}} \Omega_{n,a,b}.$$

Then $\Omega_n \in \mathcal{F}_n$. Let now $t < n$ and suppose that

$$\lim_{q\downarrow t} M_q(\omega)$$

does not exist for some $\omega \in \Omega_n$. Then there exists $a < b$, $a, b \in \mathsf{Q}$, such that

$$\liminf_{q\downarrow t} M_q(\omega) < a < b < \limsup_{q\downarrow t} M_q(\omega).$$

Hence $U_{[0,n]\cap\mathsf{Q}}[a, b](\omega) = \infty$, a contradiction. It follows that $\lim_{q\downarrow t} M_q(\omega)$ exists for all $\omega \in \Omega_n$ and all $t \in [0, n)$.

A similar argument holds for the left limits: $\lim_{q\uparrow t} M_q(\omega)$ exists for all $\omega \in \Omega_n$ for all $t \in (0, n]$. It follows that on $\Omega' = \cap_n \Omega_n$ these limits exist in $[-\infty, \infty]$ for all $t > 0$ in case of left limits and for all $t \geq 0$ in case of right limits. Note that $\Omega' \in \mathcal{F}_\infty$ and $\mathsf{P}\{\Omega'\} = 1$.

We still have to show that the limits are in fact finite. Fix $t \in T$, $n > t$. Let $\mathsf{Q}_n = [0, n] \cap \mathsf{Q}$ and let $\mathsf{Q}_{m,n}$ be a nested sequence of finitely many rational numbers increasing to Q_n, all containing 0 and n. Then $(M_s)_{s\in\mathsf{Q}_{m,n}}$ is a discrete-time supermartingale. By virtue of Theorem 2.2.10

$$\lambda\mathsf{P}\{\max_{s\in\mathsf{Q}_{m,n}} |M_s| > 3\lambda\} \leq 4\mathsf{E}|M_0| + 3\mathsf{E}|M_n|.$$

Letting $m \to \infty$ and then $\lambda \to \infty$, by virtue of the monotone convergence theorem for sets

$$\sup_{s\in\mathsf{Q}_n} |M_s| < \infty, \quad \text{a.s.}$$

This implies that the limits are finite.

Put $\Omega''_n = \{\omega \mid \sup_{s\in\mathsf{Q}_n} |M_s|(\omega) < \infty\}$. By the above, $\Omega''_n \in \mathcal{F}_n$ and $\mathsf{P}\{\Omega''_n\} = 1$. Hence, $\Omega'' := \cap_n \Omega''_n$ is a set of probability 1, belonging to $\mathcal{F}_\infty$. Finally, set $\Omega^* = \Omega'' \cap \Omega'$. This is a set of probability 1, belonging to $\mathcal{F}_\infty$. QED

Corollary 2.3.3 *There exists an $\mathcal{F}_\infty$-measurable set Ω^*, $\mathsf{P}\{\Omega^*\} = 1$, such that every sample path of a right-continuous supermartingale is cadlag on Ω^*.*

Proof. See Exercise 2.22. QED

> Our aim is now to construct a modification of a supermartingale that is a supermartingale with a.s. cadlag sample paths itself, under suitable conditions. To this end read LN §1.6, definitions 1.6.2 (right-continuity of a filtration) and 1.6.3 (usual conditions)

Given a supermartingale M, define for every $t \geq 0$

$$M_{t^+}(\omega) = \begin{cases} \lim_{q\downarrow t, q\in\mathsf{Q}} M_q(\omega), & \text{if this limit exists and is finite} \\ 0, & \text{otherwise.} \end{cases}$$

The random variables M_{t^+} are well-defined by Theorem 2.3.2. By inspection of the proof of this theorem, one can check that M_{t^+} is $\mathcal{F}_{t^+}$-measurable.

We have the following result concerning the process $(M_{t^+})_{t\geq 0}$.

Lemma 2.3.4 *Let M be a supermartingale.*

i) *Then $\mathsf{E}|M_{t+}| < \infty$ for every t and*

$$\mathsf{E}(M_{t+} \mid \mathcal{F}_t) \leq M_t, \quad \text{a.s.}$$

If in addition $t \to \mathsf{E}M_t$ is right-continuous, then this inequality is an equality.

ii) *The process $(M_{t+})_{t\geq 0}$ is a supermartingale with respect to the filtration $(\mathcal{F}_{t+})_{t\geq 0}$ and it is a martingale if M is a martingale.*

Proof. Fix $t \geq 0$. Let $q_n \downarrow t$ be a sequence of rational numbers decreasing to t. Then the process $(M_{q_n})_n$ is a backward discrete-time supermartingale like we considered in Theorem 2.2.16, with $\sup_n \mathsf{E}M_{q_n} \leq \mathsf{E}M_t < \infty$. By that theorem, M_{t+} is integrable, and $M_{q_n} \stackrel{\mathsf{L}^1}{\to} M_{t+}$. As in the proof of Theorem 2.2.14, L^1-convergence allows to take the limit $n \to \infty$ in the inequality

$$\mathsf{E}(M_{q_n} \mid \mathcal{F}_t) \leq M_t \quad \text{a.s.}$$

yielding

$$\mathsf{E}(M_{t+} \mid \mathcal{F}_t) \leq M_t \quad \text{a.s.}$$

L^1 convergence also implies that $\mathsf{E}M_{q_n} \to \mathsf{E}M_{t+}$. So, if $s \to \mathsf{E}M_s$ is right-continuous, then $\mathsf{E}M_{t+} = \lim_{n\to\infty} \mathsf{E}M_{q_n} = \mathsf{E}M_t$. But this implies (cf. Exercise 2.7 b) that $\mathsf{E}M_{t+} \mid \mathcal{F}_t) = M_t$ a.s.

To prove the final statement, let $s < t$ and let $q'_n \leq t$ be a sequence of rational numbers decreasing to s. Then

$$\mathsf{E}(M_{t+} \mid \mathcal{F}_{q'_n}) = \mathsf{E}(\mathsf{E}(M_{t+} \mid \mathcal{F}_t)|\mathcal{F}_{q'_n}) \leq \mathsf{E}(M_t \mid \mathcal{F}_{q'_n}) \leq M_{q'_n} \quad \text{a.s.}$$

with equality if M is a martingale. The right-hand side of inequality converges to M_{s+} as $n \to \infty$. The process $\mathsf{E}(M_{t+} \mid \mathcal{F}_{q'_n})$ is a backward martingale satisfying the conditions of Theorem 2.2.16. Hence, $\mathsf{E}(M_{t+} \mid \mathcal{F}_{q'_n})$ converges to $\mathsf{E}(M_{t+} \mid \mathcal{F}_{s+})$ a.s. QED

We can now prove the main regularisation theorem for supermartingales with respect to filtrations satisfying the usual conditions.

The idea is the following. In essence we want ensure that all sample paths are cadlag a priori (that the cadlag paths are a measurable set w.r.t to each σ-algebra $\mathcal{F}_t$). In view of Corollary 2.3.3 this requires to complete $\mathcal{F}_0$ with all P-null sets in $\mathcal{F}_\infty$. On the other hand, we want to make out of $(M_{t+})_t$ a cadlag modification of M. This is guaranteed if the filtration involved is right-continuous.

> Can one enlarge a given filtration to obtain a filtration satisfying the usual conditions? If yes, can this procedure destroy properties of interest - like the supermartingale property, independence properties? For some details on this issue see BN §9.

Theorem 2.3.5 (Doob's regularity theorem) *Let M be a supermartingale with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$ satisfying the usual conditions. Then M has a cadlag modification $\widetilde{M}$ (such that $\{M_t - \widetilde{M}_t \neq 0\}$ is a null set contained in $\mathcal{F}_0$) if and only if $t \to \mathsf{E}M_t$ is right-continuous. In that case $\widetilde{M}$ is a supermartingale with respect to $(\mathcal{F}_t)_{t\geq 0}$ as well. If M is a martingale then $\widetilde{M}$ is a martingale.*

Proof. '⇐' By Theorem 2.3.2 and the fact that $(\mathcal{F}_t)$ satisfies the usual conditions, there exists an event $\Omega^* \in \mathcal{F}_0$ (!) of probability 1, on which the limits

$$M_{t^-} = \lim_{q \uparrow t} M_q, \quad M_{t^+} = \lim_{q \downarrow t} M_q$$

exist for every t. Define the process $\widetilde{M}$ by $\widetilde{M}_t(\omega) = M_{t^+}(\omega)\mathbf{1}_{\{\Omega^*\}}(\omega)$ for $t \geq 0$. Then $\widetilde{M}_t = M_{t^+}$ a.s. and they differ at most on a null-set contained in $\mathcal{F}_0$. Since M_{t^+} is $\mathcal{F}_{t^+}$-measurable, we have that $\widetilde{M}_t$ is $\mathcal{F}_{t^+}$-measurable. It follows that $\widetilde{M}_t = \mathsf{E}(M_{t^+} \,|\, \mathcal{F}_{t^+})$ a.s. By right-continuity of the filtration, right-continuity of the map $t \to \mathsf{E}M_t$ and the preceding lemma, we get that a.s. $\widetilde{M}_t = \mathsf{E}(M_{t^+} \,|\, \mathcal{F}_{t^+}) = \mathsf{E}(M_{t^+} \,|\, \mathcal{F}_t) = M_t$ a.s. In other words, $\widetilde{M}$ is a modification of M. Right-continuity of the filtration implies further that $\widetilde{M}$ is adapted to $(\mathcal{F}_t)$. The process $\widetilde{M}$ is cadlag as well as a supermartingale (see Exercise 2.23).

'⇒' see Exercise 2.24. QED

Corollary 2.3.6 *A martingale with respect to a filtration that satisfies the usual conditions has a cadlag modification, which is a martingale w.r.t the same filtration.*

We will next give two example showing what can go wrong without right-continuity of the filtration, or without continuity of the expectation as a function of time.

Example
Let $\Omega = \{-1, 1\}$, $\mathcal{F}_t = \{\Omega, \emptyset\}$ for $t \leq 1$ and $\mathcal{F}_t = \{\Omega, \emptyset, \{1\}, \{-1\}\}$ for $t > 1$. Let $\mathsf{P}(\{1\}) = \mathsf{P}(\{-1\}) = 1/2$.

Note that $\mathcal{F}_t$ is *not right-continuous*, since $\mathcal{F}_1 \neq \mathcal{F}_{1^+}$! Define

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

Now, $Y = (Y_t)_t$ is a martingale, but it is not right-continuous, whereas $X = (X_t)_t$ is a right-continuous process. One does have that $\mathsf{E}Y_t = 0$ is a right-continuous function of t.

Moreover, $Y_{t^+} = X_t$ and $\mathsf{P}\{X_1 = Y_1\} = 0$. Hence X is not a cadlag modification of Y, and in particular Y cannot have a cadlag modification.

By Lemma 2.3.4 it follows that $\mathsf{E}(X_t \,|\, \mathcal{F}_t) = Y_t$, so that X cannot be a martingale w.r.t the filtration $(\mathcal{F}_t)_t$. By the same lemma, X is a right-continuous martingale, w.r.t to $(\mathcal{F}_{t^+})_t$. On the other hand, Y is *not* a martingale w.r.t. to $(\mathcal{F}_{t^+})_t$!

Example
Let $\Omega = \{-1, 1\}$, $\mathcal{F}_t = \{\Omega, \emptyset\}$ for $t < 1$ and $\mathcal{F}_t = \{\Omega, \emptyset, \{1\}, \{-1\}\}$ for $t \geq 1$. Let $\mathsf{P}(\{1\}) = \mathsf{P}(\{-1\}) = 1/2$. Define

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

In this case the filtration $(\mathcal{F}_t)_t$ is right-continuous and Y and X are both supermartingales w.r.t $(\mathcal{F}_t)_t$. Furthermore $X_t = \lim_{q \downarrow t} Y_q$ for $t \geq 0$, but $\mathsf{P}\{X_1 = Y_1\} = 0$ and hence X is not a modification of Y.

2.3.3 Convergence theorems

In view of the results of the previous section, we will only consider *everywhere right-continuous martingales* from this point on. Under this assumption, many of the discrete-time theorems can be generalised to continuous time.

Theorem 2.3.7 *Let M be a right-continuous supermartingale that is bounded in L^1. Then M_t converges a.s. to a finite $\mathcal{F}_\infty$-measurable limit M_∞, as $t \to \infty$, with $\mathsf{E}|M_\infty| < \infty$.*

Proof. The first step to show is that we can restrict to take a limit along rational time-sequences. In other words, that $M_t \to M_\infty$ a.s. as $t \to \infty$ if and only if

$$\lim_{q\to\infty} M_q = M_\infty \quad \text{a.s.} \tag{2.3.1}$$

To prove the non-trivial implication in this assertion, assume that (2.3.1) holds. Fix $\epsilon > 0$ and $\omega \in \Omega$ for which $M_q(\omega) \to M_\infty(\omega)$. Then there exists a number $a = a_{\omega,\epsilon} > 0$ such that $|M_q(\omega) - M_\infty(\omega)| < \epsilon$ for all $q > a$. Now let $t > a$ be arbitrary. Since M is right-continuous, there exists $q' > t$ such that $|M_{q'}(\omega) - M_t(\omega)| < \epsilon$. By the triangle inequality, it follows that $|M_t(\omega) - M_\infty(\omega)| \leq |M_{q'}(\omega) - M_\infty(\omega)| + |M_t(\omega) - M_{q'}(\omega)| < 2\epsilon$. This proves that $M_t(\omega) \to M_\infty(\omega)$, $t \to \infty$.

To prove convergence to a finite $\mathcal{F}_\infty$-measurable, integrable limit, we may assume that M is indexed by the countable set Q_+. The proof can now be finished by arguing as in the proof of Theorem 2.2.13, replacing Doob's discrete-time upcrossing inequality by Lemma 2.3.1. QED

Corollary 2.3.8 *A non-negative, right-continuous supermartingale M converges a.s. as $t \to \infty$, to a finite, integrable, $\mathcal{F}_\infty$-measurable random variable.*

Proof. Simple consequence of Theorem 2.3.7. QED

> The following continuous-time extension of Theorem 2.2.14 can be derived by reasoning as in discrete-time. The only slight difference is that for a *continuous-time* process X L^1-convergence of X_t as $t \to \infty$ need not imply that X is UI.

Theorem 2.3.9 *Let M be a right-continuous supermartingale that is bounded in L^1.*

i) *If M is uniformly integrable, then $M_t \to M_\infty$ a.s. and in L^1, and*

$$\mathsf{E}(M_\infty \mid \mathcal{F}_t) \leq M_t \quad \text{a.s.}$$

with equality if M is a martingale.

ii) *If M is a martingale and $M_t \to M_\infty$ in L^1 as $t \to \infty$, then M is uniformly integrable.*

Proof. See Exercise 2.25. QED

2.3.4 Inequalities

Dobb's submartingale inequality and L^p-inequality are very easily extended to the setting of general right-continuous martingales.

Theorem 2.3.10 (Doob's submartingale inequality) *Let M be a right-continuous submartingale. Then for all $\lambda > 0$ and $t \geq 0$*

$$\mathsf{P}\{\sup_{s\leq t} M_s \geq \lambda\} \leq \tfrac{1}{\lambda}\mathsf{E}|M_t|.$$

Proof. Let T be a countable, dense subset of $[0,t]$ and choose an increasing sequence of finite subsets $T_n \subseteq T$, $0, t \in T_n$ for every n and $T_n \uparrow T$ as $n \to \infty$. By right-continuity of M we have that

$$\sup_n \max_{s\in T_n} M_s = \sup_{s\in T} M_s = \sup_{s\in[0,t]} M_s.$$

This implies that $\{\max_{s\in T_n} M_s > c\} \uparrow \{\sup_{s\in T} M_s > c\}$ and so by monotone convergence of sets $\mathsf{P}\{\max_{s\in T_n} M_s > c\} \uparrow \mathsf{P}\{\sup_{s\in T} M_s > c\}$. By the discrete-time version of the submartingale inequality for each $m > 0$ sufficiently large

$$\begin{aligned}\mathsf{P}\{\sup_{s\in[0,t]} M_s > \lambda - \tfrac{1}{m}\} &= \mathsf{P}\{\sup_{s\in T} M_s > \lambda - \tfrac{1}{m}\} \\ &= \lim_{n\to\infty} \mathsf{P}\{\max_{s\in T_n} M_s > \lambda - \tfrac{1}{m}\} \\ &\leq \tfrac{1}{\lambda - 1/m}\mathsf{E}|M_t|.\end{aligned}$$

Let m tend to infinity. QED

By exactly the same reasoning, we can generalise the L^p-inequality to continuous time.

Theorem 2.3.11 (Doob's L^p-inequality) *Let M be a right-continuous martingale or a right-continuous, nonnegative submartingale. Then for all $p > 1$ and $t \geq 0$*

$$\mathsf{E}\Big(\sup_{s\leq t} |M_s|^p\Big) \leq \Big(\frac{p}{p-1}\Big)^p \mathsf{E}|M_t|^p.$$

2.3.5 Optional sampling

We will now discuss the continuous-time version of the optional stopping theorem.

Theorem 2.3.12 (Optional sampling theorem) *Let M be a right-continuous, uniformly integrable supermartingale. Then for all stopping times $\sigma \leq \tau$ we have that M_τ and M_σ are integrable and*

$$\mathsf{E}(M_\tau \mid \mathcal{F}_\sigma) \leq M_\sigma \quad \text{a.s.}.$$

with equality if M is martingale.

Proof. By Lemma 1.6.17 there exist stopping times σ_n and τ_n taking only finitely many values, such that $\sigma_n \leq \tau_n$, and $\sigma_n \downarrow \sigma$, $\tau_n \downarrow \tau$.

By the discrete-time optional sampling theorem applied to the supermartingale $(M_{k/2^n})_{k\in\mathbf{Z}_+}$ (it is uniformly integrable!)

$$\mathsf{E}(M_{\tau_n} \mid \mathcal{F}_{\sigma_n}) \leq M_{\sigma_n} \quad \text{a.s.}$$

Since $\sigma \leq \sigma_n$ it holds that $\mathcal{F}_\sigma \subseteq \mathcal{F}_{\sigma_n}$. It follows that

$$\mathsf{E}(M_{\tau_n} \mid \mathcal{F}_\sigma) = \mathsf{E}\Big(\mathsf{E}(M_{\tau_n} \mid \mathcal{F}_{\sigma_n}) \mid \mathcal{F}_\sigma\Big) \leq \mathsf{E}(M_{\sigma_n} \mid \mathcal{F}_\sigma) \quad \text{a.s.} \tag{2.3.2}$$

Similarly

$$\mathsf{E}(M_{\tau_n} \mid \mathcal{F}_{\tau_{n+1}}) \leq M_{\tau_{n+1}} \quad \text{a.s.}$$

Hence, $(M_{\tau_n})_n$ is a 'backward' supermartingale in the sense of the Lévy-Doob downward theorem 2.2.16. Since $\sup_n \mathsf{E}M_{\tau_n} \leq \mathsf{E}M_0$, this theorem implies that $(M_{\tau_n})_n$ is uniformly integrable and it converges a.s. and in L^1. By right-continuity $M_{\tau_n} \to M_\tau$ a.s. Hence $M_{\tau_n} \to M_\tau$ in L^1. Similarly $M_{\sigma_n} \to M_\sigma$ in L^1. Now take $A \in \mathcal{F}_\sigma$. By equation (2.3.2) it holds that

$$\int_A M_{\tau_n} d\mathsf{P} \leq \int_A M_{\sigma_n} d\mathsf{P}.$$

By L^1-convergence, this yields

$$\int_A M_\tau d\mathsf{P} \leq \int_A M_\sigma d\mathsf{P},$$

if we let n tend to infinity. This completes the proof. QED

So far we have not yet addressed the question whether stopped (sub-, super-) martingales are (sub-, super-) martingales. The next theorem prepares the way to prove this.

Theorem 2.3.13 *A right-continuous, adapted process M is a supermartingale (resp. a martingale) if and only if for all bounded stopping times τ, σ with $\sigma \leq \tau$, the random variables M_τ and M_σ are integrable and $\mathsf{E}M_\tau \leq \mathsf{E}M_\sigma$ (resp. $\mathsf{E}M_\tau = \mathsf{E}M_\sigma$).*

Proof. Suppose that M is a supermartingale. Since τ is bounded, there exists a constant $K > 0$ such that $\tau \leq K$ a.s.

As in the construction in the proof of Lemma 1.6.17 there exist stopping times $\tau_n \downarrow \tau$ and $\sigma_n \downarrow \sigma$ that are bounded by K and take finitely many values. In particular $\tau_n, \sigma_n \in D_n = \{K \cdot k \cdot 2^{-n}, k = 0, 1, \ldots, 2^n\}$. Note that $(D_n)_n$ is an increasing sequence of sets.

By bounded optional stopping for discrete-time martingales, we have that $\mathsf{E}(M_{\tau_n} \mid \mathcal{F}_{\tau_{n+1}}) \leq M_{\tau_{n+1}}$, a.s. (consider M restricted to the discrete time points $\{k \cdot 2^{-(n+1)}), k \in \mathbf{Z}_+\}$. We can apply the Lévy-Doob downward theorem 2.2.16, to obtain that $(M_{\tau_n})_n$ is a uniformly integrable supermartingale, converging a.s. and in L^1 to an integrable limit as in the proof of the previous theorem. By right-continuity the limit is M_τ. Analogously, we obtain that M_σ is integrable.

Bounded optional stopping for discrete-time supermartingales similarly yields that $\mathsf{E}(M_{\tau_n} \mid M_{\sigma_n}) \leq M_{\sigma_n}$ a.s. Taking expectations and using L^1-convergence proves that $\mathsf{E}M_\tau \leq \mathsf{E}M_\sigma$.

The reverse statement is proved by arguing as in the proof of Theorem 2.2.5. Let $s \leq t$ and let $A \in \mathcal{F}_s$. Choose stopping times $\sigma = s$ and $\tau = \mathbf{1}_{\{A\}}t + \mathbf{1}_{\{A^c\}}s$. QED

Corollary 2.3.14 *If M is a right-continuous (super)martingale and τ is an $(\mathcal{F}_t)_t$ stopping time, then the stopped process is a (super)martingale as well.*

Proof. Note that M^τ is right-continuous. By Lemmas 1.6.14 and 1.6.16 it is adapted.

Let $\sigma \leq \xi$ be bounded stopping times. By applying the previous theorem to the supermartingale M, and using that $\sigma \wedge \tau$, $\xi \wedge \tau$ are bounded stopping times, we find that

$$\mathsf{E} M^\tau_\sigma = \mathsf{E} M_{\tau\wedge\sigma} \leq \mathsf{E} M_{\tau\wedge\xi} = \mathsf{E} M^\tau_\xi.$$

Since σ and ξ were arbitrary bounded stopping times, another application of the previous theorem yields the desired result. QED

Just as in discrete time the assumption of uniform integrability is crucial for the optional sampling theorem. If this condition is dropped, we only have an inequality in general. Theorem 2.2.19 carries over to continuous time by using the same arguments as in the proof of Theorem 2.3.12.

Theorem 2.3.15 *Let M be a right-continuous non-negative supermartingale and let $\sigma \leq \tau$ be stopping times. Then*

$$\mathsf{E}(M_\tau \mid \mathcal{F}_\sigma) \leq M_\sigma, \quad \text{a.s.}$$

A consequence of this result is that non-negative right-continuous supermartingales stay at zero once they have hit it.

Corollary 2.3.16 *Let M be a non-negative, right-continuous supermartingale and define $\tau = \inf\{t \mid M_t = 0 \text{ or } M_{t^-} = 0\}$. Then $\mathsf{P}\{M_{\tau+t}\mathbf{1}_{\{\tau<\infty\}} = 0, t \geq 0\} = 1$, where $M_{t^-} = \limsup_{s\uparrow t} M_s$.*

Proof. Positive supermartingales are bounded in L^1. By Theorem 2.3.7, M converges a.s. to an integrable limit M_∞ say.

Note that $\limsup_{s\uparrow t} M_s = 0$ implies that $\lim_{s\uparrow t} M_s$ exists and equals 0. Hence by right-continuity $\tau(\omega) \leq t$ if and only if $\inf\{M_q(\omega) \mid q \in \mathsf{Q} \cap [0,t]\} = 0$ or $M_t(\omega) = 0$. Hence τ is a stopping time. Define

$$\tau_n = \inf\{t \mid M_t < n^{-1}\}.$$

Then τ_n is a stopping time with $\tau_n \leq \tau$. Furthermore, for all $q \in \mathsf{Q}_+$, we have that $\tau + q$ is a stopping time. Then by the foregoing theorem

$$\mathsf{E} M_{\tau+q} \leq \mathsf{E} M_{\tau_n} \leq \tfrac{1}{n}\mathsf{P}\{\tau_n < \infty\} + \mathsf{E} M_\infty \mathbf{1}_{\{\tau_n=\infty\}}.$$

On the other hand, since $\tau_n = \infty$ implies $\tau + q = \infty$ we have

$$\mathsf{E} M_\infty \mathbf{1}_{\{\tau_n=\infty\}} \leq \mathsf{E} M_\infty \mathbf{1}_{\{\tau+q=\infty\}}.$$

Combination yields for all $n \in \mathbf{Z}_+$ and all $q \in \mathsf{Q}_+$

$$\begin{aligned} \mathsf{E} M_{\tau+q}\mathbf{1}_{\{\tau+q<\infty\}} &= \mathsf{E} M_{\tau+q} - \mathsf{E} M_\infty \mathbf{1}_{\{\tau+q=\infty\}} \\ &\leq \tfrac{1}{n}\mathsf{P}\{\tau_n < \infty\}. \end{aligned}$$

Taking the limit $n \to \infty$ yields $\mathsf{E} M_{\tau+q}\mathbf{1}_{\{\tau+q<\infty\}} = 0$ for all $q \in \mathsf{Q}_+$. By non-negativity of M we get that $M_{\tau+q}\mathbf{1}_{\{\tau+q<\infty\}} = 0$ a.s. for all $q \in \mathsf{Q}_+$. But then also $\mathsf{P}\{\cap_q\{M_{\tau+q} = 0\} \cup \{\tau < \infty\}\} = 1$. By right continuity

$$\cap_q\{M_{\tau+q} = 0\} \cap \{\tau < \infty\} = \{M_{\tau+t} = 0, \tau < \infty, t \geq 0\}.$$

Note that the set where $M_{\tau+t} = 0$ belongs to $\mathcal{F}_\infty$! QED

2.4 Applications to Brownian motion

In this section we apply the developed theory to the study of Brownian motion.

2.4.1 Quadratic variation

The following result extends the result of Exercise 1.15 of Chapter 1.

Theorem 2.4.1 *Let W be a Brownian motion and fix $t > 0$. For $n \in \mathbf{Z}_+$, let π_n be a partition of $[0, t]$ given by $0 = t_0^n \leq t_1^n \leq \cdots \leq t_{k_n}^n = t$ and suppose that the mesh $\|\pi_n\| = \max_k |t_k^n - t_{k-1}^n|$ tends to zero as $n \to \infty$. Then*

$$\sum_k (W_{t_k^n} - W_{t_{k-1}^n})^2 \xrightarrow{\mathsf{L}^2} t, \quad t \to \infty.$$

If the partitions are nested we have

$$\sum_k (W_{t_k^n} - W_{t_{k-1}^n})^2 \xrightarrow{\text{a.s.}} t, \quad t \to \infty.$$

Proof. For the first statement see Exercise 1.15 in Chapter 1. To prove the second one, denote the sum by X_n and put $\mathcal{F}_n = \sigma(X_n, X_{n+1}, \ldots)$. Then $\mathcal{F}_{n+1} \subset \mathcal{F}_n$ for every $n \in \mathbf{Z}_+$. Now suppose that we can show that $\mathsf{E}(X_n \mid \mathcal{F}_{n+1}) = X_{n+1}$ a.s. Then, since $\sup \mathsf{E} X_n < \infty$, the Lévy-Doob downward theorem 2.2.16 implies that X_n converges a.s. to a finite limit X_∞. By the first statement of the theorem the X_n converge in probability to t. Hence, we must have $X_\infty = t$ a.s.

So it remains to prove that $\mathsf{E}(X_n \mid \mathcal{F}_{n+1}) = X_{n+1}$ a.s. Without loss of generality, we assume that the number of elements of the partition π_n equals n. In that case, there exists a sequence t_n such that the partition π_n has the numbers $t_1, \ldots, t_n$ as its division points: the point t_n is added to π_{n-1} to form the next partition π_n. Now fix n and consider the process W' defined by

$$W'_s = W_{s\wedge t_{n+1}} - (W_s - W_{s\wedge t_{n+1}}).$$

By Exercise 1.12 of Chapter 1, W' is again a BM. For W', denote the analogous sums X_k by X'_k. Then it is easily seen for $k \geq n+1$ that $X'_k = X_k$. Moreover, it holds that $X'_n - X'_{n+1} = X_{n+1} - X_n$ (check!). Since both W and W' are BM's, the sequences $(X_1, X_2, \ldots)$ and $(X'_1, X'_2, \ldots)$ have the same distribution. It follows that a.s.

$$\begin{aligned}
\mathsf{E}(X_n - X_{n+1} \mid \mathcal{F}_{n+1}) &= \mathsf{E}(X'_n - X'_{n+1} \mid X'_{n+1}, X'_{n+2}, \ldots) \\
&= \mathsf{E}(X'_n - X'_{n+1} \mid X_{n+1}, X_{n+2}, \ldots) \\
&= \mathsf{E}(X_{n+1} - X_n \mid X_{n+1}, X_{n+2}, \ldots) \\
&= -\mathsf{E}(X_{n+1} - X_n \mid \mathcal{F}_{n+1}).
\end{aligned}$$

This implies that $\mathsf{E}(X_n - X_{n+1} \,|\, \mathcal{F}_{n+1}) = 0$ a.s. QED

A real-valued function f is said to be of *finite variation* on an interval $[a,b]$, if there exists a finite number $K > 0$, such that for every finite partition $a = t_0 < \cdots < t_n = b$ of $[a,b]$ it holds that

$$\sum_k |f(t_k) - f(t_{k-1})| < K.$$

Roughly speaking, this means that the graph of the function f on $[a,b]$ has finite length. Theorem 2.4.1 shows that the sample paths of BM have positive, finite quadratic variation. This has the following consequence.

Corollary 2.4.2 *Almost every sample path of* BM *has unbounded variation on every interval.*

Proof. Fix $t > 0$. Let π_n be nested partitions of $[0,t]$ given by $0 = t_0^n \le t_1^n \le \cdots \le t_{k_n}^n = t$. Suppose that the mesh $\|\pi_n\| = \max_k |t_k^n - t_{k-1}^n| \to 0$ as $n \to \infty$. Then

$$\sum_k (W_{t_k^n} - W_{t_{k-1}^n})^2 \le \max_k |W_{t_k^n} - W_{t_{k-1}^n}| \cdot \sum_k |W_{t_k^n} - W_{t_{k-1}^n}|.$$

By uniform continuity of Brownian sample paths, the first factor on the right-hand side converges to zero a.s., as $n \to \infty$. Hence, if the Brownian motion would have finite variation on $[0,t]$ with positive probability, then $\sum_k (W_{t_k^n} - W_{t_{k-1}^n})^2$ would converge to 0 with positive probability. This contradicts Theorem 2.4.1. QED

2.4.2 Exponential inequality

Let W be a Brownian motion. We have the following exponential inequality for the tail properties of the running maximum of the Brownian motion.

Theorem 2.4.3 *For every $t \ge 0$ and $\lambda > 0$*

$$\mathsf{P}\Big\{\sup_{s \le t} W_s \ge \lambda\Big\} \le e^{-\lambda^2/2t}$$

and

$$\mathsf{P}\Big\{\sup_{s \le t} |W_s| \ge \lambda\Big\} \le 2e^{-\lambda^2/2t}$$

Proof. For $a > 0$ consider the exponential martingale M defined by $M_t = \exp\{aW_t - a^2 t/2\}$ (see Example 2.1.4). Observe that

$$\mathsf{P}\{\sup_{s \le t} W_s \ge \lambda\} \le \mathsf{P}\Big\{\sup_{s \le t} M_s \ge e^{a\lambda - a^2 t/2}\Big\}.$$

By the submartingale inequality, the probability on the right-hand side is bounded by

$$e^{a^2 t/2 - a\lambda} \mathsf{E} M_t = e^{a^2 t/2 - a\lambda} \mathsf{E} M_0 = e^{a^2 t/2 - a\lambda}.$$

The proof of the first inequality is completed by minimising the latter expression in $a > 0$. To prove the second one, note that

$$\begin{aligned}\mathsf{P}\Big\{\sup_{s\le t}|W_s|\ge\lambda\Big\} &\le \mathsf{P}\{\sup_{s\le t}W_s\ge\lambda\}+\mathsf{P}\{\inf_{s\le t}W_s\le-\lambda\}\\ &= \mathsf{P}\{\sup_{s\le t}W_s\ge\lambda\}+\mathsf{P}\{\sup_{s\le t}-W_s\ge\lambda\}.\end{aligned}$$

The proof is completed by applying the first inequality to the BM's W and $-W$. QED

The exponential inequality also follows from the fact that $\sup_{s\le t}W_s \stackrel{\mathsf{d}}{=} |W_t|$ for every fixed t. We will prove this equality in distribution in the next chapter.

2.4.3 The law of the iterated logarithm

The law of the iterated logarithm describes how BM oscillates near zero and infinity. In the proof we will need the following simple lemma.

Lemma 2.4.4 *For every $a > 0$*

$$\int_a^\infty e^{-x^2/2}dx \ge \frac{a}{1+a^2}e^{-a^2/2}.$$

Proof. The proof starts from the inequality

$$\int_a^\infty \frac{1}{x^2}e^{-x^2/2}dx \le \frac{1}{a^2}\int_a^\infty e^{-x^2/2}dx.$$

Integration by parts shows that the left-hand side equals

$$\begin{aligned}-\int_a^\infty e^{-x^2/2}d(\frac{1}{x}) &= \frac{1}{a}e^{-a^2/2}+\int_a^\infty \frac{1}{x}d(e^{-x^2/2})\\ &= \frac{1}{a}e^{-a^2/2}-\int_a^\infty e^{-x^2/2}dx.\end{aligned}$$

Hence we find that

$$\Big(1+\frac{1}{a^2}\Big)\int_a^\infty e^{-x^2/2}dx \ge \frac{1}{a}e^{-a^2/2}.$$

Thid finishes the proof. QED

Theorem 2.4.5 (Law of the iterated logarithm) *It almost surely holds that*

$$\limsup_{t\downarrow 0}\frac{W_t}{\sqrt{2t\log\log 1/t}}=1,\qquad \liminf_{t\downarrow 0}\frac{W_t}{\sqrt{2t\log\log 1/t}}=-1,$$

$$\limsup_{t\to\infty}\frac{W_t}{\sqrt{2t\log\log t}}=1,\qquad \liminf_{t\to\infty}\frac{W_t}{\sqrt{2t\log\log t}}=-1.$$

Proof. It suffices to prove the first statement. The second follows by applying the first to the BM $-W$. The third and fourth statements follow by applying the first two to the BM $tW_{1/t}$ (cf. Theorem 1.4.4).

Put $h(t) = \sqrt{2t \log\log 1/t}$. We will first prove that $\limsup_{s\downarrow 0} W(s)/h(s) \le 1$. Choose two numbers $\theta, \delta \in (0,1)$. We put

$$\alpha_n = (1+\delta)\theta^{-n} h(\theta^n), \quad \beta_n = h(\theta^n)/2.$$

Use the submartingale inequality applied to the exponential martingale $M_s = \exp\{\alpha_n W_s - \alpha_n^2 s/2\}$:

$$\begin{aligned}
\mathsf{P}\{\sup_{s\le 1}(W_s - \alpha_n s/2) \ge \beta_n\} &= \mathsf{P}\{\sup_{s\le 1} M_s \ge e^{\alpha_n\beta_n}\} \\
&\le e^{-\alpha_n\beta_n}\mathsf{E}M_1 = e^{-\alpha_n\beta_n} \\
&\le K_\theta n^{-(1+\delta)},
\end{aligned}$$

for some constant $K_\theta > 0$ that depends on θ but not on n. Applying the Borel-Cantelli lemma yields the existence of a set $\Omega^*_{\theta,\delta} \in \mathcal{F}_\infty$, $\mathsf{P}\{\Omega^*_{\theta,\delta}\} = 1$, such that for each $\omega \in \Omega^*_{\theta,\delta}$ there exists n_ω such that

$$\sup_{s\le 1}(W_s(\omega) - \alpha_n s/2) \le \beta_n,$$

for $n \ge n_\omega$.

One can verify that h is increasing for $t \in (0, e^{-c}]$ with c satisfying $c = e^{1/c}$. Hence, for $n \ge n_\omega$ and $s \in [\theta^n, \theta^{n-1}]$

$$W_s(\omega) \le \frac{\alpha_n s}{2} + \beta_n \le \frac{\alpha_n \theta^{n-1}}{2} + \beta_n = \Big(\frac{1+\delta}{2\theta} + \frac{1}{2}\Big)h(\theta^n) \le \Big(\frac{1+\delta}{2\theta} + \frac{1}{2}\Big)h(s).$$

It follows that for all $n \ge n_\omega \vee -c/\log\theta$

$$\sup_{\theta^n \le s \le \theta^{n-1}} \frac{W_s(\omega)}{h(s)} \le \Big(\frac{1+\delta}{2\theta} + \frac{1}{2}\Big).$$

It follows that $\limsup_{s\downarrow 0} W_s(\omega)/h(s) \le \Big(\frac{1+\delta}{2\theta} + \frac{1}{2}\Big)$, for $\omega \in \Omega^*_{\theta,\delta}$. Write $\Omega^*_m = \Omega^*_{\theta=1-1/m,\delta=1/m}$ and put $\Omega^* = \cap_m \Omega^*_m$. Then $\limsup_{s\downarrow 0} W_s(\omega)/h(s) \le 1$ for $\omega \in \Omega^*$.

To prove the reverse inequality, choose $\theta \in (0,1)$ and consider the events

$$A_n = \{W_{\theta^n} - W_{\theta^{n+1}} \ge (1-\sqrt{\theta})h(\theta^n)\}.$$

By the independence of the incements of BM, the events A_n are independent. Note that

$$\frac{W_{\theta^n} - W_{\theta^{n+1}}}{\sqrt{\theta^n - \theta^{n+1}}} \stackrel{\mathsf{d}}{=} \mathsf{N}(0,1).$$

Hence,

$$\begin{aligned}
\mathsf{P}\{A_n\} &= \mathsf{P}\Big\{\frac{W_{\theta^n} - W_{\theta^{n+1}}}{\sqrt{\theta^n - \theta^{n+1}}} \ge \frac{(1-\sqrt{\theta})h(\theta^n)}{\sqrt{\theta^n - \theta^{n+1}}}\Big\} \\
&= \frac{1}{\sqrt{2\pi}} \int_a^\infty e^{-x^2/2} dx,
\end{aligned}$$

with

$$a = \frac{(1-\sqrt{\theta})h(\theta^n)}{\sqrt{\theta^n - \theta^{n+1}}} = (1-\sqrt{\theta})\sqrt{\frac{2\log\log\theta^{-n}}{1-\theta}}.$$

By Lemma 2.4.4 it follows that

$$\sqrt{2\pi}\mathsf{P}\{A_n\} \geq \frac{a}{1+a^2}e^{-a^2/2}.$$

It is easily seen that the right-hand side is of order

$$n^{-\frac{(1-\sqrt{\theta})^2}{1-\theta}} = n^{-\alpha},$$

with $\alpha < 1$. It follows that $\sum_n \mathsf{P}\{A_n\} = \infty$ and so by the 2d Borel-Cantelli Lemma there exists a set $\Omega_\theta \in \mathcal{F}_\infty$, $\mathsf{P}\{\Omega_\theta\} = 1$, such that

$$W_{\theta^n}(\omega) \geq (1-\sqrt{\theta})h(\theta^n) + W_{\theta^{n+1}}(\omega)$$

for infinitely many n, for all $\omega \in \Omega_\theta$. Since $-W$ is also a BM, the first part of the proof implies the existence of a set $\Omega_0 \in \mathcal{F}_\infty$, $\mathsf{P}\{\Omega_0\} = 1$, such that for each $\omega \in \Omega_0$ there exists n_ω with

$$-W_{\theta^{n+1}}(\omega) \leq 2h(\theta^{n+1}), \quad n \geq n_\omega.$$

Note that for $n \geq 2/\log\theta^{-1}$

$$\log\theta^{-(n+1)} = (n+1)\log\theta^{-1} \leq 2n\log\theta^{-1} \leq n^2(\log\theta^{-1})^2 = (\log\theta^{-n})^2.$$

Hence

$$\log\log\theta^{-(n+1)} \leq \log(\log\theta^{-n})^2 \leq 2\log\log\theta^{-n}.$$

and so we find that

$$h(\theta^{n+1}) \leq \sqrt{2}\theta^{(n+1)/2}\sqrt{2\log\log\theta^{-n}} \leq 2\sqrt{\theta}h(\theta^n).$$

Combining this with the preceding inequality yields for $\omega \in \Omega_0 \cap \Omega_\theta$ that

$$W_{\theta^n}(\omega) \geq (1-\sqrt{\theta})h(\theta^n) - 2h(\theta^{n+1}) \geq h(\theta^n)(1-5\sqrt{\theta}),$$

for inifinitely many n. Hence

$$\limsup_{t\downarrow 0}\frac{W_t(\omega)}{h(t)} \geq 1-5\sqrt{\theta}, \quad \omega \in \Omega_0 \cap \Omega_\theta.$$

Finally put $\Omega^* = \Omega_0 \cap \cap_{k\geq 1}\Omega_{1/k}$. Clearly $\mathsf{P}\{\Omega^*\} = 1$ and $\limsup_{t\downarrow 0}\frac{W_t(\omega)}{h(t)} \geq \lim_{k\to\infty}(1 - 5/\sqrt{k}) = 1$. QED

As a corollary we have the following result regarding the zero set of the BM that was considered in Exercise 1.28 of Chapter 1.

Corollary 2.4.6 *The point 0 is an accumulation point of the zero set of the* BM*, i.e. for every* $\epsilon > 0$*, the* BM *visits 0 infinitely often in the time interval* $[0, \epsilon)$*.*

Proof. By the law of the iterated logarithm, there exist sequences t_n and s_n converging monotonically to 0, such that

$$\frac{W_{t_n}}{\sqrt{2t_n\log\log 1/t_n}} \to 1, \quad \frac{W_{s_n}}{\sqrt{2s_n\log\log 1/s_n}} \to -1, \quad n\to\infty.$$

The corollary follows from the continuity of Brownian motion paths. QED

2.4.4 Distribution of hitting times

Let W be a standard Brownian motion and, for $a > 0$, let τ_a be the (a.s. finite) hitting time of level a (cf. Example 1.6.9).

Theorem 2.4.7 *For $a > 0$ the Laplace transform of the hitting time τ_a is given by*

$$\mathsf{E}e^{-\lambda\tau_a} = e^{-a\sqrt{2\lambda}}, \quad \lambda \geq 0.$$

Proof. For $b \geq 0$, consider the exponential martingales $M_t = \exp(bW_t - b^2t/2)$ (see Example 2.1.4. The stopped process M^{τ_a} is again a martingale (see Corollary 2.3.14) and is bounded by $\exp(ab)$. A bounded martingale is uniformly integrable. Hence, by the optional stopping theorem

$$\mathsf{E}M_{\tau_a} = \mathsf{E}M_\infty^{\tau_a} = \mathsf{E}M_0^{\tau_a} = \mathsf{E}M_0 = 1.$$

Since $W_{\tau_a} = a$, it follows that

$$\mathsf{E}e^{ba - b^2\tau_a/2} = 1.$$

The expression for the Laplace transform now follows by substituting $b^2 = 2\lambda$. QED

We will later see that τ_a has the density

$$x \to \frac{ae^{-a^2/2x}}{\sqrt{2\pi x^3}} \mathbf{1}_{\{x\geq 0\}}.$$

This can be shown by inverting the Laplace transform of τ_a.

A formula for the inversion of Laplace transforms are given in BN§4.

We will however use an alternative method in the next chapter. At this point we only prove that although the hitting times τ_a are a.s. finite, we have $\mathsf{E}\tau_a = \infty$ for every $a > 0$. A process with this property is called *null recurrent*.

Corollary 2.4.8 *For every $a > 0$ it holds that $\mathsf{E}\tau_a = \infty$.*

Proof. Denote the distribution function of τ_a by F. By integration by parts we have for every $\lambda > 0$

$$\mathsf{E}e^{-\lambda\tau_a} = \int_0^\infty e^{-\lambda x} dF(x) = e^{-\lambda x}F(x)\Big|_0^\infty - \int_0^\infty F(x)d(e^{-\lambda x}) = -\int_0^\infty F(x)d(e^{-\lambda x}).$$

Combination with the fact that

$$-1 = \int_0^\infty d(e^{-\lambda x})$$

it follows that

$$\frac{1 - \mathsf{E}e^{-\lambda\tau_a}}{\lambda} = -\frac{1}{\lambda}\int_0^\infty (1 - F(x))d(e^{-\lambda x}) = \int_0^\infty (1 - F(x))e^{-\lambda x}dx.$$

Now suppose that $\mathsf{E}\tau_a < \infty$. Then by dominated convergence the right-hand side converges to $\mathsf{E}\tau_a$ as $\lambda \to 0$. In particular

$$\lim_{\lambda\downarrow 0} \frac{1 - \mathsf{E}e^{-\lambda\tau_a}}{\lambda}$$

is finite. However, the preceding theorem shows that this is not the case. QED

2.5 Poisson process and the PASTA property

Let N be a right continuous Poisson process with parameter μ on $(\Omega, \mathcal{F}, \mathsf{P})$. Here the space Ω are right-continuous, non-decreasing integer valued paths, such that for each path ω one has $\omega_0 = 0$ as well as $\omega_t \leq \lim_{s\uparrow t} \omega_s + 1$, for all $t > 0$ (cf. construction in Chapter 1.1). The path properties imply all paths in Ω to have at most finitely many discontinuities in each bounded time interval. The σ-algebra $\mathcal{F}$ is the associated σ-algebra that makes the projections on the t-coordinate measurable.

As we have seen in Example 2.1.5 , $\{N_t - \mu t\}_t$ is a martingale. This implies that N_t has a decomposition as the sum of a martingale and an increasing process, called Doob-Meyer decomposition.

Lemma 2.5.1

$$\frac{N_t}{t} \stackrel{\text{a.s.}}{\to} \mu, \quad t \to \infty.$$

Proof. See Exercise 2.34. QED

The structure of Poisson paths implies that each path can be viewed as to represent the 'distribution function' of a counting measure, that gives measure 1 to each point where $N(\omega)$ has a discontinuity. In fact, the measure ν_ω, with $\nu_\omega([0,t]) = N_t(\omega)$ is a Lebesgue-Stieltjes measure generated by the trajectory $N(\omega)$, and $\nu_\omega(A)$ 'counts' the number of jumps of $N(\omega)$ occurring in $A \in \mathcal{B}[0,\infty)$.

Denote these successive jumps by $S_n(\omega)$, $n = 1, \ldots$. Let $f : [0,t] \to \mathbf{R}$ be any bounded or non-negative measurable function. Then for each $\omega \in \Omega$, $t \geq 0$, we define

$$\int_0^t f(s) dN_s(\omega) = \sum_{n=0}^{N_t(\omega)} f(S_n(\omega)) = \sum_{n=0}^{\infty} f(S_n(\omega)) \mathbf{1}_{\{S_n(\omega) \leq t, N_t(\omega) < \infty\}}. \tag{2.5.1}$$

This can be derived using the 'standard machinery'.

The PASTA property Let us now consider some $(E, \mathcal{E})$-valued stochastic process X on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, \mathsf{P})$, where $(E, \mathcal{E})$ is some measure space. Let $B \in \mathcal{E}$.

The aim is to compare the fraction of time that X-process spends in set B, with the fraction of time points generated by the Poisson process that the X-process is in B. We need to introduce some notation: $U_t = \mathbf{1}_{\{X_t \in B\}}$, and, as usual, λ stands for the Lebesgue measure on $(\mathbf{R}_+, \mathcal{B}(\mathbf{R}_+))$.

Assumption A U has ladcag paths, i.o.w. all paths of U are left continuous and have right limits.

We further define

$$\begin{aligned}
\overline{U}_t &= \frac{1}{t}\int_0^t U_s d\lambda(s) \\
A_t &= \int_0^t U_s dN_s \\
\bar{A}_t &= \frac{A_t}{N_t} \mathbf{1}_{\{N_t > 0\}} \\
\mathcal{F}_t &= \sigma(U_s, s \leq t, N_s, s \leq t).
\end{aligned}$$

Here $\overline{U}_t$ stands for the fraction of time during $(0,t]$ that X spends in set B; A_t is the amount of Poisson time points before t at which X is in set B, and $\overline{A}_t$ is the fraction of Poisson time points upto time t at which X is in B.

Assumption B Lack of anticipation property $\sigma(N_{t+u}-N_t, u\geq 0)$ and $\mathcal{F}_t$ are independent for all $t\geq 0$.

Theorem 2.5.2 *Under assumptions A and B, there exists a finite random variable $\overline{U}_\infty$ such that $\overline{U}_t \stackrel{\text{a.s.}}{\to} \overline{U}_\infty$, iff there exists a finite random variable $\bar{A}_\infty$ such that $\bar{A}_t \stackrel{\text{a.s.}}{\to} \bar{A}_\infty$ and then $\bar{A}_\infty \stackrel{\text{a.s.}}{=} \overline{U}_\infty$.*

The proof requires a number of steps.

Lemma 2.5.3 *Suppose that assumptions A and B hold.*
Then $\mathsf{E}A_t = \lambda t \mathsf{E}\overline{U}_t = \lambda \mathsf{E}\int_0^t U_s d\lambda(s)$.

Proof. In view of (2.5.1) and the sample properties of N, this implies that we can approximate A_t by

$$A_{n,t} = \sum_{k=0}^{n-1} U_{\frac{kt}{n}}[N_{\frac{(k+1)t}{n}} - N_{\frac{kt}{n}}].$$

In other words, $A_{n,t} \stackrel{\text{a.s.}}{\to} A_t$. Now, evidently $0\leq A_{n,t}\leq N_t$. Since $\mathsf{E}|N_t| = \mathsf{E}N_t < \infty$, we can apply the dominated convergence theorem and Assumption B to obtain that

$$\mathsf{E}A_t = \lim_{n\to\infty}\mathsf{E}A_{n,t} = \lim_{n\to\infty}\frac{\mu t}{n}\sum_{k=0}^{n-1}\mathsf{E}U_{\frac{tk}{n}}. \tag{2.5.2}$$

Similarly, one can derive that

$$\frac{\mu t}{n}\sum_{k=0}^{n-1}\mathsf{E}U_{\frac{tk}{n}} \stackrel{\text{a.s.}}{\to} \mu\int_0^t U_s d\lambda(s), \quad n\to\infty.$$

Using that

$$0 \leq \frac{\mu t}{n}\sum_{k=0}^{n-1}\mathsf{E}U_{\frac{tk}{n}} \leq \mu t,$$

we can apply the dominated convergence theorem to obtain, that

$$\mathsf{E}\lim_{n\to\infty}\frac{\mu t}{n}\sum_{k=0}^{n-1}U_{\frac{tk}{n}} = \mu\mathsf{E}\int_0^t U_s d\lambda(s). \tag{2.5.3}$$

Combining (2.5.2) and (2.5.3) yields

$$\mathsf{E}A_t = \mu\mathsf{E}\int_0^t U_s d\lambda(s).$$

QED

Corollary 2.5.4 *Suppose that assumptions A and B hold.* $\mathsf{E}(A_t - A_s \mid \mathcal{F}_s) = \mu \mathsf{E}(\int_s^t U_v d\lambda(v) \mid \mathcal{F}_s)$ *a.s.*

Proof. The lemma implies that $\mathsf{E}(A_t - A_s) = \mu \mathsf{E} \int_s^t U_v d\lambda(v)$. Define

$$A_{n,s,t} = \sum_{ns/t \le k \le n-1} U_{\frac{kt}{n}} [N_{\frac{(k+1)t}{n}} - N_{\frac{kt}{n}}].$$

Then, analogously to the above proof, $A_{n,s,t} \stackrel{\text{a.s.}}{\to} A_t - A_s = \int_s^t U_v dN_v$.

We use conditional dominated convergence (BN Theorem 7.2 (vii)). This implies that $\mathsf{E}(A_{n,s,t} \mid \mathcal{F}_s) \to \mathsf{E}(A_t - A_s \mid \mathcal{F}_s)$ a.s. On the other hand

$$\mathsf{E}(A_{n,s,t} \mid \mathcal{F}_s) = \mathsf{E}(\tfrac{\mu t}{n} \sum_{ns/t \le k \le n-1} U_{\frac{kt}{n}} \mid \mathcal{F}_s).$$

By another application of conditional dominated convergence, using boundedness of the function involved, the right-hand side converges a.s. to $\mathsf{E}(\mu \int_s^t U_v dv \mid \mathcal{F}_s)$. QED

Next define $R_t = A_t - \mu t \overline{U}_t$. By virtue of Corollary 2.5.4 $\{R_t\}_t$ is an $(\mathcal{F}_t)_t$-adapted martingale.

Lemma 2.5.5 *Suppose that assumptions A and B hold. Then* $R_t/t \stackrel{\text{a.s.}}{\to} 0$.

Proof. Note that $\{R_{nh}\}_n$ is an $(\mathcal{F}_{nh})_n$-adapted discrete time martingale for any $h > 0$. By virtue of Theorem 2.2.22

$$\frac{R_{nh}}{n} \to 0,$$

on the set

$$A = \{\sum_{k=1}^{\infty} \frac{1}{k^2} \mathsf{E}((R_{(k+1)h} - R_{kh})^2 \mid \mathcal{F}_{kh}) < \infty\}.$$

Note that

$$|R_t - R_s| \le N_t - N_s + \mu(t-s). \tag{2.5.4}$$

It follows that

$$\mathsf{E}(R_t - R_s)^2 \le \mathsf{E}(N_t - N_s)^2 + 3\mu^2(t-s)^2 = 4\mu^2(t-s)^2 + \mu(t-s). \tag{2.5.5}$$

Consider the random variables $Y_n = \sum_{k=1}^n \frac{1}{k^2} \mathsf{E}((R_{(k+1)h} - R_{kh})^2 \mid \mathcal{F}_{kh})$, $n = 1, \ldots$. By (2.5.5),

$$\mathsf{E}|Y_n| = \mathsf{E}Y_n \le \sum_{k=1}^{\infty} (4\mu^2 h^2 + \mu h) \frac{1}{k^2} < \infty,$$

for all n, hence $\{Y_n\}_n$ is bounded in L_1. Y_n is an increasing sequence that converges to a limit Y_∞, that is possibly not finite everywhere. By monotone convergence and L_1-boundednes $\mathsf{E}Y_n \to \mathsf{E}Y_\infty < \infty$. As a consequence Y_∞ must be a.s. finite. In other words, $\mathsf{P}\{A\} = 1$. That is $R(nh)/n \stackrel{\text{a.s.}}{\to} 0$.

Let Ω^* be the intersection of A and the set where $N_t/t \to \mu$. By Lemma 2.5.1 $\mathsf{P}\{\Omega^*\} = 1$. The lemma is proved if we show that $R_t/t \to 0$ on the set Ω^*. Let $\omega \in \Omega^*$. Fix t and let n_t be such that $t \in [n_t h, (n_t + 1)h)$. By virtue of(2.5.4)

$$|R_t(\omega) - R_{n_t h}(\omega)| \le N_t(\omega) - N_{n_t h}(\omega) + \mu h.$$

By another application of Lemma 2.5.1

$$\frac{R_t(\omega)}{t} \le \frac{n_t h}{t} \cdot \frac{R_{n_t h}(\omega) + N(t,\omega) - N_{n_t h})(\omega) + \mu h}{n_t h} \to 0, \to \infty.$$

QED

Now we can finish the proof of the theorem. It follows from the relation

$$\frac{R_t}{t} = \frac{A_t}{N_t}\frac{N_t}{t}\mathbf{1}_{\{N_t>0\}} - \mu\overline{U}_t,$$

whilst noting that $(N_t/t)\mathbf{1}_{\{N_t>0\}} \to \mu$ on Ω^*.

2.6 Exercises

Discrete-time martingales

Exercise 2.1 Prove the assertion in Example 2.1.2.

Exercise 2.2 Prove the assertion in Example 2.1.3.

Exercise 2.3 Show that the processes defined in Example 2.1.4 are indeed martingales.

Exercise 2.4 (Kolmogorov 0-1 Law) Let $X_1, X_2, \ldots$ be i.i.d. random variables and consider the tail σ-algebra defined by

$$\mathcal{T} = \bigcap_{n=1}^{\infty} \sigma(X_n, X_{n-1}, \ldots).$$

a) Show that for every n, $\mathcal{T}$ is independent of the σ-algebra $\sigma(X_1, \ldots, X_n)$ and conclude that for every $A \in \mathcal{T}$

$$\mathsf{P}\{A\} = \mathsf{E}(\mathbf{1}_{\{A\}} \,|\, X_1, \ldots, X_n), \quad \text{a.s.}$$

b) Give a "martingale proof" of Kolmogorov's $0-1$ law: for every $A \in \mathcal{T}$, $\mathsf{P}\{A\} = 0$ or $\mathsf{P}\{A\} = 1$.

c) Give an example of an event $A \in \mathcal{T}$.

Exercise 2.5 (Law of large Numbers) In this exercise we present a "martingale proof" of the law of large numbers. Let $X_1, X_2, \ldots$ be random variables with $\mathsf{E}|X_1| < \infty$. Define $S_n = \sum_{i=1}^{n} X_i$ and $\mathcal{F}_n = \sigma(S_n, S_{n+1}, \ldots)$.

a) Note that for $i = 1, \ldots, n$, the distribution of the pair (X_i, S_n) is independent of i. From this fact, deduce that $\mathsf{E}(X_n \,|\, \mathcal{F}_n) = S_n/n$, and that consequently

$$\mathsf{E}(\tfrac{1}{n} S_n \,|\, \mathcal{F}_n) = \tfrac{1}{n+1} S_{n+1}, \quad \text{a.s.}$$

ii) Show that S_n/n converges almost surely to a finite limit.

iv) Derive from Kolmogorov's $0-1$ law that the limit must be a constant and determine its value.

Exercise 2.6 Consider the proof of Theorem 2.2.18. Prove that for the stopping time τ and the event $A \in \mathcal{F}_\tau$ it holds that $A \cap \{\tau < \infty\} \in \mathcal{G}$.

Exercise 2.7 Let X, Y be two integrable random variables defined on the same space Ω. Let $\mathcal{F}$ be a σ-algebra on Ω.

a) Suppose that X, Y are both $\mathcal{F}$-measurable. Show that $X \geq Y$ a.s. if and only if $\mathsf{E}\mathbf{1}_{\{A\}} X \geq \mathsf{E}\mathbf{1}_{\{A\}} Y$ for all $A \in \mathcal{F}$.

b) Suppose that Y is $\mathcal{F}$-measurable. Show that $\mathsf{E}(X \,|\, \mathcal{F}) \leq Y$ a.s. together with $\mathsf{E}X = \mathsf{E}Y$ implies $\mathsf{E}(X \,|\, \mathcal{F}) = Y$ a.s.

Exercise 2.8 Let M be a martingale such that $\{M_{n+1} - M_n\}_{n\geq 1}$ is a bounded process. Let Y be a bounded predictable process. Let $X = Y \cdot M$. Show that $\mathsf{E}X_\tau = 0$ for τ a finite stopping time with $\mathsf{E}\tau < \infty$.

Exercise 2.9 Let $X_1, X_2, \ldots$ be an i.i.d. sequence of Bernouilli random variables with probability of success equal to p. Put $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$, $n \geq 1$. Let M be a martingale adapted to the generated filtration. Show that the *Martingale Representation Property* holds: there exists a constant m and a predictable process Y such that $M_n = m + (Y \cdot S)_n$, $n \geq 1$, where $S_n = \sum_{k=1}^n (X_k - p)$.

Exercise 2.10 Let $X_1, \ldots$ be a sequence of independent random variables with $\sigma_n^2 = \mathsf{E}X_n^2 < \infty$ and $\mathsf{E}X_n = 0$ for all $n \geq 1$. Consider the filtration generated by X and define the martingale M by $M_n = \sum_{i=1}^n X_i$. Determine $\langle M \rangle$.

Exercise 2.11 Let M be a martingale with $\mathsf{E}M_n^2 < \infty$ for every n. Let C be a bounded predictable process and define $X = C \cdot M$. Show that $\mathsf{E}X_n^2 < \infty$ for every n and that $\langle X \rangle = C^2 \cdot \langle M \rangle$.

Exercise 2.12 Let M be a martingale with $\mathsf{E}M_n^2 < \infty$ for every n and let τ be a stopping time. We know that the stopped process is a martingale as well. Show that $\mathsf{E}(M_n^\tau)^2 < \infty$ for all n and that $\langle M^\tau \rangle_n = \langle M \rangle_{n\wedge\tau}$.

Exercise 2.13 Let $(C_n)_n$ be a predictable sequence of random variables with $\mathsf{E}C_n^2 < \infty$ for all n. Let $(\epsilon_n)_n$ be a sequence with $\mathsf{E}\epsilon_n = 0$, $\mathsf{E}\epsilon_n^2 = 1$ and ϵ_n independent of $\mathcal{F}_{n-1}$ for all n. Let $M_n = \sum_{i\leq n} C_i\epsilon_i$, $n \geq 0$. Compute the conditional variance process A of M. Take $p > 1/2$ and consider $N_n = \sum_{i\leq n} C_i\epsilon_i/(1 + A_i)^p$. Show that there exists a random variable N_∞ such that $N_n \to N_\infty$ a.s. Show (use Kronecker's lemma) that $M_n/(1 + A_n)^p$ has an a.s. finite limit.

Exercise 2.14 **i)** Show that the following generalisation of the optional stopping Theorem 2.2.18 holds. Let M be a uniformly integrable supermartingale. Then the family of random variables $\{M_\tau \,|\, \tau \text{ is a finite stopping time}\}$ is UI and $\mathsf{E}(M_\tau \,|\, \mathcal{F}_\sigma) \leq M_\sigma$, a.s. for stopping times $\sigma \leq \tau$. Hint: use Doob decomposition.

ii) Give an example of a non-negative martingale for which $\{M_\tau \,|\, \tau \text{ stopping time}\}$ is not UI.

Exercise 2.14* Show for a non-negative supermartingale M that for all $\lambda > 0$

$$\lambda\mathsf{P}\{\sup_n M_n \geq \lambda\} \leq \mathsf{E}(M_0).$$

Exercise 2.15 Consider the unit interval $I = [0,1]$ equipped with the Borel-σ-algebra $\mathcal{B}([0,1])$ and the Lebesgue measure. Let f be an integrable function on I. Let for $n = 1, 2, \ldots$

$$f_n(x) = 2^n \int_{(k-1)2^{-n}}^{k2^{-n}} f(y)dy, \quad (k-1)2^{-n} \leq x < k2^{-n},$$

and define $f_n(1) = 1$ (the value $f_n(1)$ is not important). Finally, we define $\mathcal{F}_n$ as the σ-algebra generated by intervals of the form $[(k-1)2^{-n}, k2^{-n})$, $1 \leq k \leq 2^n$.

i) Argue that $\mathcal{F}_n$ is an increasing sequence of σ-algebras.

ii) Show that $(f_n)_n$ is a martingale.

iii) Use Lévy's Upward Theorem to prove that $f_n \to f$, a.s. and in L_1, as $n \to \infty$.

Exercise 2.16 (Martingale formulation of Bellman's optimality principle) Suppose your winning per unit stake on game n are ϵ_n, where the ϵ_n are i.i.d. r.v.s with

$$\mathsf{P}\{\epsilon_n = 1\} = p = 1 - \mathsf{P}\{\epsilon_n = -1\},$$

with $p > 1/2$. Your bet α_n on game n must lie between 0 and Z_{n-1}, your capital at time $n-1$. Your object is to maximise your 'interest rate' $\mathsf{E}\log(Z_N/Z_0)$, where $N =$*length of the game* is finite and Z_0 is a given constant. Let $\mathcal{F}_n = \sigma(\epsilon_1, \ldots, \epsilon_n)$ be your 'history' upto time n. Let $\{\alpha_n\}_n$ be an admissible strategy, i.o.w. a predictable sequence. Show that $\log(Z_n) - n\alpha$ is a supermartingale with α the entropy given by

$$\alpha = p\log p + (1-p)\log(1-p) + \log 2.$$

Hence $\mathsf{E}\log(Z_n/Z_0) \leq n\alpha$. Show also that for some strategy $\log(Z_n) - n\alpha$ is a martingale. What is the best strategy?

Exercise 2.17 Consider a monkey typing one of the numbers $0, 1, \ldots, 9$ at random at each of times $1, 2, \ldots$ U_i denotes the i-th number that the monkey types. The sequence of numbers $U_1, U_2, \ldots$, form an i.i.d. sequence uniformly drawn from the 10 possible numbers.

We would like to know how long it takes till the first time T that the monkey types the sequence **1231231231**. More formally,

$$T = \min\{n \mid n \geq 10, U_{n-9}U_{n-8}\cdots U_n = \mathbf{1231231231}\}.$$

First we need to check that T is an a.s. finite, integrable r.v. There are many ways to do this.

a) Show that T is an a.s. finite, integrable r.v. A possibility for showing this, is to first show that the number of consecutive 10–number words typed till the first occurrence of **1231231231** is a.s. finite, with finite expectation.

In order to actually compute $\mathsf{E}T$, we will associate a gambling problem with it.

Just before each time $t = 1, 2, 3, \ldots$, a new gambler arrives into the scene, carrying €1 in his pocket. He bets €1 that the next number (i.e. the t-th number) will be **1**. If he loses, he leaves; if he wins his receives 10 times his bet, and so he will have a total capital of €10. He next bets all of his capital on the event that the $(t+1)$-th number will be **2**. If he loses, he leaves; if he wins, he will have a capital of €10^2. This is repeated throughout the sequence **1231231231**. So, if the gambler wins the second time, his third bet is on the number **3**, and so on, till the moment that either he loses, or the monkey has typed the desired sequence. Note that any gambler entering the game *after* the monkey typed the desired sequence, cannot play anymore, he merely keeps his initial capital intact.

b) Define a martingale $\{X_n^1 - 1\}_{n=0,\ldots}$, such that X_n^1 is the total capital of the first gambler after his n-th game and hence $X_n^1 - 1$ his total gain. Similarly associate with the k-th gambler (who enters the game at time k) a martingale $X_n^k - 1$, where X_n^k is his capital after the n-th number that the monkey typed. Write $M_n = \sum_{k=1}^n X_n^k$. Then argue that $(M_n - n)_n$ is a bounded martingale associated with the total gain of all gamblers that entered the game at time n latest.

c) Argue now that $\mathsf{E}T = 10^{10} + 10^7 + 10^4 + 10$.

Exercise 2.18 Let X_i, $0 = 1, 2 \ldots$, be independent, integer valued random variables, with $X_1, \ldots$ identically distributed. Assume that $\mathsf{E}|X_1|$, $\mathsf{E}X_1^2 < \infty$. Let $a < b$, $a, b \in \mathbf{Z}$, and assume that $X_i \not\equiv 0$. Consider the stochastic process $S = (S_n)_{n \in \mathrm{N}}$, with $S_n = \sum_{i=0}^n X_i$ the $(n+1)$-th partial sum.

We desire to show the intuitively clear assertion that the process leaves (a, b) in finite expected time, given that it starts in (a, b). Define $\tau_{a,b} = \min\{n \mid S_n \notin (a, b)\}$ and

$$f(x) = \mathsf{P}\{S_i \in (a, b), i = 0, 1, 2, \ldots \mid S_0 = x\}.$$

Note that $f(x) = 0$ whenever $x \notin (a, b)$! Let now $S_0 \equiv x_0 \in (a, b)$.

a) Show that $(f(S_{n \wedge \tau_{a,b}}))_n$ is a martingale.

b) Show that this implies that $\tau_{a,b}$ is a.s. finite. Hint: consider the maximum of f on (a, b) and suppose that it is strictly positive. Derive a contradiction.

Fix $x_0 \in (a, b)$, let $S_0 \equiv x_0$, and assume that X_1 is bounded.

c) Show that $\tau_{a,b}$ is an a.s. finite and integrable r.v. Hint: you may consider the processes $(S_n - n\mathsf{E}X_1)_{n \in \mathrm{N}}$, and, if $\mathsf{E}X_1 = 0$, $(S_n^2 - n\mathsf{E}X_1^2)_{n \in \mathrm{N}}$.

d) Show that $\tau_{a,b}$ is a.s. finite and integrable also if X_1 is not bounded.

e) Derive an expression for $\mathsf{E}\tau_{a,b}$ (in terms of x_0, a and b) in the special case that $\mathsf{P}\{X_1 = 1\} = \mathsf{P}\{X_1 = -1\} = 1/2$.

f) Now assume that $\mathsf{E}X_1 \leq 0$. Let $\tau_a = \min\{n \geq 0 \mid S_n \leq a\}$. Show that τ_a is a.s. finite. Hint: consider $\tau_{a,n}$ and let n tend to ∞.

Exercise 2.18' Another approach of the first part of Exercise 2.18. Let X_i, $i = 0, \ldots$, all be defined on the same underlying probability space $(\Omega, \mathcal{F}, \mathsf{P})$. Let

$$f = \mathbf{1}_{\{S_i \in (a,b), i=0,\ldots\}}.$$

a') The stochastic process $M_n = \mathsf{E}(f \mid S_0, \ldots, S_n)$ is a martingale that converges a.s. and in L^1 to $M_\infty = \mathsf{E}\{f \mid S_0, \ldots\}$. Argue that

$$\mathsf{E}\{f \mid S_0 = x_0, \ldots, S_n = x_n\} = \mathsf{E}\{f \mid S_0 = x_n\},$$

for all $x_0, \ldots, x_{n-1} \in (a, b)$.

Use this to show for all $x \in (a, b)$ that

$$a_x = \sum_y \mathsf{P}\{X_1 = y\} a_{x+y},$$

where $a_x = \mathsf{E}\{f \mid S_0 = x\}$ (note: a_x is a real number!).

b') Show that this implies that $a_x = 0$ for all $x \in (a, b)$. Hint: consider the point $x^* = \{x \in (a, b) \mid a_x = \max_{y \in (a,b)} a_y\}$. Let now $S_0 = x_0 \in (a, b)$ be given. Conclude from the previous that $\tau_{a,b}$ is a.s. finite.

Exercise 2.19 (Galton-Watson process) This is a simple model for population growth, growth of the number of cells, etc.

A population of cells evolves as follows. In every time step, every cell splits into 2 cells with probability p or it dies with probability $1-p$, independently of the other cells and of the population history. Let N_t denote the number of cells at time t, $t = 0, 1, 2, \ldots$. Initially, there is only 1 cell, i.e. $N_0 = 1$.

We can describe this model formally by defining Z_t^n, $n = 1, \ldots, N_t$, $t = 0, 1, \ldots$, to be i.i.d. random variables with

$$\mathsf{P}\{Z_t^n = 2\} = p = 1 - \mathsf{P}\{Z_t^n = 0\},$$

and then $N_{t+1} = \sum_{n=1}^{N_t} Z_t^n$. Let $\{\mathcal{F}_t\}_{t=0,1,\ldots}$ be the natural filtration generated by $\{N_t\}_{t=0,1,\ldots}$.

i) Argue or prove that

$$\begin{aligned}\mathsf{P}\{\mathbf{1}_{\{N_{t+1}=2y\}} \,|\, \mathcal{F}_t\} &= \mathsf{E}\{\mathbf{1}_{\{N_{t+1}=2y\}} \,|\, \mathcal{F}_t\} \\ &= \mathsf{E}\{\mathbf{1}_{\{N_{t+1}=2y\}} \,|\, N_t\} = \mathsf{P}\{N_{t+1} = 2y \,|\, N_t\} = \binom{N_t}{y} p^y (1-p)^{N_t - y}.\end{aligned}$$

Hence, conditional on $\mathcal{F}_t$, $N_{t+1}/2$ has a binomial distribution with parameters N_t and p.

ii) Let $\mu = \mathsf{E}\{N_1\}$. Show that N_t/μ^t is a martingale with respect to $\{\mathcal{F}_t\}$, bounded in L^1.

iii) Assume that $\mu < 1$. Show that $\mathsf{E}N_t = \mu^t$ and that the population dies out a.s. in the long run.

iv) Assume again that $\mu < 1$. Show that $M_t = \alpha^{N_t}\mathbf{1}_{\{N_t>0\}}$ is a contracting supermartingale for some $\alpha > 1$, i.e. there exist $\alpha > 1$ and $0 < \beta < 1$ such that

$$\mathsf{E}(M_{t+1} \,|\, \mathcal{F}_t) \le \beta M_t, \quad t = 1, 2, \ldots.$$

v) Show that this implies that $\mathsf{E}(T) < \infty$ with $T = \min\{t \ge 1 \,|\, N_t = 0\}$ the extinction time.

Exercise 2.20 (Continuation of Exercise 2.19) From now on, assume the critical case $\mu = 1$, and so N_t is itself a martingale. Define $\tau_{0,N} = \min\{t \,|\, N_t = 0 \text{ or } N_t \ge N\}$. Further define

$$M_t = N_t \cdot \mathbf{1}_{\{N_1,\ldots,N_t \in \{1,\ldots,N-1\}\}}.$$

i) Argue that M_t is a supermartingale and that there exists a constant $\alpha < 1$ (depending on N) such that

$$\mathsf{E}M_{t+N} \le \alpha \mathsf{E}N_t.$$

Show that this implies that $\mathsf{P}\{\tau_{0,N} = \infty\} = 0$.

ii) Show that $\mathsf{P}\{N_{\tau_{0,N}} \ge N\} \le 1/N$. Show that this implies the population to die out with probability 1.

iii) Is $\{N_t\}_t$ UI in the case of $\mu < 1$? And if $\mu = 1$?

Exercise 2.21 Let X_i, $i = 0, \ldots$ be independent $\mathbf{Z}$-valued random variables, with $\mathsf{E}|X_i| < \infty$. We assume that $X_1, X_2, \ldots$ are identically distributed with values in $\{-1, 0, 1\}$, that all have positive probability. Then $S_n = \sum_{i=0}^n X_i$ is a discrete time Markov chain taking values in $\mathbf{Z}$. Suppose that X_0 has distribution $\nu = \delta_x$, with $x \in (a, b) \subset \mathbf{Z}_+$. Define $\tau_y = \min\{n > 0 : S_n = y\}$ and let $\tau = \tau_a \wedge \tau_b$. We want to compute $\mathsf{P}\{\tau_b < \tau_a\}$ and $\mathsf{E}\{\tau\}$ (recall that $S_0 \equiv x$!). Let first $\mathsf{E}X_1 \neq 0$.

i) Show that τ is a stopping time w.r.t. a suitable filtration. Show that τ is finite a.s. Hint: use the law of large numbers.

ii) We want to define a function $f : \mathbf{Z} \to \mathbf{R}$, such that $\{f(S_n)\}_n$ is a discrete-time martingale. It turns out that we can take $f(z) = e^{\alpha z}$, $z \in \mathbf{R}$, for suitably chosen α.

Show that there exists $\alpha \neq 0$, such that $e^{\alpha S_n}$ is a martingale. Use this martingale to show that

$$\mathsf{P}\{\tau_b < \tau_a\} = \frac{e^{\alpha x} - e^{\alpha a}}{e^{\alpha b} - e^{\alpha a}}.$$

If $\mathsf{P}\{X_i = 1\} = p = 1 - \mathsf{P}\{X_i = -1\}$ (that is: X_i takes only values ± 1), then

$$e^{\alpha} = \frac{1-p}{p} \qquad \text{or } \alpha = \log(1-p) - \log p.$$

Show this.

iii) Show that $S_n - n\mathsf{E}X_1$ is martingale. Show that

$$\mathsf{E}\{\tau\} = \frac{(e^{\alpha x} - e^{\alpha b})(x-a) + (e^{\alpha x} - e^{\alpha a})(b-x)}{(e^{\alpha b} - e^{\alpha a})\mathsf{E}X_1}.$$

iv) Let now $\mathsf{E}X_1 = 0$. Show that $\tau < \infty$ a.s. (hint: use the Central Limit Theorem). Show for $x \in (a, b)$ that

$$\mathsf{P}\{\tau_a < \tau_b\} = \frac{b-x}{b-a}$$

and

$$\mathsf{E}\{\tau\} = \frac{(x-a)(b-x)}{\mathsf{E}X_1^2},$$

by constructing suitable martingales.

Continuous-time martingales

Exercise 2.22 Prove Corollary 2.3.3. Hint: we know from Theorem 2.3.2 that the left limits exist for all t on an $\mathcal{F}_\infty$-measurable subset $\Omega^*$of probability 1, along rational sequences. You now have to consider arbitrary sequences.

Exercise 2.23 Show that the process constructed in the proof of Theorem 2.3.5 is cadlag and a supermartingale, if M is a supermartingale.

Exercise 2.24 Prove the 'only if' part of Theorem 2.3.5.

Exercise 2.25 Prove Theorem 2.3.9 from LN. You may use Theorem 2.3.7.

Exercise 2.26 Show that for every $a \neq 0$, the exponential martingale of Example 2.1.4 converges to 0 a.s, as $t \to \infty$. (Hint: use for instance the recurrence of Brownian motion) Conclude that these martingales are not uniformly integrable.

Exercise 2.27 Give an example of two stopping times $\sigma \leq \tau$ and a martingale M that is bounded in L^1 but not uniformly integrable, for which the equality $\mathsf{E}(M_\tau \,|\, \mathcal{F}_\sigma) = M_\sigma$ a.s. fails. (Hint: see Exercise 2.26).

Exercise 2.28 Let M be a positive, continuous martingale that converges a.s. to zero as t tends to infinity.

a) Prove that for every $x > 0$

$$\mathsf{P}\{\sup_{t\geq 0} M_t > x \,|\, \mathcal{F}_0\} = 1 \wedge \tfrac{M_0}{x} \quad \text{a.s.}$$

(Hint: stop the martingale when it gets to above the level x).

b) Let W be a standard BM. Using the exponential martingales of Example 2.1.4, show that for every $a > 0$ the random variable

$$\sup_{t\geq 0}(W_t - \tfrac{1}{2}at)$$

has an exponential distribution with parameter a.

Exercise 2.29 Let W be a BM and for $a \in \mathbf{R}$ let τ_a be the first time that W hits a. Suppose that $a > 0 > b$. By considering the stopped martingale $W^{\tau_a \wedge \tau_b}$, show that

$$\mathsf{P}\{\tau_a < \tau_b) = \frac{-b}{a-b}.$$

Exercise 2.30 Consider the setup of the preceding exercise. By stopping the martingale $W_t^2 - t$ at an appropriate stopping time, show that $\mathsf{E}(\tau_a \wedge \tau_b) = -ab$. Deduce that $\mathsf{E}\tau_a = \infty$.

Exercise 2.31 Let W be a BM and for $a > 0$, let $\tilde{\tau}_a$ be the first time that $|W|$ hit the level a.

a) Show that for every $b > 0$, the process $M_t = \cosh(b|W_t|)\exp\{b^2 t/2)$ is a martingale.

b) Find the Laplace transform of the stopping time $\tilde{\tau}_a$.

c) Calculate $\mathsf{E}\tilde{\tau}_a$.

Exercise 2.32 (Emperical distributions) Let $X_1, \ldots, X_n$ be i.i.d. random variables, each with the uniform distribution on $[0,1]$. For $0 \leq t < 1$ define

$$G_n(t) = \frac{1}{n}\#\{k \leq n \,|\, X_k \leq t\} = \frac{1}{n}\sum_{k=1}^{n} \mathbf{1}_{\{X_k \leq t\}}.$$

In words, $G_n(t)$ is the fraction of X_k that have value at most t. Denote

$$\mathcal{F}_n(t) = \sigma(\mathbf{1}_{\{X_1 \leq s\}}, \ldots, \mathbf{1}_{\{X_n \leq s\}}, s \leq t\}$$

and $\mathcal{G}_n(t) = \sigma(G_n(s), s \leq t)$.

i) Prove that for $0 \le t \le u < 1$ we have

$$\mathsf{E}\{G_n(u)\,|\,\mathcal{F}_n(t)\} = G_n(t) + (1 - G_n(t))\frac{u-t}{1-t}.$$

Show that this implies that

$$\mathsf{E}\{G_n(u)\,|\,\mathcal{G}_n(t)\} = G_n(t) + (1 - G_n(t))\frac{u-t}{1-t}.$$

ii) Show that the stochastic process $M_n = (M_n(t))_{t\in[0,1)}$ defined by

$$M_n(t) = \frac{G_n(t) - t}{1-t}$$

is a continuous-time martingale w.r.t $\mathcal{G}_n(t)$, $t \in [0, 1)$.

iii) Is M_n UI? Hint: compute $\lim_{t\uparrow 1} M_n(t)$.

iv) We extend the process M_n to $[0, \infty)$, by putting $M_n(t) = 1$ and $\mathcal{G}_n(t) = \sigma(X_1, \ldots, X_n)$, for $t \ge 1$. Show that $(M_n(t))_{t\in[0,\infty)}$ is an L^1-bounded submartingale relative to $(\mathcal{G}_n(t))_{t\in[0,\infty)}$, that converges in L^1, but is not UI.

Exercise 2.33 Let $(W_t)_t$ be a standard Brownian motion, and define

$$X_t = W_t + ct,$$

for some constant c. The process X_t is called *Brownian motion with drift*. Fix some $\lambda > 0$.

i) Show that

$$M_t := e^{\theta X_t - \lambda t}$$

is a martingale (with respect to the natural filtration) if and only if $\theta = \sqrt{c^2 + 2\lambda} - c$ or $\theta = -\sqrt{c^2 + 2\lambda} - c$.

Next, let $H_x = \inf\{t > 0\,|\,X_t = x\}$.

ii) Argue for $x \neq 0$ that H_x is a stopping time.

iii) Show that

$$\mathsf{E}(e^{-\lambda H_x}) = \begin{cases} e^{-x(\sqrt{c^2+2\lambda}-c)}, & x > 0 \\ e^{-x(-\sqrt{c^2+2\lambda}-c)}, & x < 0. \end{cases}$$

iv) Use the result from (iii) to prove that for $x > 0$

$$\mathsf{P}\{H_x < \infty\} = \begin{cases} 1, & c \ge 0 \\ e^{-2|c|x}, & c < 0. \end{cases}$$

v) Explain why this result is reasonable.

Exercise 2.34 Let $N = \{N(t)\}_{t\ge 0}$ be a Poisson process (see Definition in Ch1.1). Show that $\{N(t) - \lambda t\}_{t\ge 0}$ is a martingale. Then prove Lemma 2.5.1. Hint: use the martingale LLN given in Section 2.2.6.

Chapter 3

Markov Processes

3.1 Basic definitions: a mystification?

Notational issues To motivate the conditions used lateron to define a Markov process, we will recall the definition of a discrete-time and discrete-space Markov chain.

Let E be a discrete space, and $\mathcal{E}$ the σ-algebra generated by the one-point sets: $\mathcal{E} = \sigma\{\{x\} \mid x \in E\}$. Let $X = \{X_n\}_{n=0,1,...}$ be an $(E, \mathcal{E})$-valued stochastic process defined on some underlying probability space $(\Omega, \mathcal{F}, \mathsf{P})$. In Markov chain theory, it is preferred not to fix the distribution of X_0, i.e. the initial distribution. In our notation we will therefore incorporate the dependence on the initial distribution.

> The *initial distribution of the process* is always denoted by ν in these notes. The associated probability law of X and corresponding expectation operator will be denoted by P_ν and E_ν, to make the dependence on initial distribution visible in the notation. If $X_0 = x$ a.s. then we write $\nu = \delta_x$ and use the shorthand notation P_x and E_x (instead of P_{δ_x} and E_{δ_x}). E is called the *state space*.

Assume hence that X is a stochastic process on $(\Omega, \mathcal{F}, \mathsf{P}_\nu)$. Then X is called a Markov chain with initial distribution ν, if there exists an $E \times E$ stochastic matrix P[1], such that

i) $\mathsf{P}_\nu\{X_0 \in B\} = \nu(B)$ for all $B \in \mathcal{E}$;

ii) The Markov property holds, i.e. for all $n = 0, 1, \ldots, x_0, \ldots, x_n, x_{n+1} \in E$

$$\mathsf{P}_\nu\{X_{n+1} = x_{n+1} \mid X_0 = x_0, \ldots, X_n = x_n\} = \mathsf{P}_\nu\{X_{n+1} = x_{n+1} \mid X_n = x_n\} = P(x_n\,, x_{n+1}).$$

Recall that

$$\mathsf{P}_\nu\{X_{n+1} = x_{n+1} \mid \sigma(X_n)\} = \mathsf{E}_\nu\{\mathbf{1}_{\{x_{n+1}\}}(X_{n+1}) \mid \sigma(X_n)\}$$

is a function of X_n. In the Markov case this is $P(X_n, x_{n+1})$. Then $\mathsf{P}_\nu\{X_{n+1} = x_{n+1} \mid X_n = x_n\} = P(x_n, x_{n+1})$ is simply the evaluation of that function at the point $X_n = x_n$. These conditional probabilities can be computed by

$$\mathsf{P}_\nu\{X_{n+1} = x_{n+1} \mid X_n = x_n\} = \begin{cases} \dfrac{\mathsf{P}_\nu\{X_{n+1} = x_{n+1}, X_n = x_n\}}{\mathsf{P}\{X_n = x_n\}}, & \text{if } \mathsf{P}_\nu\{X_n = x_n\} > 0 \\ \text{anything you like} & \text{if } \mathsf{P}_\nu\{X_n = x_n\} = 0 \end{cases}$$

[1]The $E \times E$ matrix P is stochastic if it is a non-negative matrix with row sums equal to 1.

We can now rephrase the Markov property as follows: for all $n \in \mathbf{Z}_+$ and $y \in E$

$$\mathsf{P}_\nu\{X_{n+1} = y \mid \mathcal{F}_n^X\} = \mathsf{P}_\nu\{X_{n+1} = y \mid \sigma(X_n)\} = P(X_n, y), \quad \text{a.s.} \tag{3.1.1}$$

It is a straightforward computation that

$$\mathsf{P}_\nu\{X_{n+m} = y \mid \mathcal{F}_n^X\} = P^m(X_n, y),$$

is the (X_n, y)-th element of the m-th power of P. Indeed, for $m = 2$

$$\begin{aligned} \mathsf{P}_\nu\{X_{n+2} = y \mid \mathcal{F}_n^X\} &= \mathsf{E}_\nu(\mathbf{1}_{\{y\}}(X_{n+2}) \mid \mathcal{F}_n^X) \\ &= \mathsf{E}_\nu(\mathsf{E}_\nu \mathbf{1}_{\{y\}}(X_{n+2}) \mid \mathcal{F}_{n+1}^X) \mid \mathcal{F}_n^X) \\ &= \mathsf{E}_\nu(P(X_{n+1}, y) \mid \mathcal{F}_n^X) \\ &= \mathsf{E}_\nu(\sum_{x \in E} \mathbf{1}_{\{x\}}(X_{n+1}) \cdot P(X_{n+1}, y) \mid \mathcal{F}_n^X) \\ &= \sum_{x \in E} P(x, y) \mathsf{E}_\nu(\mathbf{1}_{\{x\}}(X_{n+1}) \mid \mathcal{F}_n^X) \\ &= \sum_{x \in E} P(x, y) P(X_n, x) = P^2(X_n, y). \end{aligned} \tag{3.1.2}$$

In steps (3.1.2) and after, we use discreteness of the state space as well linearity of conditional expectations. In fact we have proved a more general version of the Markov property to hold. To formulate it, we need some more notation. But first we will move on to Markov chains on a general measurable space.

Discrete time Markov chains on a general state space The one point sets need not be measurable in general space. The notion of a stochastic matrix generalises to the notion of a transition kernel.

Definition 3.1.1 Let $(E, \mathcal{E})$ be a measurable space. A *transition kernel* on E is a map $P : E \times \mathcal{E} \to [0, 1]$ such that

i) for every $x \in E$, the map $B \mapsto P(x, B)$ is a probability measure on $(E, \mathcal{E})$,

ii) for every $B \in \mathcal{E}$, the map $x \mapsto P(x, B)$ is $\mathcal{E}/\mathcal{B}$-measurable.

Let X be an $(E, \mathcal{E})$ valued stochastic process defined on some underlying probability space $(\Omega, \mathcal{F}, \mathsf{P}_\nu)$. Then X is a Markov chain with initial distribution ν if (i) $\mathsf{P}_\nu\{X_0 \in B\} = \nu(B)$ for all $B \in \mathcal{E}$; (ii) if there exists a transition kernel P such that the Markov property holds:

$$\mathsf{P}_\nu\{X_{n+1} \in B \mid \mathcal{F}_n^X\} = \mathsf{P}_\nu\{X_{n+1} \in B \mid \sigma(X_n)\} = P(X_n, B), \quad B \in \mathcal{E}, n = 0, 1, 2 \ldots. \tag{3.1.3}$$

Remark If E is a discrete space, and $\mathcal{E}$ is the σ-algebra generated by the one-point sets, then for each set $\{y\}$, $y \in E$ we write $P(x, y)$ instead of $P(x, \{y\})$. Moreover, $P(x, B) = \sum_{y \in B} P(x, y)$, and so the transition kernel is completely specified by $P(x, y)$, $x, y \in E$.

As in the above, we would like to infer that

$$\mathsf{P}_\nu\{X_{n+m} \in B \mid \mathcal{F}_n^X\} = P^m(X_n, B), \quad B \in \mathcal{E}, n = 0, 1, 2 \ldots, \tag{3.1.4}$$

where P^m is defined inductively by

$$P^m(x,B) = \int P(y,B)P^{m-1}(x,dy), m = 2,3,\ldots.$$

For $m = 2$ we get (cf.(3.1.2))

$$\mathsf{P}_\nu\{X_{n+2} \in B \mid \mathcal{F}_n^X\} = \mathsf{E}_\nu(\mathsf{E}_\nu(\mathbf{1}_{\{B\}}(X_{n+2}) \mid \mathcal{F}_{n+1}^X) \mid \mathcal{F}_n^X\} = \mathsf{E}_\nu(P(X_{n+1},B) \mid \mathcal{F}_n^X).$$

Our definition of the Markov property does not allow to infer (3.1.4) for $m = 2$ directly, which we do expect to hold. However, we can prove that it does. In fact, more general relations hold. Let us introduce some notation.

Notation Integrals of the form $\int f d\nu$ are often written in operator notation as νf. A similar notation for transition kernels is as follows. If $P(x,dy)$ is a transition kernel on measurable space $(E,\mathcal{E})$ and f is a non-negative (or bounded), measurable function on E, we define the function Pf by

$$Pf(x) = \int f(y)P(x,dy).$$

Then $P(x,B) = \int \mathbf{1}_{\{B\}}P(x,dy) = P\mathbf{1}_{\{B\}}(x)$. For notational convenience, write $b\mathcal{E}$ for the space of bounded, measurable functions $f : E \to \mathbf{R}$ and $m\mathcal{E}$ for the measurable functions $f : E \to \mathbf{R}$.

Note that Pf is bounded, for $f \in b\mathcal{E}$. Since P is a transition kernel, $P\mathbf{1}_{\{B\}} \in b\mathcal{E}$. Applying the standard machinery yields that $Pf \in b\mathcal{E}$ for all $f \in b\mathcal{E}$, in particular Pf is $\mathcal{E}/\mathcal{E}$-measurable. In other words P is a linear operator mapping $b\mathcal{E}$ to $b\mathcal{E}$.

Look up in BN section 3 Measurability what we mean by the 'standard machinery'.

The Markov property (3.1.3) can now be reformulated as

$$\mathsf{E}(\mathbf{1}_{\{B\}}(X_{n+1}) \mid \mathcal{F}_n^X) = P\mathbf{1}_{\{B\}}(X_n), \quad B \in \mathcal{E}, n = 0,1,\ldots$$

Applying the standard machinery once more, yields

$$\mathsf{E}(f(X_{n+1}) \mid \mathcal{F}_n^X) = Pf(X_n), \quad f \in b\mathcal{E}, n = 0,1,\ldots.$$

This has two consequences. The first is that now

$$\mathsf{P}\{X_{n+2} \in B \mid \mathcal{F}_n^X\} = \mathsf{E}(\mathsf{E}(\mathbf{1}_{\{B\}}(X_{n+2}) \mid \mathcal{F}_{n+1}^X) \mid \mathcal{F}_n^X\} = \mathsf{E}(P\mathbf{1}_{\{B\}}(X_{n+1}) \mid \mathcal{F}_n^X) = P(P\mathbf{1}_{\{B\}})(X_n).$$

If $X_n = x$, the latter equals

$$\int_E P\mathbf{1}_{\{B\}}(y)P(x,dy) = \int_E P(y,B)P(x,dy) = P^2(x,B).$$

It follows that $\mathsf{P}\{X_{n+2} \in B \mid \mathcal{F}_n^X\} = P^2(X_n,B)$. Secondly, it makes sense to define *the Markov property straightaway for bounded, measurable functions!*

Continuous time Markov processes on a general state space Let us now go to the continuous time case. Then we cannot define one stochastic matrix determining the whole probabilistic evolution of the stochastic process considered. Instead, we have a collection of transition kernels $(P_t)_{t\in T}$ that should be related through the so-called Chapman-Kolmogorov equation to allow the Markov property to hold.

Definition 3.1.2 Let $(E,\mathcal{E})$ be a measurable space. A collection of transition kernels $(P_t)_{t\geq 0}$ is called a *(homogeneous) transition function* if for all $s,t\geq 0$, $x\in E$ and $B\in\mathcal{E}$

$$P_{t+s}(x,B)=\int P_s(x,dy)P_t(y,B).$$

This relation is known as the *Chapman-Kolmogorov relation.*

Translated to operator notation, the Chapman-Kolmogorov equation states that for a transition function $(P_t)_{t\geq 0}$ it holds that for every non-negative (or bounded) measurable function f and $s,t\geq 0$ we have

$$P_{t+s}f=P_t(P_sf)=P_s(P_tf).$$

In other words, the linear operators $(P_t)_{t\geq 0}$ form a *semigroup* of operators on the space of non-negative (or bounded) functions on E. In the sequel we will not distinguish between this semigroup and the corresponding (homogeneous) transition function on $(E,\mathcal{E})$, since there is a one-to-one relation between the two concepts.

Further notation Some further notation is enlightening. Let f,g,h be bounded (non-negative) measurable functions on E. *As argued before, P_tf is bounded, $\mathcal{E}/\mathcal{E}$-measurable.* Hence multiplying by g gives gP_tf, which is bounded, measurable. Here

$$gP_tf(x)=g(x)\cdot P_tf(x)=g(x)\int_y f(y)P_t(x,dy).$$

Then we can apply P_s to this function, yielding the bounded, measurable function P_sgP_tf, with

$$P_sgP_tf(x)=\int_y g(y)P_tf(y)P_s(x,dy)=\int_y g(y)\int_z f(z)P_t(y,dz)P_s(x,dy).$$

hP_sgP_tf is again bounded, measurable and we can integrate over the probability distribution ν on $(E,\mathcal{E})$:

$$\begin{aligned}\nu hP_sgP_tf &= \int_x h(x)P_sgP_tf(x)\nu(dx)\\ &= \int_x h(x)\int_y g(y)\int_z f(z)P_t(y,dz)P_s(x,dy)\nu(dx).\end{aligned}$$

To summarise, we get the following alternative notation for Markov processes that will be interchangedly used.

Put $\nu = \delta_x$, then

$$\begin{aligned}\mathsf{P}_x\{X_s \in A\} = \mathsf{E}_x \mathbf{1}_{\{A\}}(X_s) &= \mathsf{E}_x \mathsf{E}_x(\mathbf{1}_{\{A\}}(X_{s+0}) \,|\, \mathcal{F}_0) \\ &= \mathsf{E}_x P_s \mathbf{1}_{\{A\}}(X_0) = P_s \mathbf{1}_{\{A\}}(x) \\ &= \int_{y\in E} \mathbf{1}_{\{A\}}(y) P_s(x, dy) = P_s(x, A). \qquad (3.1.5)\end{aligned}$$

This can be generalised to: for $f \in b\mathcal{E}$

$$\mathsf{E}_x f(X_t) = \int f(y) P_t(x, dy) = P_t f(x),$$

and

$$\mathsf{E}_\nu f(X_t) = \int_x \int_y f(y) P_t(x, dy)\nu(dx) = \int_x P_t f(x)\nu(dx) = \nu P_t f.$$

Note that by the 'standard machinery', we may replace bounded f by non-negative f in the definition.

We can also let the initial distribution be given by the value X_s for some $s \geq 0$. This results in the following notation. For $f \in b\mathcal{E}$

$$\mathsf{E}_{X_s} f(X_t) = \int f(y) P_t(X_s, dy) = P_t f(X_s),$$

which we understand as being equal to $\mathsf{E}_x f(X_t) = P_t f(x)$ on the event $\{X_s = x\}$.

We can now give the definition of a Markov process.

Definition 3.1.3 Let $(E, \mathcal{E})$ be a measurable space and let X be an $(E, \mathcal{E})$-valued stochastic process that is adapted to some underlying filtered space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, \mathsf{P}_\nu)$. X is a **Markov process with initial distribution** ν, if

i) $\mathsf{P}_\nu\{X_0 \in B\} = \nu(B)$ for every $B \in \mathcal{E}$;

ii)(Markov property) there exists a transition function $(P_t)_t$, such that for all $s, t \geq 0$ and $f \in b\mathcal{E}$

$$\mathsf{E}_\nu(f(X_{t+s}) \,|\, \mathcal{F}_s) = P_t f(X_s)(= \mathsf{E}_{X_s} f(X_t)!) \quad \mathsf{P}_\nu - \text{a.s.} \qquad (3.1.6)$$

Definition 3.1.4 Let $(E, \mathcal{E})$ be a measurable space and let $X : (\Omega, \mathcal{F}) \to (E^{\mathbb{R}_+}, \mathcal{E}^{\mathbb{R}_+})$ be a map that is adapted to the filtration $(\mathcal{F}_t)_t$, with $\mathcal{F}_t \subset \mathcal{F}$, $t \geq 0$. X is a **Markov process**, if there exists a transition function $(P_t)_t$, such that for each distribution ν on $(E, \mathcal{E})$ there exists a probability distribution P_ν on $(\Omega, \mathcal{F})$ with

i) $\mathsf{P}_\nu\{X_0 \in B\} = \nu(B)$ for every $B \in \mathcal{E}$;

ii)(Markov property) for all $s, t \geq 0$ and $f \in b\mathcal{E}$

$$\mathsf{E}_\nu(f(X_{t+s}) \,|\, \mathcal{F}_s) = P_t f(X_s)(= \mathsf{E}_{X_s} f(X_t)), \quad \mathsf{P}_\nu - \text{a.s.} \qquad (3.1.7)$$

A main question is whether such processes exist, and whether sufficiently regular versions of these processes exist. As in the first chapter we will address this question by first showing that the fdd's of a Markov process (provided it exists) are determined by transition function and initial distribution. You have to realise further that a stochastic process with a transition function $(P_t)_t$ need not be Markov in general. The Markov property really is a property of the underlying stochastic process (cf. Example 3.1.1).

Lemma 3.1.5 *Let X be an $(E,\mathcal{E})$-valued stochastic process with transition function $(P_t)_{t\geq 0}$, adapted to the filtration $(\mathcal{F}_t)_t$. Let ν be a distribution on $(E,\mathcal{E})$.*

If X is Markov with initial distribution ν, then for all $0 = t_0 < t_1 < \cdots < t_n$, and all functions $f_0,\ldots,f_n \in b\mathcal{E}$, $n \in \mathbf{Z}_+$,

$$\mathsf{E}_\nu \prod_{i=0}^n f_i(X_{t_i}) = \nu f_0 P_{t_1-t_0} f_1 \cdots P_{t_n-t_{n-1}} f_n. \tag{3.1.8}$$

Vice versa, suppose that (3.1.8) holds for all $0 = t_0 < t_1 < \cdots < t_n$, and all functions $f_0,\ldots,f_n \in b\mathcal{E}$, $n \in \mathbf{Z}_+$. Then X is Markov with respect to the natural *filtration $(\mathcal{F}_t^X)_t$.*

In either case, (3.1.8) also holds for non-negative functions $f_0,\ldots,f_n \in m\mathcal{E}$.

Remark: the proof of the Lemma shows that is it sufficient to check (3.1.8) for indicator functions.

Proof. Let X be a Markov process with initial distribution ν. Then

$$\begin{aligned}\mathsf{E}_\nu \prod_{i=0}^n f_i(X_{t_i}) &= \mathsf{E}_\nu \mathsf{E}_\nu\Big(\prod_{i=0}^n f_i(X_{t_i}) \,|\, \mathcal{F}_{t_{n-1}}\Big) \\ &= \mathsf{E}_\nu \prod_{i=0}^{n-1} f_i(X_{t_i}) \mathsf{E}_\nu(f(X_{t_n}) \,|\, \mathcal{F}_{t_{n-1}}) \\ &= \mathsf{E}_\nu \prod_{i=0}^{n-1} f_i(X_{t_i}) P_{t_n-t_{n-1}} f_n(X_{t_{n-1}}).\end{aligned}$$

Now, $P_{t_n-t_{n-1}} f_n \in b\mathcal{E}$, and so one has

$$\begin{aligned}\mathsf{E}_\nu \prod_{i=0}^{n-1} f_i(X_{t_i}) P_{t_n-t_{n-1}} f_n(X_{t_{n-1}}) &= \mathsf{E}_\nu \prod_{i=0}^{n-2} f_i(X_{t_i}) \mathsf{E}_\nu(f_{n-1}(X_{t_{n-1}}) P_{t_n-t_{n-1}} f_n(X_{t_{n-1}}) \,|\, \mathcal{F}_{t_{n-2}}) \\ &= \mathsf{E}_\nu \prod_{i=0}^{n-2} f_i(X_{t_i}) P_{t_{n-1}-t_{n-2}} f_{n-1} P_{t_n-t_{n-1}} f_n(X_{t_{n-2}}).\end{aligned}$$

Iterating this yields

$$\begin{aligned}\mathsf{E}_\nu \prod_{i=0}^n f_i(X_{t_i}) &= \mathsf{E}_\nu f_0(X_{t_0}) P_{t_1-t_0} f_1 P_{t_2-t_1} f_1 \cdots P_{t_n-t_{n-1}} f_n(X_0) \\ &= \nu f_0 P_{t_1-t_0} f_1 P_{t_2-t_1} f_1 \cdots P_{t_n-t_{n-1}} f_n.\end{aligned}$$

Conversely, assume that (3.1.8) holds for all $0 = t_0 < t_1 < \cdots < t_n$, all functions $f_0,\ldots,f_n \in b\mathcal{E}$. We have to show that (i) $\mathsf{P}_\nu\{X_0 \in B\} = \nu(B)$ for all $B \in \mathcal{E}$, and that (ii) for any $s,t \geq 0$, all sets $A \in \mathcal{F}_s^X$

$$\mathsf{E}_\nu \mathbf{1}_{\{A\}} f(X_{t+s}) = \mathsf{E}_\nu \mathbf{1}_{\{A\}} P_t f(X_s). \tag{3.1.9}$$

Let $B \in \mathcal{E}$, put $n = 0$, $f_0 = \mathbf{1}_{\{B\}}$. (i) immediately follows.

We will show (ii). To derive (3.1.9), it is sufficient to check this for a π-system containing Ω and generating $\mathcal{F}_s^X$. As the π-system we take

$$\Big\{A = \{X_{t_0} \in A_0, X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n\} \,|\, t_0 = 0 < t_1 < \cdots < t_n \leq s, \\ A_i \in \mathcal{E}, i = 0, \ldots, n, n = 0, \ldots\Big\}$$

Let $f_i = \mathbf{1}_{\{A_i\}}$, then $\prod_{i=0}^n f_i(X_{t_i}) = \mathbf{1}_{\{X_{t_0} \in A_0, \ldots, X_{t_n} \in A_n\}}$ and so, assuming that $t_n < s$

$$\begin{aligned} \mathsf{E}_\nu \prod_{i=0}^n & \mathbf{1}_{\{A_i\}}(X_{t_i}) \mathbf{1}_{\{E\}}(X_s) f(X_{t+s}) \\ &= \nu \mathbf{1}_{\{A_0\}} P_{t_1 - t_0} \mathbf{1}_{\{A_1\}} P_{t_2 - t - 1} \cdots P_{t_n - t_{n-1}} \mathbf{1}_{\{A_n\}} P_{t+s-t_n} f \\ &= \nu \mathbf{1}_{\{A_0\}} P_{t_1 - t_0} \mathbf{1}_{\{A_1\}} \cdots P_{s - t_n}(P_t f) \\ &= \mathsf{E}_\nu \prod_{i=0}^n \mathbf{1}_{\{A_i\}}(X_{t_i})(P_t f)(X_s), \end{aligned}$$

which we wanted to prove. The reasoning is similar if $t_n = s$.

This implies that (3.1.9) holds for all sets A in a π-system generating $\mathcal{F}_s^X$, hence it holds for $\mathcal{F}_s^X$. Consequently, $\mathsf{E}_\nu(f(X_{t+s}) \,|\, \mathcal{F}_s^X) = P_t f(X_s)$, a.s. QED

Example 3.1.1 (Not a Markov process) Consider the following space

$$\mathsf{S} = \{(1,1,1), (2,2,2), (3,3,3), (1,2,3), (1,3,2), (2,3,1), (2,1,3), (3,1,2), (3,2,1)\},$$

with σ-algebra $\mathcal{S} = 2^S$. Putting $\mathsf{P}\{x\} = 1/9$ defines a probability measure on this space.

Define a sequence of i.i.d. random vectors $Z_k = (X_{3k}, X_{3k+1}, X_{3k+2})$, $k = 0, \ldots$ on $(\mathsf{S}, \mathcal{S}, \mathsf{P})$. Then the sequence $\{X_n\}_n$ is an $(E = \{1,2,3\}, \mathcal{E} = 2^E)$-valued stochastic process on $\{\mathsf{S}, \mathcal{S}, \mathsf{P}\}$ in discrete time. Then $\mathsf{P}\{X_{n+1} = j \,|\, \sigma(X_n)\} = 1/3$ for each $j \in \{1,2,3\}$ and n, meaning that the motion is determined by the 3×3 stochastic matrix with all elements equal to $1/3$.

Let $\{\mathcal{F}_n^X = \sigma(X_k, k \leq n)\}_n$ be the natural filtration generated by $\{X_n\}_n$. $\{X_n\}_n$ is not a Markov chain w.r.t $\{\mathcal{F}_n^X\}_n$, since $\mathsf{P}\{X_2 = 1 \,|\, \sigma(X_0, X_1)\} = f(X_0, X_1)$ with

$$f(X_0, X_1) = \begin{cases} 1, & (X_0, X_1) \in \{(1,1), (2,3), (3,2)\} \\ 0, & \text{otherwise.} \end{cases}$$

Hence $f(1,1) \neq f(2,1)$, thus showing that the Markov property lacks.

Example 3.1.2 (A (BM process)) Let W be a standard BM on an underlying probability space $(\Omega, \mathcal{F}, \mathsf{P})$. Let X_0 be a measurable random variable with distribution $\nu = \delta_x$, for some $x \in \mathbf{R}$, independent of W. Define $X_t = X_0 + W_t$, $t \geq 0$. Then $X = (X_t)_t$ is a Markov process with initial distribution ν with respect to its natural filtration. Note that $X_t - X_s = W_t - W_s$ is independent of $\mathcal{F}_s^X$.

To see that X is a Markov process, let f be a bounded, measurable function (on $\mathbf{R}$). Write $Y_t = W_{t+s} - W_s$. Then $Y_t \stackrel{\mathsf{d}}{=} \mathsf{N}(0,t)$ is independent of $\mathcal{F}_s^X$ and so with BN Lemma 7.5

$$\mathsf{E}_\nu(f(X_{t+s}) \,|\, \mathcal{F}_s^X) = \mathsf{E}_\nu\Big(f(Y_t + W_s + x) \,|\, \mathcal{F}_s^X) = g(X_s)$$

for the function g given by

$$\begin{aligned} g(z) &= \int_y \frac{1}{\sqrt{2\pi t}} f(y+z) e^{-y^2/2t} dy \\ &= \int_y \frac{1}{\sqrt{2\pi t}} f(y) e^{-(y-z)^2/2t} dy \\ &= P_t f(z) \end{aligned}$$

with P_t defined by

$$P_t f(z) = \int f(y) p(t,z,y) dy,$$

where

$$p(t,z,y) = \frac{1}{\sqrt{2\pi t}} e^{-(y-z)^2/2t}.$$

Hence

$$\mathsf{E}(f(X_{t+s}) \,|\, \mathcal{F}_s^X) = g(X_s) = P_t f(X_s) \quad \text{a.s.}$$

It is easily shown that P_t is a transition function. Fix $t \geq 0$. Measurability of $P_t(x,B)$ in x for each set $B = (-\infty, b)$, $b \in \mathbf{R}$, follows from continuity arguments. Together with $\mathbf{R}$, these sets form a π-system for $\mathcal{B}$. Then apply the d-system recipe to the set $\mathcal{S} = \{B \in \mathcal{B} \,|\, x \mapsto P_t(x,B) \text{is } \mathcal{E}/\mathcal{B}[0,1]\text{-measurable}\}$.

Example 3.1.3 ((B) Ornstein-Uhlenbeck process) Let W be a standard Brownian motion. Let $\alpha, \sigma^2 > 0$ and let X_0 be a $\mathbf{R}$-valued random variable with distribution ν that is independent of $\sigma(W_t, t \geq 0)$. Define the scaled Brownian motion by

$$X_t = e^{-\alpha t}(X_0 + W_{\sigma^2(\exp\{2\alpha t\}-1)/2\alpha}).$$

If $\nu = \delta_x$, $X = (X_t)_t$ a Markov process with the P_ν distribution of X_t a normal distribution with mean $\exp\{-\alpha t\}x$ and variance $\sigma^2(1 - e^{-2\alpha t})/2\alpha$. Note that $X_t \xrightarrow{\mathcal{D}} \mathsf{N}(0, \sigma^2/2\alpha)$.

If $X_0 \stackrel{\mathsf{d}}{=} \mathsf{N}(0, \sigma^2/2\alpha)$ then X_t is a Gaussian, Markov process with mean $m(t) = 0$ and covariance function $r(s,t) = \sigma^2 \exp\{-\alpha|t-s|\}/2\alpha$. See Exercise 3.1.

Example 3.1.4 ((C) Geometric Brownian motion) Let W be a standard BM on an underlying probability space $(\Omega, \mathcal{F}, \mathsf{P})$. Let X_0 be a measurable positive random variable with distribution ν, independent of W. Let $\mu \in \mathbf{R}$ and $\sigma^2 \in (0,\infty)$ and define $X_t = X_0 e^{(\mu-\sigma^2/2)t+\sigma W_t}$, $t \geq 0$. Then X is a Markov process. Depending on the value μ it is a (super/sub) martingale. Geometric Brownian motion is used in financial mathematics to model stock prices.

Example 3.1.5 (Poisson process) Let N be a Poisson process on an underlying probability space $(\Omega, \mathcal{F}, \mathsf{P})$. Let X_0 be a measurable random variable with distribution $\nu = \delta_x$, for some $x \in \mathbf{Z}_+$, independent of N. Define $X_t = X_0 + N_t$, $t \geq 0$. Then $X = (X_t)_{t\geq 0}$ is a Markov process with initial distribution ν, w.r.t. the natural filtration.

This can be shown in precisely the same manner as for BM (example 3.1.2A). In this case the transition function P_t is a stochastic matrix, $t \geq 0$, with

$$P_t(x, y) = \mathsf{P}\{N_t = y - x\}, \quad y \geq x.$$

(cf. Exercise 3.8).

In general it is not true that a function of a Markov process with state space $(E, \mathcal{E})$ is a Markov process. The following lemma gives a sufficient condition under which this is the case.

Lemma 3.1.4 *Let X be a Markov process with state space $(E, \mathcal{E})$, initial distribution ν and transition function $(P_t)_t$, defined on an underlying filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathsf{P}_\nu)$. Suppose that $(E', \mathcal{E}')$ is a measurable space and let $\phi : E \to E'$ be measurable and onto. If $(Q_t)_t$ is a collection of transition kernels such that*

$$P_t(f \circ \phi) = (Q_t f) \circ \phi$$

for all bounded, measurable functions f on E', then $Y = \phi(X)$ is a Markov process with respect to $(\mathcal{F}_t)_{t\geq 0}$, with state space $(E', \mathcal{E}')$, initial measure ν', with $\nu'(B') = \nu(\phi^{-1}(B'))$, $B' \in \mathcal{E}'$, and transition function (Q_t).

Proof. Let f be a bounded, measurable function on E'. By assumption and the semi-group property of (P_t),

$$(Q_t Q_s f) \circ \phi = P_t((Q_s f) \circ \phi) = P_t P_s (f \circ \phi) = P_{t+s}(f \circ \phi) = (Q_{t+s} f) \circ \phi.$$

Since ϕ is onto, this implies that $(Q_t)_t$ is a semigroup. It is easily verified that Y has the Markov property (see Exercise 3.3). QED

Example 3.1.6 (W_t^2 is a Markov process) We apply Lemma 3.1.6. In our example one has the function $\phi : E = \mathbf{R} \to E' = \mathbf{R}_+$ given by $\phi(x) = x^2$. The corresponding σ-algebras are simply the Borel-σ-algebras on the respective spaces.

If we can find a transition kernel Q_t, $t \geq 0$, such that

$$P_t(f \circ \phi)(x) = (Q_t f) \circ \phi(x), x \in \mathbf{R} \tag{3.1.10}$$

for all bounded, measurable functions f on $E' = \mathbf{R}_+$, then $\phi(W_t) = W_t^2$, $t \geq 0$, is a Markov process (w.r.t. its natural filtration).

Let f be a bounded, measurable function on $\mathbf{R}_+$. Then for $x \in \mathbf{R}$

$$\begin{aligned} P_t(f \circ \phi)(x) &= \int_{-\infty}^{\infty} p(t, x, y) f(y^2) dy \\ &= \int_0^{\infty} (p(t, x, y) + p(t, x, -y)) f(y^2) dy \\ &\overset{u=y^2 \Rightarrow y=\sqrt{u}, dy=du/2\sqrt{u}}{=} \int_0^{\infty} (p(t, x, \sqrt{u} + p(t, x, -\sqrt{u}) \frac{1}{2\sqrt{u}} f(u) du. \end{aligned} \tag{3.1.11}$$

Define for $y \in \mathbf{R}_+$, $B \in \mathcal{E}' = \mathcal{B}(\mathbf{R}_+)$

$$Q_t(y,B) = \int_B (p(t,\sqrt{y},\sqrt{u}) + p(t,\sqrt{y},-\sqrt{u}))\frac{1}{2\sqrt{u}}du.$$

One can check that $(Q_t)_{t\geq 0}$, is a transition kernel. Moreover, from (3.1.11) it follows for $x \in \mathbf{R}_+$ that

$$(Q_t f) \circ \phi(x) = (Q_t f)(x^2) = P_t(f \circ \phi)(x).$$

For $x < 0$ one has $p(t,x,y) + p(t,x,-y) = p(t,-x,y) + p(t,-x,-y)$ and so $P_t(f \circ \phi)(x) = P_t(f \circ \phi)(-x)$. Since $(Q_t f) \circ \phi(x) = (Q_t f)(x^2) = (Q_t f) \circ \phi(-x)$, the validity of (3.1.10) follows immediately.

3.2 Existence of a canonical version

The question is whether we can construct processes satisfying definition 3.1.3. In this section we show that this is indeed the case. In other words, for a given transition function $(P_t)_t$ and probability measure ν on a measurable space $(E,\mathcal{E})$, we can construct a so-called canonical Markov process X which has initial distribution ν and transition function $(P_t)_t$. We go back to the construction in Chapter 1. Note that, although we consider processes in continuous time, the results are valid as well for discrete time processes.

Recall that an E-valued process can be viewed as a random element of the space E^{R_+} of E-valued functions f on $\mathbf{R}_+$, or of a subspace $\Gamma \subset E^{\mathrm{R}_+}$ if X is known to have more structure. The σ-algebra $\Gamma \cap \mathcal{E}^{\mathrm{R}_+}$ is the smallest σ-algebra that makes all projections $f \to f(t)$ measurable.

As in Chapter 1, let $\Omega = \Gamma$ and $\mathcal{F} = \Gamma \cap \mathcal{E}^{\mathrm{R}_+}$. Consider the process $X = (X_t)_{t\geq 0}$ defined as the identity map

$$X(\omega) = \omega,$$

so that $X_t(\omega) = \omega_t$ is projection on the t-th coordinate. By construction $X : (\Omega,\mathcal{F}) \to (\Omega,\mathcal{F})$ and $X_t : (\Omega,\mathcal{F}) \to (E,\mathcal{E})$ are measurable maps. The latter implies that X is a stochastic process in the sense of Definition 1.1.1. X is adapted to the natural filtration $(\mathcal{F}_t^X = \Gamma \cap \mathcal{E}^{[0,t]})_t$. In a practical context, the path space, or a subspace, is the natural space to consider as it represents the process itself evolving in time.

Note that we have not yet defined a probability measure on $(\Omega,\mathcal{F})$. The Kolmogorov consistency theorem 1.2.3 validates the existence of a process on $(\Omega,\mathcal{F})$ with given fdds. Hence, we have to specify appropriate fdds based on the given transition function $(P_t)_t$ and initial distribution ν.

In order to apply this theorem, from this point on we will assume that $(E,\mathcal{E})$ is a Polish space, endowed with its Borel σ-algebra.

Corollary 3.2.2 (to the Kolmogorov consistency theorem) *Let $(P_t)_t$ be a transition function and let ν be a probability measure on $(E,\mathcal{E})$. Then there exists a unique probability measure P_ν on $(\Omega,\mathcal{F})$ such that under P_ν the canonical process X is a Markov process with initial distribution ν and transition function $(P_t)_t$ with respect to its natural filtration $(\mathcal{F}_t^X)_t$.*

Proof. For any n and all $0 = t_0 < t_1 < \cdots < t_n$ we define a probability measure on $(E^{n+1}, \mathcal{E}^{n+1})$ (see Exercise 3.6) by

$$\mu_{t_0\ldots,t_n}(A_0 \times A_1 \times \cdots \times A_n) = \nu \mathbf{1}_{\{A_0\}} P_{t_1-t_0} \mathbf{1}_{\{A_1\}} \cdots P_{t_n-t_{n-1}} \mathbf{1}_{\{A_n\}}, \quad A_0, \ldots, A_n \in \mathcal{E},$$

and on $(E^n, \mathcal{E}^n)$ by

$$\mu_{t_1,\ldots,t_n}(A_1 \times \cdots \times A_n) = \nu \mathbf{1}_{\{E\}} P_{t_1-t_0} \mathbf{1}_{\{A_1\}} \cdots P_{t_n-t_{n-1}} \mathbf{1}_{\{A_n\}}, \quad A_1, \ldots, A_n \in \mathcal{E}.$$

By the Chapman-Kolmogorov equation these probability measures form a consistent system (see Exercise 3.6). Hence by Kolmogorov's consistency theorem there exists a probability measure P_ν on $(\Omega, \mathcal{F})$, such that under P_ν the measures $\mu_{t_1,\ldots,t_n}$ are precisely the fdd's of the canonical process X.

In particular, for any n, $0 = t_0 < t_1 < \cdots < t_n$, and $A_0, \ldots, A_n \in \mathcal{E}$

$$\mathsf{P}\{X_{t_0} \in A_0, \ldots, X_{t_n} \in A_n\} = \nu \mathbf{1}_{\{A_0\}} P_{t_1-t_0} \cdots P_{t_n-t_{n-1}} \mathbf{1}_{\{A_n\}}.$$

By virtue of the remark following Lemma 3.1.5 this implies that X is Markov w.r.t. its natural filtration. QED

As the initial measure ν we can choose the Dirac measure δ_x at $x \in E$. By the above there exists a measure P_x on $(\Omega, \mathcal{F})$, such that the canonical process X has distribution P_x. This distributions has all mass on paths ω starting at x: $\omega_0 = x$. In words, we say that under P_x the process X starts at point x. Note that (cf. (3.1.5))

$$\mathsf{P}_x\{X_t \in A\} = P_t(x, A) = \int P_t(y, A)\delta_x(dy)$$

is a measurable function in x. In particular, since any distribution ν can be obtained as a convex combination of Dirac measures, we get

$$\mathsf{P}_\nu\{X_t \in A\} = \int P_t(y, A)\nu(dy) = \int \mathsf{P}_y\{X_t \in A\}\nu(dy).$$

Similarly, the fdd's of X under P_ν can be written as convex combination of the fdd's of X under P_x, $x \in E$. The next lemma shows that this applies to certain functions of X as well.

Lemma 3.2.3 *Let Z be an $\mathcal{F}^X_\infty$-measurable random variable, that is either non-negative or bounded. Then the map $x \to \mathsf{E}_x Z$ is $\mathcal{E}/\mathcal{B}$-measurable and for every initial distribution ν*

$$\mathsf{E}_\nu Z = \int_x \mathsf{E}_x Z \, \nu(dx).$$

Review BN §3 on monotone class theorems

Proof. Consider the collection of sets

$$\mathcal{S} = \{\Gamma \in \mathcal{F}^X_\infty \,|\, x \to \mathsf{E}_x \mathbf{1}_{\{\Gamma\}} \text{ is measurable and } \mathsf{E}_\nu \mathbf{1}_{\{\Gamma\}} = \int \mathsf{E}_x \mathbf{1}_{\{\Gamma\}} \nu(dx)\}.$$

It is easily checked that this is a d-system. The collection of sets

$$\mathcal{G} = \{\{X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n\} \mid A_1 \ldots, A_n \in \mathcal{E}, 0 \leq t_1 < \cdots < t_n, n \in \mathbf{Z}_+\}$$

is a π-system for $\mathcal{F}^X_\infty = \mathcal{E}^{\mathbb{R}_+}$. So if we can show that $\mathcal{G} \subset \mathcal{S}$, then by BN Lemma 3.8 $\mathcal{F}^X_\infty \subset \mathcal{S}$. But this follows from Lemma 3.1.5.

It follows that the statement of the lemma is true for $Z = \mathbf{1}_{\{\Gamma\}}$, $\Gamma \in \mathcal{F}^X_\infty$. Apply the standard machinery to obtain the validity of the lemma for $\mathcal{F}^X_\infty$-measurable bounded or non-negative random variables Z. See also Exercise 3.7. QED

This lemma allows to formulate a more general version of the Markov property. For any $t \geq 0$ we define the *translation or shift operator* $\theta_t : E^{\mathbb{R}_+} \to E^{\mathbb{R}_+}$ by

$$(\theta_t \omega)_s = \omega_{t+s}, \quad s \geq 0, \quad \omega \in E^{\mathbb{R}_+}.$$

So θ_t just cuts off the part of ω before time t and shifts the remainder to the origin. Clearly $\theta_t \circ \theta_s = \theta_{t+s}$.

Let $\Gamma \subset E^{\mathbb{R}_+}$ be such that $\theta_t(\Gamma) \subset \Gamma$ for each $t \geq 0$. Assume that X is a canonical Markov process on $(\Omega = E^{\mathbb{R}_+}, \mathcal{F} = \mathcal{E}^{\mathbb{R}_+} \cap \Gamma)$. In other words, for each distribution ν on $(E, \mathcal{E})$, there exists a probability distribution P_ν on $(\Omega, \mathcal{F})$, such that X is the canonical Markov process on $(\Omega, \mathcal{F}, \mathsf{P}_\nu)$ with initial distribution ν.
Note that $\mathcal{F}^X_t = \mathcal{E}^{[0,t]} \cap \Gamma$ and θ_t is $\mathcal{F}$-measurable for every $t \geq 0$ (why?).

Theorem 3.2.4 (Generalised Markov property for canonical process) *Assume that X is a canonical Markov process with respect to a filtration $(\mathcal{F}_t)_t$. Let Z be an $\mathcal{F}^X_\infty$-measurable random variable, non-negative or bounded. Then for every $t > 0$ and any initial distribution ν*

$$\mathsf{E}_\nu(Z \circ \theta_t \mid \mathcal{F}_t) = \mathsf{E}_{X_t} Z, \quad \mathsf{P}_\nu - \text{a.s.} \tag{3.2.1}$$

Before turning to the proof, note that we introduced new notation: $\mathsf{E}_{X_t} Z$ is a random variable with value $\mathsf{E}_x Z$ on the event $\{X_t = x\}$. By Lemma 3.2.3 this is a measurable function of X_t.

Proof. Fix an initial probability measure ν. We will first show that (3.2.1) holds for all $Z = \mathbf{1}_{\{B\}}$, $B \in \mathcal{F}^X_\infty$. Let

$$\mathcal{S} = \{B \in \mathcal{F}^X_\infty \mid \mathsf{E}_\nu(\mathbf{1}_{\{B\}} \circ \theta_t \mid \mathcal{F}_t) = \mathsf{E}_{X_t} \mathbf{1}_{\{B\}}, \quad \mathsf{P}_\nu - \text{a.s.}\}.$$

Then $\mathcal{S}$ is a d-system, since (i) $\Omega \in \mathcal{S}$, (ii)$B, B' \in \mathcal{S}$, $B \subseteq B'$, implies $B' \setminus B \in \mathcal{S}$, and (iii) for $B_n, n = 1, \ldots, \in \mathcal{F}^X_\infty$ a non-decreasing sequence of sets with $B_n \in \mathcal{S}$, $n = 1, 2, \ldots$, one has $\cup_n B_n \in \mathcal{S}$. Indeed, (ii) and (iii) follow from linearity of integrals and monotone convergence.

Recall that collection of all finite-dimensional rectangles, $\mathcal{A}$ say, is a π-system generating $\mathcal{F}^X_\infty$. Note that $B \in \mathcal{A}$ whenever there exist $n \in \mathbf{Z}_+$, $0 = s_0 < s_1 < \cdots < s_m$, $B_0, \ldots, B_m \in \mathcal{E}$, $n \in \mathbf{Z}_+$, such that $B = \{X_{s_0} \in B_0, \ldots, X_{s_m} \in A_m\}$. If we can show that $\mathcal{A} \subset \mathcal{S}$, then by BN Lemma 3.4 it follows that $\sigma(\mathcal{A}) = \mathcal{F}^X_\infty \subseteq \mathcal{S}$.

Take a finite-dimensional rectangle $B = \{X_{s_1} \in B_1, \ldots, X_{s_m} \in B_m\}$, where $0 \leq s_1 < \cdots < s_m$, $B_i \in \mathcal{E}$, $i = 1, \ldots, m$. Using Lemma 3.1.5, it follows that

$$\begin{aligned}
\mathsf{E}_\nu(\mathbf{1}_{\{B\}} \circ \theta_t \,|\, \mathcal{F}_t) &= \mathsf{E}_\nu(\prod_{i=1}^m \mathbf{1}_{\{B_i\}}(X_{s_i}) \circ \theta_t \,|\, \mathcal{F}_t) \\
&= \mathsf{E}_\nu(\prod_{i=1}^m \mathbf{1}_{\{B_i\}}(X_{t+s_i}) \,|\, \mathcal{F}_t) \\
&= \mathsf{E}_\nu(\mathbf{1}_{\{B_1\}} P_{s_2-s_1} \mathbf{1}_{\{B_2\}} \cdots P_{s_{m-1}-s_{m-2}} \mathbf{1}_{\{B_{m-1}\}} P_{s_m-s_{m-1}} \mathbf{1}_{\{B_m\}}(X_{t+s_1}) \,|\, \mathcal{F}_t) \\
&= P_{s_1} \mathbf{1}_{\{B_2\}} P_{s_2-s_1} \cdots P_{s_{m-1}-s_{m-2}} \mathbf{1}_{\{B_{m-1}\}} P_{s_m-s_{m-1}} \mathbf{1}_{\{B_m\}}(X_t) \\
&= \mathsf{E}_{X_t} \prod_{i=1}^m \mathbf{1}_{\{B_i\}}(X_{s_i}),
\end{aligned}$$

where we have consecutively conditioned on the σ-algebras $\mathcal{F}_{t+s_{m-1}}, \ldots, \mathcal{F}_{t+s_1}$, and used the Markov property. For the last equality we have used (3.1.8).

This results in having proved (3.2.1) for indicator functions Z. Apply the standard machinery to prove it for step functions (by linearity of integrals), non-negative functions, and bounded functions Z. QED

We end this section with an example of a Markov process with a countable state space.

Example 3.2.1 (Markov jump process) Let E be a countable state space with σ-algebra $\mathcal{E} = 2^E$ generated by the one-point sets. Let P be an $E \times E$ stochastic matrix. We define the transition function $(P_t)_t$ as follows:

$$P_t(x, y) = \sum_{n=0}^\infty e^{-\lambda t} \frac{(\lambda t)^n}{n!} P^{(n)}(x, y), \quad x, y \in E$$

where $P^{(n)} = (P)^n$ is the n-th power of P, and $P^{(0)} = \mathbf{I}$ is the identity matrix.

By virtue of Corollary 3.2.2 the canonical process X on $(E^{\mathbb{R}_+}, \mathcal{E}^{\mathbb{R}_+})$ with initial distribution ν is a Markov process with respect to its natural filtration.

The construction is as follows. Construct independently of X_0, a Poisson process N (cf. Chapter 1), starting at 0 and, independently, a discrete-time Markov chain Y with transition matrix P, with initial distribution δ_x. If $N_t = n$, then $X_t = Y_n$. Formally $X_t = \sum_{n=0}^\infty \mathbf{1}_{\{N_t=n\}} Y_n$. By construction X_t has right-continuous paths.

3.3 Strong Markov property

3.3.1 Strong Markov property

Let X be an $(E, \mathcal{E})$-valued Markov process on $(\Omega, \mathcal{F})$, adapted to the filtration $(\mathcal{F}_t)_t$ and transition function $(P_t)_t$. Assume that X has everywhere right-continuous paths, and that E is a Polish space, with $\mathcal{E}$ the Borel-σ-algebra.

Definition 3.3.1 X is said to have the **strong Markov property** if for every function $f \in b\mathcal{E}$, any adapted stopping time σ and any initial distribution ν and any $t \geq 0$

$$\mathbf{1}_{\{\sigma<\infty\}} \mathsf{E}_\nu(f(X_{\sigma+t}) \,|\, \mathcal{F}_\sigma) = \mathbf{1}_{\{\sigma<\infty\}} \mathsf{E}_{X_\sigma} f(X_t), \quad \mathsf{P}_\nu \text{ a.s.} \tag{3.3.1}$$

Note that we have to exclude the event $\{\sigma = \infty\}$, since X_t maybe not have a limit as $t \to \infty$ and hence it may not be possible to define the value X_∞ appropriately.

Lemma 3.3.2 *Let X be an $(E, \mathcal{E})$-valued Markov process. Then (3.3.1) holds for any function $f \in b\mathcal{E}$, any initial distribution ν and any stopping time σ, for which there exists a countable subset $S \subset [0, \infty)$ such that $\sigma \in S \cup \{\infty\}$.*

Proof. Any stopping time is optional, and so by the fact that σ is countably valued, it is easily checked that $A \in \mathcal{F}_\sigma$ if and only if $A \cap \{\sigma = s\} \in \mathcal{F}_s$ for each $s \in S$.

It is directly checked that $\mathbf{1}_{\{\sigma<\infty\}} X_\sigma$ is $\mathcal{F}_\sigma$-measurable. Use Lemma 3.2.3 to derive that the map $\omega \mapsto \mathbf{1}_{\{\sigma(\omega)<\infty\}} \mathsf{E}_{X_{\sigma(\omega)}} f(X_t)$ is $\mathcal{F}_\sigma/\mathcal{B}$-measurable as a composition of measurable maps. The next step is to show that

$$\mathsf{E}_\nu \mathbf{1}_{\{A\}} \mathbf{1}_{\{\sigma<\infty\}} f(X_{\sigma+t}) = \mathsf{E}_\nu \mathbf{1}_{\{A\}} \mathbf{1}_{\{\sigma<\infty\}} \mathsf{E}_{X_\sigma} f(X_t), \quad A \in \mathcal{F}_\sigma.$$

If $A \in \mathcal{F}_\sigma$ with $A \subset \{\sigma = s\}$ for some $s \in S$, then $A \in \mathcal{F}_s$. By the Markov property Theorem 3.2.4

$$\mathsf{E}_\nu \mathbf{1}_{\{A\}} \mathbf{1}_{\{\sigma<\infty\}} f(X_{\sigma+t}) = \mathsf{E}_\nu \mathbf{1}_{\{A\}} f(X_{s+t}) = \mathsf{E}_\nu \mathbf{1}_{\{A\}} \mathsf{E}_{X_s} f(X_t) = \mathsf{E}_\nu \mathbf{1}_{\{A\}} \mathbf{1}_{\{\sigma<\infty\}} \mathsf{E}_{X_\sigma} f(X_t).$$

Let $A \in \mathcal{F}_\sigma$ be arbitrary. By the previous $A \cap \{\sigma = s\} \in \mathcal{F}_\sigma$. Use that $A \cap \{\sigma < \infty\} = \cup_{s\in S}(A \cap \{\sigma = s\})$ and linearity of expectations. QED

Corollary 3.3.3 *Any discrete time Markov chain has the strong Markov property.*

Theorem 3.3.4 *Let X be a an $(E, \mathcal{E})$-valued Markov process adapted to the filtration $(\mathcal{F}_t)_t$ with E a Polish space and $\mathcal{E}$ the Borel-σ-algebra, and with right-continuous paths. Suppose that $x \mapsto \mathsf{E}_x f(X_s) = P_s f(x)$ is continuous, or, more generally, that $t \mapsto \mathsf{E}_{X_t} f(X_s)$ is right-continuous (everywhere) for each bounded continuous function f. Then the strong Markov property holds.*

Proof. Let σ be an $(\mathcal{F}_t)_t$-stopping time. We will first show the strong Markov property for $f \in b\mathcal{E}$, continuous. Then for indicator functions $\mathbf{1}_{\{B\}}$, $B \in \mathcal{E}$, with B closed, by an approximation argument. Since the closed sets form a π-system for $\mathcal{E}$, the d-system recipe provides us the strong Markov property for all of $\mathcal{E}$. The standard machinery finally provides us the result for $f \in b\mathcal{E}$.

Let first f be a bounded and continuous function, let $t \geq 0$. $\mathcal{F}_\sigma$-measurability of $\mathbf{1}_{\{\sigma<\infty\}} X_\sigma$ and $\mathbf{1}_{\{\sigma<\infty\}} \mathsf{E}_{X_\sigma} f(X_t)$ can be checked analogously to the previous lemma.

Consider

$$\sigma_m = \sum_{k=1}^{\infty} \frac{k}{2^m} \cdot \mathbf{1}_{\{\frac{k-1}{2^m}<\sigma\leq\frac{k}{2^m}\}} + \infty \cdot \mathbf{1}_{\{\sigma=\infty\}}.$$

Then σ_m takes countably many different values and $\sigma_m \downarrow \sigma$. By virtue of Lemma 3.3.2 for all $A \in \mathcal{F}_{\sigma_m}$

$$\mathsf{E}_\nu \mathbf{1}_{\{A\}} \mathbf{1}_{\{\sigma_m<\infty\}} f(X_{\sigma_m+t}) = \mathsf{E}_\nu \mathbf{1}_{\{A\}} \mathbf{1}_{\{\sigma_m<\infty\}} \mathsf{E}_{X_{\sigma_m}} f(X_t).$$

Next, use that if $A \in \mathcal{F}_\sigma$, then $A \in \mathcal{F}_{\sigma_m}$. Moreover,

$$\mathbf{1}_{\{A\}} \mathbf{1}_{\{\sigma_m<\infty\}} f(X_{\sigma_m+t}) \to \mathbf{1}_{\{A\}} \mathbf{1}_{\{\sigma<\infty\}} f(X_{\sigma+t})$$

and

$$\mathbf{1}_{\{A\}}\mathbf{1}_{\{\sigma_m<\infty\}}\mathsf{E}_{X_{\sigma_m}}f(X_t) \to \mathbf{1}_{\{A\}}\mathbf{1}_{\{\sigma<\infty\}}\mathsf{E}_{X_\sigma}f(X_t), \quad m \to \infty.$$

Apply dominated convergence.

Next we will show that the strong Markov property holds for $\mathbf{1}_{\{B\}}$, $B \in \mathcal{E}$. Let $B \in \mathcal{E}$ be a closed set and let f^m be given by

$$f^m(x) = 1 - m \cdot (m^{-1} \wedge d(x, B)), \quad m = 1, \ldots,$$

where d is a metric on E, consistent with the topology. Then $f^m \in b\mathcal{E}$ is continuous and by the previous

$$\mathsf{E}\mathbf{1}_{\{A\}}\mathbf{1}_{\{\sigma<\infty\}}f^m(X_t) \circ \theta_\sigma = \mathsf{E}\mathbf{1}_{\{A\}}\mathbf{1}_{\{\sigma<\infty\}}\mathsf{E}_{X_\sigma}f^m(X_t), \quad A \in \mathcal{F}_\sigma.$$

The random variable on the left-hand side converges pointwise to $\mathbf{1}_{\{A\}}\mathbf{1}_{\{\sigma<\infty\}}\mathbf{1}_{\{B\}}\circ\theta_\sigma$, the one on right-hand side converges pointwise to $\mathbf{1}_{\{A\}}\mathbf{1}_{\{\sigma<\infty\}}\mathsf{E}_{X_\sigma}\mathbf{1}_{\{B\}}$. Use monotone convergence. QED

Corollary 3.3.5 *Assume that X is a right-continuous process with a countable state space E, equipped with the discrete topology, and $\mathcal{E} = 2^E$. Then X has the strong Markov property.*

The corollary implies that the Poisson process has the strong Markov property, as well as the right-continuous Markov jump process.

Corollary 3.3.6 BM, *the Ornstein-Uhlenbeck process, and geometric* BM *have the strong Markov property.*

Without the required continuity properties, the strong Markov property may fail, as illustrated in Example 3.4.1. We discuss some other aspects of the strong Markov property in connection with optional times.

3.3.2 Intermezzo on optional times: Markov property and strong Markov property

Sometimes (eg. the book by Karatzas and Shreve) the strong Markov property is defined through optional times. In other words, the strong Markov property is defined to hold if for every function $f \in b\mathcal{E}$, any adapted optional time σ and any initial distribution ν and any $t \geq 0$

$$\mathbf{1}_{\{\sigma<\infty\}}\mathsf{E}_\nu(f(X_{\sigma+t}) \,|\, \mathcal{F}_{\sigma^+}) = \mathbf{1}_{\{\sigma<\infty\}}\mathsf{E}_{X_\sigma}f(X_t), \quad \mathsf{P}_\nu \text{ a.s.} \tag{3.3.2}$$

The conditions under which this alternative strong Markov property holds, are analogous to the conditions guaranteeing the strong Markov property (3.3.1) to hold by Lemma 1.6.7 (see Chapter 1 the new observation on the σ-algebra $\mathcal{F}_{\tau^+}$ and its characterisations in Exercises 1.34 and 1.35). In this case, you have to replace $\mathcal{F}_\sigma$ etc. by $\mathcal{F}_{\sigma^+}$ in the above statements. However, one needs to show that the Markov property holds with respect to the filtration $(\mathcal{F}_{t^+})_t$. This holds true under the conditions of Theorem 3.3.4.

Corollary 3.3.7 *Assume the conditions of Theorem 3.3.4.*

i) *Then X is Markov with respect to $(\mathcal{F}_{t^+})_t$.*

ii) *Let τ be an $(\mathcal{F}_t)_t$-optional time. Then (3.3.2) holds for any function $f \in b\mathcal{E}$.*

iii) *Let τ be a finite adapted optional time. Let Y be a bounded $\mathcal{F}_{\tau+}$-measurable random variable. Then*

$$\mathsf{E}_\nu Y f(X_{\tau+t}) = \mathsf{E}_\nu(Y \mathsf{E}_{X_\tau} f(X_t)).$$

Proof. See Exercise 3.10. QED

An interesting consequence is the following.

Lemma 3.3.8 *Assume the conditions of Theorem 3.3.4.*

i) Blumenthal's 0-1 Law *If $A \in \mathcal{F}^X_{0+}$ then $\mathsf{P}_x(A) = 0$ or 1 for all $x \in E$.*

ii) *If τ is an $(\mathcal{F}^X_t)_t$-optional time, then $\mathsf{P}_x\{\tau = 0\} = 0$ or 1, for all $x \in E$.*

Proof. See Exercise 3.10. QED

The generalised Markov property w.r.t. $(\mathcal{F}^X_{t+})_t$ is now an immediate consequence of Theorem 3.2.4, and the strong Markov property w.r.t. the same filtration is a consequence of Theorem 3.3.4.

3.3.3 Generalised strong Markov property for right-continuous canonical Markov processes

We are now interested in the question under what conditions the analogon of Theorem 3.2.4 hold. In order to formulate this analogon properly, we need introduce the concept of a random time shift. However, the use of time shifts restricts us to canonical processes.

Let X be a canonical Markov process w.r.t the filtration $(\mathcal{F}_t)_{t\geq 0}$, where we again assume the set-up described in the section 3.2, prior to Corollary 3.2.2 and Theorem 3.2.4. Suppose that X has everywhere right-continuous sample paths.

For a random time τ we now define θ_τ as the operator that maps the path $s \mapsto \omega_s$ to the path $s \mapsto \omega_{\tau(\omega)+s}$. If τ equals the deterministic time t, then $\tau(\omega) = t$ for all ω and so θ_τ equals the old operator θ_t.

Since the canonical process X is just the identity on the space Ω, we have for instance that $(X_t \circ \theta_\tau)(\omega) = X_t(\theta_\tau(\omega)) = (\theta_\tau)(\omega))_t = \omega_{\tau(\omega)+t} = X_{\tau(\omega)+t}(\omega)$, in other words $X_t \circ \theta_\tau = X_{\tau+t}$. So the operators θ_τ can still be viewed as time shifts.

The first results deals with countably valued stopping times.

Lemma 3.3.9 *Let X be a canonical Markov process. Then*

$$\mathbf{1}_{\{\sigma<\infty\}} \mathsf{E}_\nu(Z \circ \theta_\sigma \,|\, \mathcal{F}^X_\sigma) = \mathbf{1}_{\{\sigma<\infty\}} \mathsf{E}_{X_\sigma} Z \quad \mathsf{P}_\nu \text{ a.s..} \tag{3.3.3}$$

for any bounded or non-negative $\mathcal{F}^X_\infty$-measurable random variable Z, any initial distribution ν and any stopping time σ, for which there exists a countable subset $S \subset [0,\infty)$ such that $\sigma \in S \cup \{\infty\}$.

Proof. The map $\mathbf{1}_{\{\sigma<\infty\}}\mathsf{E}_{X_\sigma}Z$ is $\mathcal{F}^X_\sigma$-measurable if

$$\{\mathbf{1}_{\{\sigma<\infty\}}\mathsf{E}_{X_\sigma}Z \in B\}\cap\{\sigma=t\} = \{\mathsf{E}_{X_t}Z\in B\}\cap\{\sigma=t\}\in\mathcal{F}^X_t$$

for all $B\in\mathcal{B}$, $t\geq 0$. This follows from Lemma 3.2.3.

The next step is to show that

$$\mathsf{E}_\nu\mathbf{1}_{\{A\}}\mathbf{1}_{\{\sigma<\infty\}}Z\circ\theta_\sigma = \mathsf{E}_\nu\mathbf{1}_{\{A\}}\mathbf{1}_{\{\sigma<\infty\}}\mathsf{E}_{X_\sigma}Z,\quad A\in\mathcal{F}^X_\sigma.$$

This follows analogously to the similar assertion in the proof of Lemma 3.3.2, but with $Z\circ\theta_\sigma$ replacing $f(X_{\sigma+t})$. QED

As is the case with the strong Markov property, for general stopping times we need additional conditions.

Theorem 3.3.10 *Let X be a an $(E,\mathcal{E})$-valued canonical Markov process with E a Polish space and $\mathcal{E}$ the Borel-σ-algebra, and with right-continuous paths. Suppose that $x\mapsto\mathsf{E}_xf(X_s)=P_sf(x)$ is continuous, or, more generally, $t\mapsto\mathsf{E}_{X_t}f(X_s)$ is right-continuous everywhere for each bounded continuous function f. Then (3.3.3) holds for any bounded or non-negative $\mathcal{F}^X_\infty$-measurable random variable Z, any initial distribution ν and any stopping time σ.*

Proof. Let σ be an $(\mathcal{F}_t)_t$-adapted stopping time and ν an initial distribution. We will first prove the result for $Z=\prod_{i=1}^n\mathbf{1}_{\{A_i\}}(X_{t_i})$, with $n\in\mathbf{Z}_+$, $t_1<\cdots<t_n$, $A_1,\ldots,A_n\in\mathcal{E}$ by an induction argument. For $n=1$ the result follows from Theorem 3.3.4. For $n=2$

$$\begin{aligned}
\mathsf{E}_\nu(\mathbf{1}_{\{A_1\}}(X_{t_1})\mathbf{1}_{\{A_2\}}(X_{t_2})\circ\theta_\sigma\,|\,\mathcal{F}_\sigma) &= \mathsf{E}_\nu(\mathbf{1}_{\{A_1\}}(X_{\sigma+t_1})\mathbf{1}_{\{A_2\}}(X_{\sigma+t_2})\,|\,\mathcal{F}_\sigma)\\
&= \mathsf{E}_\nu(\mathbf{1}_{\{A_1\}}(X_{\sigma+t_1})\mathsf{E}_\nu(\mathbf{1}_{\{A_2\}}(X_{\sigma+t_2})\,|\,\mathcal{F}_{\sigma+t_1})\,|\,\mathcal{F}_\sigma)\\
&= \mathsf{E}_\nu(\mathbf{1}_{\{A_1\}}(X_{\sigma+t_1})\mathsf{E}_\nu(\mathbf{1}_{\{A_2\}}(X_{t_2-t_1})\circ\theta_{\sigma+t_1}\,|\,\mathcal{F}_{\sigma+t_1})\,|\,\mathcal{F}_\sigma)\\
&= \mathsf{E}_\nu(\mathbf{1}_{\{A_1\}}(X_{\sigma+t_1})\mathsf{E}_{X_{\sigma+t_1}}\mathbf{1}_{\{A_2\}}(X_{t_2-t_1})\,|\,\mathcal{F}_\sigma)\\
&= \mathsf{E}_{X_\sigma}g(X_{t_1}),
\end{aligned}$$

where on the event $X_{t_1}=x$

$$g(X_{t_1})=g(x)=\mathbf{1}_{\{A_1\}}(x)\mathsf{E}_x\mathbf{1}_{\{A_2\}}(X_{t_2-t_1})=\mathbf{1}_{\{A_1\}}P_{t_2-t_1}\mathbf{1}_{\{A_2\}}(x),$$

so that $g(X_{t_1})=\mathbf{1}_{\{A_1\}}P_{t_2-t_1}\mathbf{1}_{\{A_2\}}(X_{t_1})$. Further, on the event $X_\sigma=y$

$$\begin{aligned}
\mathsf{E}_{X_\sigma}g(X_{t_1}) &= \mathsf{E}_y\mathbf{1}_{\{A_1\}}P_{t_2-t_1}\mathbf{1}_{\{A_2\}}(X_{t_1})\\
&= P_{t_1}\mathbf{1}_{\{A_1\}}P_{t_2-t_1}\mathbf{1}_{\{A_2\}}(X_{t_1})(y)\\
&= \mathsf{E}_y\mathbf{1}_{\{A_1\}}(X_{t_1})\mathbf{1}_{\{A_2\}}(X_{t_2}).
\end{aligned}$$

This yields

$$\mathsf{E}_{X_\sigma}g(X_{t_1})=\mathsf{E}_{X_\sigma}\mathbf{1}_{\{A_1\}}(X_{t_1})\mathbf{1}_{\{A_2\}}(X_{t_2}),$$

which gives us the desired relation for $n=2$. The general induction step follows similarly.

Finally, we apply the d-system recipe to show that the strong Markov property holds for $Z=\mathbf{1}_{\{B\}}$ with $B\in\mathcal{F}^X_\infty$. Then use the standard machinery. QED

The strong Markov property has interesting consequences for right-continuous canonical Markov processes X with so-called stationary and independent increments. This means that $X_t - X_s$ is independent of $\mathcal{F}_s$ for $s \leq t$, and for each initial distribution ν, the P_ν-distribution of $X_t - X_s$ only depends on the difference $t - s$, and is independent of ν. In other words: the P_ν-distribution of $X_t - X_s$ and the P_μ distribution $X_{t-s} - X_0$ are equal for all initial distributions ν and μ, provided that E is an Abelian group.

The *Lévy processes* are a class of processes with this property of which canonical BM and the canonical Poisson process are well-known examples. In fact, one can prove that if X is a stochastic process satisfying the conditions of the lemma below, for $E = \mathbf{R}$, then X is a Gaussian process!

Lemma 3.3.11 *Let E be a Banach space. Let X be a right-continuous process with values in $(E, \mathcal{B}(E))$, defined on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, \mathsf{P})$. Suppose that X has stationary, independent increments.*

i) *X is a Markov process with initial distribution $\mathsf{P}\{X_0 \in \cdot\}$.*

ii) *Let τ be a finite $(\mathcal{F}_t)_t$-stopping time. Then the process $X(\tau) = (X_{\tau+t} - X_\tau)_{t\geq 0}$ is independent of $\mathcal{F}_\tau$. It is a Markov process, adapted to the filtration $(\mathcal{F}_{\tau+t})_t$. The distribution P_τ of $X(\tau)$ is the same as the distribution of $X - X_0$ under P_0.*

Proof. See Exercise 3.11. We give some hints for the proof of part (ii). Put $Y_t = X_{\tau+t} - X_\tau$, $t \geq 0$. For $t_1 < \cdots < t_n$ and functions $f_1, \ldots, f_n \in b\mathcal{E}$ we have

$$\begin{aligned} \mathsf{E}_\nu\Big(\prod_k f_k(Y_{t_k}) \,|\, \mathcal{F}_\tau\Big) &= \mathsf{E}_\nu\Big(\prod_k f_k(X_{\tau+t_k} - X_\tau) \,|\, \mathcal{F}_\tau\Big) \\ &= \mathsf{E}_{X_\tau} \prod_k f_k(X_{t_k} - X_0), \quad \mathsf{P}_\nu - \text{a.s.}, \end{aligned}$$

by the strong Markov property. As a consequence, the proof is complete once we have shown that for arbitrary $x \in E$

$$\mathsf{E}_x \prod_{k=1}^n f_k(X_{t_k} - X_0) = P_{t_1} f_1 \cdots P_{t_n - t_{n-1}} f_n(0),$$

(cf. Characterisation Lemma 3.1.5). Prove this by induction on n. QED

The following lemma is often useful in connection with the strong Markov property.

Lemma 3.3.12 *Let $\Omega = \Gamma, \mathcal{F} = \mathcal{E}^T \cap \Gamma$ with $\Gamma \subset E^T$ such that $\omega \in \Gamma$ implies (i) $\omega' = \theta_s(\omega) \in \Gamma$ for all $s \in T$ and (ii) for all $t \in T$ $\omega' = (\omega_{s\wedge t})_s \in \Gamma$. Let X be the canonical process, adapted to the filtration $(\mathcal{F}_t^X)_t$.*

i) *If σ and τ are finite $(\mathcal{F}_t^X)_t$-stopping times, then $\sigma + \tau \circ \theta_\sigma$ is also a finite $(\mathcal{F}_t^X)_t$-stopping time.*

ii) *If σ and τ are finite $(\mathcal{F}_t^+)_t$-optional times, then $\sigma + \tau \circ \theta_\sigma$ is also a finite $(\mathcal{F}_t^+)_t$-optional time.*

Proof. We will prove (i). By Lemma 1.6.5 (b) Galmarino test we have to show that (i) $(\sigma + \tau \circ \theta_\sigma)$ is $\mathcal{F}^X_\infty$-measurable and (ii) for all $t \geq 0$, $\sigma(\omega) + \tau(\sigma(\omega)) \leq t$ and $\omega_s = \omega'_s$, $s \leq t$, implies that $\sigma(\omega') + \tau(\sigma(\omega')) \leq t$.

The first statement is simply proved, for instance by a similar reasoning to Exercise 3.14. Let us consider the second statement.

For $t = 0$, $\sigma(\omega)+\tau(\sigma(\omega)) \leq 0$ then $\sigma(\omega), \tau(\omega) \leq 0$. Let $\omega'_0 = \omega_0$. Since σ is a stopping time, the Galmarino test applied to σ and τ implies that $\sigma(\omega'), \tau(\omega') \leq 0$, hence $\sigma(\omega')+\tau(\sigma(\omega')) \leq 0$.

Next let $t > 0$. Let ω' satisfy $\omega'_s = \omega_s$, $s \leq t$. Then $\sigma(\omega) \leq t$, and by the above and the Galmarino test, $\sigma(\omega') \leq t$.

Is het possible that $\sigma(\omega) < \sigma(\omega')$? Suppose this is the case. Then there exists $s < t$ such that $\sigma(\omega) \leq s < \sigma(\omega')$. But by the Galmarino test, $\sigma(\omega') \leq s$. A contradiction. Hence $\sigma(\omega') \leq \sigma(\omega)$. A repetition of the same argument with the roles of ω and ω' interchanged, yields that in fact $\sigma(\omega) = \sigma(\omega')$.

But then $\tau(\sigma(\omega)) = \tau(\sigma(\omega'))$. Hence $t \geq \sigma(\omega) + \tau(\sigma(\omega)) = \sigma(\omega') + \tau(\sigma(\omega'))$. This is what we had to prove.

The proof of (ii) is Exercise 3.14. QED

3.4 Applications to Brownian Motion

3.4.1 Reflection principle

The first example that we give, is the so-called reflection principle (compare with Ch.1, Exercise 1.12). First note that Lemma 3.3.11 implies that $(W_{\tau+t} - W_\tau)_t$ is BM, provided τ is an everywhere finite stopping time.

Recall that we denote the hitting time of $x \in \mathbf{R}$ by τ_x. This is an a.s. finite stopping time with respect to the natural filtration of the BM (see Example 1.6.1). The problem is that τ_x is not necessarily finite everywhere and hence we cannot conclude that $(W_{\tau_x+t} - W_\tau)_t$ is a BM, since it need not be defined everywhere.

The solution is to restrict to a smaller underlying space, but this trick might cause problems, if we need consider different initial distributions (a null-set under one initial distribution need not necessarily be a null-set under different initial distribution...). The simplest approach is via approximations: $(W_{\tau_x\wedge n+t} - W_{\tau_x\wedge n})_t$ is a BM, for each $n \geq 0$.

Theorem 3.4.1 (Reflection principle) *Let W be a Brownian motion with continuous paths. Let $x \in \mathbf{R}$ be given. Define the process W' by*

$$W'_t = \begin{cases} W_t, & t \leq \tau_x \leq \infty \\ 2x - W_t, & t > \tau_x. \end{cases}$$

Then W' is a standard BM *with continuous paths.*

Proof. If $x = 0$, $\tau_0 = 0$, and so the assertion is equivalent to symmetry of BM.

Let $x \neq 0$. Define processes Y^n and Z^n by $Y = W^{\tau_x \wedge n}$ and $Z^n_t = W_{\tau_x\wedge n+t} - W_{\tau_x\wedge n}$, $t \geq 0$. By Theorem 3.3.11 the processes Y^n and Z^n are independent and Z^n is a standard BM. By symmetry of BM, it follows that $-Z^n$ is also a BM that is independent of Y^n, and so the two pairs (Y^n, Z^n) and $(Y^n, -Z^n)$ have the same distribution (i.e. the fdd's are equal).

Now, for $t \geq 0$

$$W_t = Y_t^n + Z_{t-\tau_x \wedge n}^n \mathbf{1}_{\{t > \tau_x \wedge n\}}.$$

Define for $t \geq 0$

$$W_t'^{,n} = Y_t^n - Z_{t-\tau_x \wedge n}^n \mathbf{1}_{\{t > \tau_x \wedge n\}}.$$

The $W'^{,n}$ is W reflected about the value $W_{\tau_x \wedge n}$. By the continuity of Brownian motion paths, $W_t'^{,n} \to W_t'$, $n \to \infty$ a.s., for all $t \geq 0$.

Write $C_0[0,\infty) = \{\omega \in C[0,\infty) \mid \omega_0 = 0\}$. We have $W = \phi^n(Y^n, Z^n)$ and $W'^{,n} = \phi^n(Y^n, -Z^n)$, where $\phi^n : C[0,\infty) \times C_0[0,\infty) \to C[0,\infty)$ is given by

$$\phi^n(y,z)(t) = y(t) + z(t - \psi^n(y))\mathbf{1}_{\{t > \psi(y)\}},$$

where $\psi^n : C[0,\infty) \to [0,\infty]$ is defined by $\psi^n(y) = n \wedge \inf\{t > 0 \mid y(t) = x\}$. Consider the induced σ-algebra on $C[0,\infty)$ and $C_0[0,\infty)$ inherited from the σ-algebra $\mathcal{B}^{[0,\infty)}$ on $\mathbf{R}^{[0,\infty)}$. It is easily verified that ψ^n is a Borel-measurable map, and that ϕ^n is measurable as the composition of measurable maps (cf. Exercise 3.17). Since $(Y^n, Z^n) \stackrel{\mathsf{d}}{=} (Y^n, -Z^n)$, it follows that $W = \phi^n(Y^n, Z^n) \stackrel{\mathsf{d}}{=} \phi^n(Y^n, -Z^n) = W'^{,n}$.

On the time-interval $[0,n]$ clearly $W'^{,m}$ and W' have equal paths for $m > n$, and so they have the same fdd on $[0,n]$. This implies that the fdd of W' are multivariate Gaussian with the right mean and covariance functions. Since W' has continuous paths, we can invoke Lemma 1.4.3 to conclude that W' is BM with everywhere continuous paths. Note that the constructed process is not canonical). QED

The reflection principle allows us to calculate the distributions of certain functionals related to the hitting times of BM. We first consider the joint distribution of W_t and the running maximum

$$S_t = \sup_{s \leq t} W_s.$$

Corollary 3.4.2 *Let W be a standard* BM *and S its running maximum. Then*

$$\mathsf{P}\{W_t \leq x, S_t \geq y\} = \mathsf{P}\{W_t \leq x - 2y\}, \quad x \leq y.$$

The pair (W_t, S_t) has joint density

$$(x,y) \mapsto \frac{(2y-x)e^{-(2y-x)^2/2t}}{\sqrt{\pi t^3/2}} \mathbf{1}_{\{x \leq y\}},$$

with respect to the Lebesgue measure.

Proof. Let first $y > 0$. Let W' be the process obtained by reflecting W at the hitting time τ_y. Observe that $S_t \geq y$ if and only if $t \geq \tau_y$. Hence, the probability of interest equals $\mathsf{P}\{W_t \leq x, t \geq \tau_y\}$. On the event $\{t \geq \tau_y\}$ we have $W_t = 2y - W_t'$, and so we have to calculate $\mathsf{P}\{W_t' \geq 2y - x, t \geq \tau_y\}$. Since $x \leq y$, we have $2y - x \geq y$. hence $\{W_t' \geq 2y - x\} \subseteq \{W_t' \geq y\} \subseteq \{t \geq \tau_y\}$. It follows that $\mathsf{P}\{W_t' \geq 2y - x, t \geq \tau_y\} = \mathsf{P}\{W_t' \geq 2y - x\}$. By the reflection principle and symmetry of BM this proves the first statement for $y > 0$.

For $y = 0$, we have $\tau_y = 0$, and $W_t'' = -W_t$, showing the first statement directly.

The second statement follows from Exercise 3.18. QED

It follows from the preceding corollary that for all $x > 0$ and $t \geq 0$,

$$\mathsf{P}\{S_t \geq x\} = \mathsf{P}\{\tau_x \leq t\} = 2\mathsf{P}\{W_t \geq x\} = \mathsf{P}\{|W_t| \geq x\}$$

(see Exercise 3.18). This shows in particular that $S_t \stackrel{\mathsf{d}}{=} |W_t|$ for every $t \geq 0$. This allows to construct an example of a Markov process that lacks the strong Markov property.

Example 3.4.1 (Strong Markov property fails (Yushkevich)) Consider Example 3.1.2 (A). Instead of the canonical process X:

$$X_t = X_0 + W_t,$$

with $X_0 \stackrel{\mathsf{d}}{=} \nu$, independent of W, we consider

$$\tilde{X}_t = X_0 + \mathbf{1}_{\{X_0 \neq 0\}} W_t = \mathbf{1}_{\{X_0 \neq 0\}} X_t.$$

For initial distribution $\nu = \delta_x$, $x \neq 0$, the underlying distributions of X and $\tilde{X}$ are equal.

$\tilde{X}$ is a Markov process with respect to the same filtration and with transition function

$$\tilde{P}_t(x, B) = \begin{cases} \displaystyle\int_B \frac{1}{\sqrt{2\pi t}} e^{-(y-x)^2/2t} dy, & x \neq 0 \\ \delta_x\{B\}, & x = 0. \end{cases}$$

Suppose that $\tilde{X}$ has the strong Markov property. Let $\tau = \inf\{t \geq 0 \,|\, \tilde{X}_t = 0\} = \inf\{t \geq 0 |\, X_t = 0\}$. It is a stopping time for both $\tilde{X}$ and X. Consider the function $f = \mathbf{1}_{\{\mathbb{R}\setminus\{0\}\}}$. Clearly $f(X_1) = f(\tilde{X}_1)$ on $\mathbf{1}_{\{X_0 \neq 0\}}$. Then,

$$\mathbf{1}_{\{\tau < \infty\}} \mathsf{E}_{\tilde{X}_\tau} f(\tilde{X}_1) = 0, \tag{3.4.1}$$

by definition. Take initial distribution $\nu = \delta_x$ for some $x > 0$, and choose $A = \mathbf{1}_{\{\tau \leq 1\}}$. Then, $A \in \mathcal{F}_\tau$ and $\tau < \infty$ on A. By (3.4.1), $\mathsf{E}_x \mathbf{1}_{\{A\}} \mathsf{E}_{\tilde{X}_\tau} f(X_1) = 0$. However,

$$\begin{aligned} \mathsf{E}_x \mathbf{1}_{\{A\}} f(\tilde{X}_{\tau+1}) &= \mathsf{E}\mathbf{1}_{\{A\}} f(X_{\tau+1}) = \mathsf{E}_x \mathbf{1}_{\{A\}} \mathsf{E}_{X_\tau} f(X_1) \\ &= \mathsf{E}_x \mathbf{1}_{\{\tau \leq 1\}} \mathsf{P}_0\{X_1 \neq 0\} = \mathsf{P}_x\{\tau \leq 1\} = \mathsf{P}\{|W_1| \geq x\} > 0, \end{aligned}$$

a contradiction with the strong Markov property. The details have to be worked out in Exercise 3.16.

3.4.2 Ratio limit result

In this subsection W is the canonical BM on $(\Omega = \mathcal{C}[0,\infty), \mathcal{F} = \mathcal{C}(0,\infty] \cap \mathcal{B}^{\mathbb{R}_+})$ with associated Markov process X (cf. Example 3.1.2). Since BM has stationary, independent increments, Corollary 3.3.11 implies that for every $(\mathcal{F}_t)_t$-stopping time τ, the proces $(X_{\tau+t} - X_\tau)_t$ is a BM. This can be used to prove an interesting ratio limit result (originally derived by Cyrus Derman 1954).

To this end, let $A \in \mathcal{B}$ be a bounded set. Define $\mu(A, \tau) = \lambda\{t \leq \tau : X_t \in A\}$, where λ is the Lebesgue measure (on $(\mathbf{R}, \mathcal{B})$ and τ a finite $(\mathcal{F}_t)_t$-stopping time, w.r.t P_0. Denote $\tau_1 = \inf\{t > 0 \,|\, X_t = 1\}$, and (by abuse of previously introduced notation) $\tau_0 = \inf\{t > 0 \,|\, t \geq \tau_1, X_t = 0\}$.

Lemma 3.4.3 **i)** *$\mu(A,\tau_1)$ is a measurable function on $(\Omega,\mathcal{F})$.*

ii) $\mu(A) := \mathsf{E}_0\mu(A,\tau_1) = 2\int_{-\infty}^0 \mathbf{1}_{\{A\}}(x)d\lambda(x) + 2\int_0^1(1-x)\mathbf{1}_{\{A\}}(x)d\lambda(x)$.

iii) $\mu'(A) := \mathsf{E}_0\mu(A,\tau_0) = 2\lambda(A)$.

Proof. See exercise 3.15. For the proof of (i), it is sufficient to show that (explain)

$$(s,\omega)\mapsto \mathbf{1}_{\{A\}}(X_s(\omega))\mathbf{1}_{\{[0,\tau(\omega)]\}}(s)$$

is $\mathcal{B}[0,\infty]\times\mathcal{F}/\mathcal{B}$-measurable. To this end, show that $(Y_s(\omega) = \mathbf{1}_{\{[0,\tau(\omega)]\}}(s))_{s\geq 0}$ is an $\mathcal{F}/\mathcal{B}$-progressively measurable stochastic process.

For the proof of (ii), note that

$$\begin{aligned}
\mathsf{E}_0\mu(A,\tau_1) &= \mathsf{E}_0\int_0^\infty \mathbf{1}_{\{A\}}(X_t)\mathbf{1}_{\{[t,\infty)\}}(\tau_1)d\lambda(t) = \int_\Omega\int_0^\infty \mathbf{1}_{\{A\}}(X_t)\mathbf{1}_{\{[t,\infty)\}}(\tau_1)dt\,d\mathsf{P}_0\\
&= \int_0^\infty\int_\Omega \mathbf{1}_{\{A\}}(X_t)\mathbf{1}_{\{(t,\infty)\}}(\tau_1)d\mathsf{P}_0\,dt\\
&= \int_0^\infty \mathsf{P}_0\{X_t\in A, t<\tau_1\}dt\\
&= \int_0^\infty\int_A w(t,x)d\lambda(x)dt\\
&= \int_A\int_0^\infty w(t,x)dtd\lambda(x),
\end{aligned}$$

where

$$w(t,x) = \begin{cases} \frac{1}{\sqrt{2\pi t}}\left(e^{-x^2/2t} - e^{-(x-2)^2/2t}\right), & x\leq 1\\ 0, & x>1.\end{cases}$$

This follows from

$$\mathsf{P}_0\{X_t\in A\cap[-\infty,1],\tau_1<t\} = \mathsf{P}_2\{X_t\in A\cap[-\infty,1)\}.$$

Why is this true? (ii) can then be shown by writing

$$w(t,x) = -\mathbf{1}_{\{(-\infty,1)\}}(x)\int_{x-2}^x \frac{u}{t^{3/2}\sqrt{2\pi}}e^{-u^2/2t}du,$$

applying Fubini, and doing a substitution $s = t^{-1/2}$. Distinguish the cases that $x\leq 0$ and $0<x\leq 1$. QED

Let $f,g:\mathbf{R}\to\mathbf{R}$ be Lebesgue measurable, integrable functions with $\int_{\mathrm{R}} g(x)d\lambda(x)\neq 0$.

Theorem 3.4.4

$$\lim_{T\to\infty}\frac{\int_0^T f(W_t)dt}{\int_0^T g(W_s)ds} = \frac{\int_R f(x)d\lambda(x)}{\int_R g(x)d\lambda(x)},\quad \text{a.s.}$$

Proof. Put $\tau_0^1 = \tau_0$ and $\tau_1^1 = \tau_1$. Inductively define for $n \geq 2$: $\tau_0^n = \inf\{t \geq \tau_1^n \mid X_t = 0\}$, and $\tau_1^n = \inf\{t \geq \tau_0^{n-1} \mid X_t = 1\}$. By virtue of the standard machinery, one has

$$\mathsf{E}_0 \int_0^{\tau_0^1} f(X_t)dt = 2 \int_{\mathbb{R}} f(x)dx.$$

Now, for any $T \geq 0$ define

$$K(T) = \max\{n \mid \tau_0^n \leq T\}.$$

Then $\lim_{T\to\infty} \int_0^T f(X_t)dt/K(T) = 2\int_{\mathbb{R}} f(x)dx$, P_0-a.s. The result then follows. QED

3.4.3 Hitting time distribution

We also may derive an explicit expression for the density of the hitting time τ_x. It is easily seen from this expression that $\mathsf{E}\tau_x = \infty$, as was proved by martingale methods in Exercise 2.30 of Chapter 2.

Corollary 3.4.5 *The first time τ_x that the standard* BM *hits the level $x > 0$ has density*

$$t \mapsto \frac{xe^{-x^2/2t}}{\sqrt{2\pi t^3}} \mathbf{1}_{\{t\geq 0\}},$$

with respect to the Lebesgue measure.

Proof. See Exercise 3.19. QED

We have seen in the first two Chapters that the zero set of standard BM is a.s. closed, unbounded, has Lebesgue measure zero and that 0 is an accumulation point of the set, i.e. 0 is not an isolated point. Using the strong Markov property we can prove that in fact the zero set contains no isolated point at all.

Corollary 3.4.6 *The zero set $Z = \{t \geq 0 \mid W_t = 0\}$ of standard* BM *is a.s. closed, unbounded, contains no isolated points and has Lebesgue measure 0.*

Proof. In view of Exercise 1.29, we only have to prove that Z contains no isolated points. For rational $q \geq 0$, define $\sigma_q = q + \tau_0 \circ \theta_q$. Hence, σ_q is the first time after (or at) time q that BM visits 0. By Lemma 3.3.12 the random time σ_q is an optional time. The strong Markov property implies that $W_{\sigma_q+t} - W_{\sigma_q}$ is a BM w.r.t $(\mathcal{F}_{t+}^X)_t$. By Corollary 2.4.6 it follows that σ_q a.s. is an accumulation point of Z. Hence, with probability 1 it holds that for every rational $q \geq 0$, σ_q is an accumulation point of Z. Now take an arbitrary point $t \in Z$ and choose rational points q_n such that $q_n \uparrow t$. Since $q_n \leq \sigma_{q_n} \leq t$, we have $\sigma_{q_n} \to t$. The limit of accumulation points is an accumulation point. This completes the proof. QED

3.4.4 Embedding a random variable in Brownian motion

This subsection discusses the set-up of a result by Skorokhod, that a discrete martingale $M = (M_n)_{n=0,\ldots}$ with independent increments can be 'embedded' in Brownian motion. Dubins[2] derived a more general result with an elegant proof, part of which we will discuss next.

[2]L.E. Dubins, On a Theorem by Skorohod, *Ann. Math. Stat.* **39**, 2094–2097, 1968. Together with L.J. Savage author of the famous book 'How to Gamble if you must'

We consider standard BM. The filtration to be considered is the natural filtration $(\mathcal{F}_t^W)_t$ generated by BM itself.

Let X be an integrable random variable (with values in $(\mathbf{R},\mathcal{B})$) on an underlying probability space. Let μ denote the induced distribution of X on $(\mathbf{R},\mathcal{B})$. Denote $\mathsf{E}X = m$.

Our aim construct an $(\mathcal{F}_t^W)_t$-stopping time τ, such that $m + W_\tau$ has distribution μ, i.e. $m + W_\tau \stackrel{\mathsf{d}}{=} X$. The method is based on constructing a suitable Doob martingale with limiting distribution μ and then duplicate the same construction within BM.

The martingale that does the trick Put $\mathcal{G}_0 = \{\Omega,\emptyset\}$, partition $G_0 = \{\mathbf{R}\}$ of $\mathbf{R}$ and let $\mathcal{S}_0 = \{m\} = \mathsf{E}(X \mid \mathcal{G}_0)(\Omega) = S_0$.

We construct a filtration and martingale iteratively as follows. We are going to associate a Doob martingale $M = (M_n)_{n=0,1,\ldots}$ with X as follows. At time n construct inductively

- a partition of $\mathbf{R}$ denoted G_n;
- the σ-algebra $\mathcal{G}_n = \sigma(X^{-1}(A), A \in G_n)$;
- $S_n = \left\{\frac{\mathsf{E}\mathbf{1}_{\{X^{-1}(A)\}}X}{\mathsf{P}\{X^{-1}(A)\}} \mid A \in G_n\right\}$.

Note that
$$\frac{\mathsf{E}\mathbf{1}_{\{X^{-1}(A)\}}X}{\mathsf{P}\{X^{-1}(A)\}} = \frac{\int_A x dP_X(x)}{P_X(A)},$$
where P_X is the induced probability distribution of X on $(\mathbf{R},\mathcal{B})$.

Given G_n, the construction of $\mathcal{G}_n$ and the finite collection of points S_n is immediate. We therefore have to prescribe the initial values, and the iterative construction of the partitions.

For $n = 0$: $G_0 = \{\mathbf{R}\}$, $\mathcal{G}_0 = \{\emptyset,\Omega\}$, $S_0 = \{m\}$. The partition G_{n+1} has the following properties. For all $A \in G_n$, let $x_A = S_n \cap A$.

- If $\mathsf{P}\{X^{-1}(\{x_A\})\} = \mathsf{P}\{X^{-1}(A)\}$ then $A \in G_{n+1}$;
- if $\mathsf{P}\{X^{-1}(\{x_A\})\} < \mathsf{P}\{X^{-1}(A)\}$ then $A \cap (-\infty, x_A]$, $A \cap (x_A,\infty) \in G_{n+1}$.

This construction defines a filtration $(\mathcal{G}_n)_n \subset \mathcal{F}$ and a Doob martingale $(M_n = \mathsf{E}(X \mid \mathcal{G}_n))_n$. By the Levy upward theorem 2.2.15 $\mathsf{E}(X \mid \mathcal{G}_n) \to \mathsf{E}(X \mid \mathcal{G}_\infty)$ a.s. and in L^1.

Lemma 3.4.7 $X \stackrel{\mathcal{D}}{=} \mathsf{E}(X \mid \mathcal{G}_\infty)$.

For the proof, see the additional exercise.

As a consequence, the constructed Doob martingale determines the distribution of a random variable uniquely. By interpreting the sequence of realised values as a branching process, built up through a binary tree, it becomes easier to understand the embedding construction.

Binary tree The root of the tree is chosen to be a node with value m. In the construction of $\mathcal{G}_1$, with each interval associated with m, we put an out-arrow and a node at level 1. With the out-arrow we associate the corresponding interval, and with the node, the corresponding value of $\mathsf{E}(X \mid \mathcal{G}_1)$. Repeat this construction.

Note that if $A_n \in G_n$ is an interval associated with the in-arrow of a vertex in level n with value s_n, then all values associated with vertices in the binary tree rooted at s_n are contained in A_n.

Furthermore the values $\mu(A_n)$, $A_n \in G_n$, $n = 0, \ldots$, can be iteratively reconstructed from the binary tree, since no two vertices have the same value. Indeed, suppose A_n is associated with the arrow from the vertex with value s_{n-1} to the vertex with value s_n. Either $s_n = s_{n-1}$ in which case $A_n = A_{n-1}$, the interval associated with the in-arrow of s_{n-1}, and so $\mu(A_n) = \mu(A_{n-1})$. Or, s_{n-1} has a second out-arrow, associated with $A_{n-1} \setminus A_n$, leading to a vertex with value s'_n say. Since $X^{-1}(A_{n-1}) \in \mathcal{G}_{n-1} \subset \mathcal{G}_n$,

$$\begin{aligned} s_{n-1} \cdot \mu(A_{n-1}) = \mathsf{E}\mathbf{1}_{\{X^{-1}(A_{n-1})\}}\mathsf{E}(X \,|\, \mathcal{G}_{n-1}) &= \mathsf{E}\mathbf{1}_{\{X^{-1}(A_{n-1})\}}X \\ &= \mathsf{E}\mathbf{1}_{\{X^{-1}(A_{n-1})\}}\mathsf{E}(X \,|\, \mathcal{G}_{n+1}) \\ &= s_n \cdot \mu(A_n) + s'_n \cdot \mu(A_{n-1} \setminus A_n). \end{aligned}$$

Calculation yields $\mu(A_n) = \frac{\mu(A_{n-1})(s_{n-1}-s'_n)}{s_n - s'_n}$.

Let μ_∞ denote the distribution of $\mathsf{E}(X \,|\, \mathcal{G}_\infty)$. By the construction of the binary tree, it follows immediately that (check!)

$$X^{-1}(A) = M_n^{-1}(A), \quad A \in G_n, \quad m \geq n, \tag{3.4.2}$$

so that $\mu(A) = \mathsf{P}\{X \in A\} = \mathsf{P}\{M_\infty \in A\} = \mu_\infty(A)$, $A \in \cup_n G_n$. A final remark is that $\mathcal{G}_n = \sigma(\mathsf{E}(X \,|\, \mathcal{G}_n))$, since $\mathsf{E}(X \,|\, \mathcal{G}_n)$ is constant on the sets $X^{-1}(A_n)$, $A_n \in G_n$. Next define

$$\mathcal{T}_n = (\mathsf{E}(X \,|\, \mathcal{G}_0), \ldots, \mathsf{E}(X \,|\, \mathcal{G}_n)(\mathbf{R}),$$

the (finite) collection of realisations of the martingale $(M_k)_k$ upto time n, or alternatively the values along all paths of the root m to level n.

Corollary 3.4.8 *Let Y be a random variable with distribution π. Define an associated Doob martingale in the same manner, and a collection $\mathcal{T}_n^Y$ of realisations upto time n, $n = 0, \ldots,$. Then $\mu = \pi$, iff $\mathcal{T}_n = \mathcal{T}_n^Y$ for all n, i.o.w. iff the associated binary trees are equal.*

Embedding of X For $\omega \in \Omega$ let

$$\tau(\omega) = \inf\{t \geq 0 \,|\, \forall n = 0, 1, 2, \ldots \exists t_0 = 0 \leq \cdots \leq t_n \leq t, \text{ s. t. } (W_{t_0}(\omega), \ldots, W_{t_n}(\omega)) \in \mathcal{T}_n - m\},$$

where $\mathcal{T}_n - m$ means that m is substracted from the values along the binary tree, we called these the 'reduced' values. In the binary tree, $(W_{t_0}(\omega), \ldots, W_{t_n}(\omega))$ is a sequence of reduced values along paths in the binary tree from the root m with end-points in level n. We further define random times τ_n iteratively by putting $\tau_0 = 0$ and

$$\tau_{n+1}(\omega) = \inf\{t \geq \tau_n(\omega) \,|\, W_{\tau_{n+1}(\omega)}(\omega) \in S_{n+1} - m\}.$$

It is easily verified that τ_n, $n \geq 0$ and τ are a.s. finite stopping times and that $\{\tau_n\}_n$ is a non-decreasing sequence converging to a.s. to τ (if $\tau_n(\omega) = \infty$, we put $\tau_m(\omega) = \infty$, for all $m \geq n$). Use unboundedness properties of BM paths, the continuity and, when viewing BM as a canonical process with the natural filtration, the Galmarino test. Note that $\tau_0 = 0$! As a consequence, $W_{\tau_n} \to W_\tau$, a.s.

Next, let $s_n^u = \max\{x \,|\, \in S_n\}$ and $s_n^l = \min\{x \,|\, x \in S_n\}$. Then $\tau_n \leq \tau_{s^l_{n+1}-m} \wedge \tau_{s^u_{n+1}-m}$, the latter having finite expectation. As a consequence $W^{\tau_n} = (W_{t \wedge \tau_n})_t$ is a bounded martingale.

Hence it is UI. By virtue of Theorem 2.2.14, $\mathsf{E}W_{\tau_n} = \mathsf{E}W_0 = 0$. Applying the optional sampling theorem 2.3.12 with stopping times τ_n and ∞ to $W^{\tau_{n+1}}$ yields

$$\mathsf{E}(W_\infty^{\tau_{n+1}} \,|\, \mathcal{F}_{\tau_n}) = W_{\tau_n}^{\tau_{n+1}}, \quad \text{a.s.}$$

i.o.w.

$$\mathsf{E}(W_{\tau_{n+1}} \,|\, \mathcal{F}_{\tau_n}) = W_{\tau_n}, \quad \text{a.s.}$$

By the tower property, it follows that $\mathsf{E}(W_{\tau_{n+1}} \,|\, \mathcal{F}_n) = W_{\tau_n}$, for $\mathcal{F}_n = \sigma(W_{\tau_n})$. By continuity of BM paths, $W_{\tau_n}(\omega)$ and $W_{\tau_{n+1}}(\omega)$ have subsequent values along a path in the binary tree at levels n and $n+1$. This implies that $\mathcal{F}_n \subset \mathcal{F}_{n+1}$. Therefore, $\{(W_{\tau_n})_n$ is an $(\mathcal{F}_n)_n$-martingale.

Theorem 3.4.9 **i)** *$\{\{W_{\tau_n}\}_n, W_\tau\}$ is an $(\mathcal{F}_n)_n$-martingale with $W_{\tau_n} \stackrel{a.s.}{\to} W_\tau$.*

ii) *$m + W_\tau \stackrel{\mathsf{d}}{=} X$.*

iii) *$\mathsf{E}\tau = \mathsf{E}X^2$.*

Under the assumption that X has finite variance, the theorem is known as *Skorokhod's first embedding theorem*. The present proof does not require this.

Proof. We first assume that $X \geq 0$ and show $\mathsf{E}W_\tau = 0$. By a.s. unboundedness of BM paths, $W_{\tau_n} \to W_\tau$ a.s. Since $\{W_{\tau_n}\}_n$ is a martingale bounded below by $-m$, with $\mathsf{E}|m + W_{\tau_n}| = \mathsf{E}(m + W_{\tau_n}) = m$, it is bounded in L^1 and so W_τ is integrable.

Let $s_n^u = \max\{x \,|\, x \in S_n\}$. If $\sup_n s_n^u < \infty$, then $\{W_{\tau_n}\}_n$ is a bounded martingale, hence UI. It follows that W_τ can be appended to the martingale sequence, and $\mathsf{E}W_\tau = 0$.

Suppose that $\sup_n s_n^u = \infty$. Let $s_n^{u-} = \max\{x \neq s_n^u \,|\, x \in S_n\}$, be the next largest value of S_n. From the binary tree we may infer for ω with $W_{\tau_n(\omega)}(\omega) \leq s_n^{u-} - m$, that $W_{\tau_m(\omega)}(\omega) \leq s_{n-1}^u - m$ for all $m \geq n$.

Consider the process $Y_l = W_{\tau_l} \mathbf{1}_{\{(-m, s_n^{u-} - m]\}}(W_{\tau_n})$, $l \geq n$. This is an $(\mathcal{F}_l)_{l \geq n}$ bounded martingale. Indeed,

$$\mathsf{E}(Y_{l+1} \,|\, \mathcal{F}_l) = \mathbf{1}_{\{(-m, s_n^{u-} - m]\}}(W_{\tau_n}) \mathsf{E}(W_{\tau_{l+1}} \,|\, \mathcal{F}_l) = \mathbf{1}_{\{(-m, s_n^{u-} - m]\}}(W_{\tau_n}) W_{\tau_l}.$$

Hence, $\{Y_l\}_{l \geq n}$ converges a.s. and in L^1 to $\mathbf{1}_{\{(-m, s_n^{u-} - m]\}}(W_{\tau_n}) W_\tau$, and

$$\mathsf{E}(\mathbf{1}_{\{(-m, s_n^{u-} - m]\}}(W_{\tau_n}) W_\tau \,|\, \mathcal{F}_n) = \mathbf{1}_{\{(-m, s_n^{u-} - m]\}}(W_{\tau_n}) W_{\tau_n}.$$

Taking expectations on both sides

$$\mathsf{E}W_\tau = -(s_n^u - m)\mathsf{P}\{W_{\tau_n} = s_n^u - m\} + \mathsf{E}\mathbf{1}_{\{(s_n^{u-} - m, \infty)\}}(W_{\tau_n}) W_\tau. \tag{3.4.3}$$

Note that $\{W_{\tau_n}\}_n$ being a martingale implies $\mathsf{P}\{W_{\tau_n} = x\} = \mathsf{P}\{X \in A_n\}$, where $x \in S_n \cap A_n - m$, for some $A_n \in G_n$. This yields

$$(s_n^u - m)\mathsf{P}\{W_{\tau_n} = s_n^u - m\} = \mathsf{E}\{(X - m)\mathbf{1}_{\{[s_{n-1}^u, \infty)\}}\} \to 0, \text{ as } n \to \infty. \tag{3.4.4}$$

Furthermore, from the binary tree

$$\mathsf{E}\mathbf{1}_{\{(s_n^{u-} - m, \infty)\}}(W_{\tau_n}) W_\tau = \mathsf{E}\mathbf{1}_{\{(s_n^{u-} - m, \infty)\}}(W_\tau) W_\tau \to 0, \quad n \to \infty. \tag{3.4.5}$$

Combining (3.4.3), (3.4.4) and (3.4.5) yields $\mathsf{E}W_\tau = 0$.

By Theorem 2.3.15 W_τ can be appended to $\{W_{\tau_n}\}_n$ to yield a supermartingale. That is, $\mathsf{E}(W_\tau \mid \mathcal{F}_n) \leq W_{\tau_n}$. Taking expectations yields equality, hence the extended sequence is a martingale, in particular a Doob martingale. In other words $W_{\tau_n} = \mathsf{E}(W_\tau \mid \mathcal{F}_n)$. The general case that X is not necessarily non-negative follows simply (see Exercise 3.20). This completes the proof of (i).

It then follows that the binary tree generated by $m + W_\tau$ is equal to the binary tree generated by X. Hence $X \stackrel{\mathsf{d}}{=} m + W_\tau$, thus proving (ii).

The proof of (iii) is an exercise (Exercise 3.20). QED

Note that for (i,ii) we have not used that the embedding process is BM. It is sufficient that it is a continuous time martingale with continuous paths that are unbounded above and below. However, essentially such a martingale is a Gaussian process, so the result is not as general as it looks.. We will finally state a simple version of the second Skorokhod embedding theorem.

Theorem 3.4.10 *Let $\{X_n\}_n$ be a sequence of i.i.d. integrable random variables. Put $S_n = \sum_{k=1}^n X_k$. Then there exists a non-decreasing sequence $\{\tau_n\}_n$ of a.s. finite stopping times, such that for all n*

i) $(W_{\tau_1}, W_{\tau_2}, \ldots, W_{\tau_n}\} \stackrel{\mathsf{d}}{=} (S_1, S_2, \ldots, S_n)$;

ii) $\tau_1, \tau_2 - \tau_1, \tau_3 - \tau_3, \ldots, \tau_n - \tau_{n-1}$ *are i.i.d. random variables with mean* $\mathsf{E}X_1^2$.

Proof. see Exercise 3.21. QED

This theorem can be e.g. used for an alternative proof of the Central Limit Theorem, and for deriving the distribution of $\max_{1 \leq k \leq n} S_k$.

3.5 Exercises

Exercise 3.1 Consider the Ornstein Uhlenbeck process in example 3.1.2(B). Show that the defined process is a Markov process which converges in distribution to an $\mathsf{N}(0, \sigma^2/2\alpha)$ distributed random variable. If $X_0 \stackrel{\mathrm{d}}{=} \mathsf{N}(0, \sigma^2/2\alpha)$, show that $X_t \stackrel{\mathrm{d}}{=} \mathsf{N}(0, \sigma^2/2\alpha)$ (in other words: the $\mathsf{N}(0, \sigma^2/2\alpha)$ distribution is an *invariant distribution* for the Markov process). Show that X_t is a Gaussian process with the given mean and covariance functions.

Exercise 3.2 Consider geometric BM in example 3.1.2(C). Show that it is a Markov process. Characterise the values of μ for which it is a super- a sub- or a mere martingale. Compute the transition function.

Exercise 3.3 Complete the proof of Lemma 3.1.4.

Exercise 3.4 Let W be a BM. Show that the *reflected Brownian motion* defined by $X = |X_0 + W|$ is a Markov process with respect to its natural filtration and compute its transition function. (Hint: calculate the conditional probability $\mathsf{P}_\nu\{X_t \in B \mid \mathcal{F}_s^X)$ by conditioning further on $\mathcal{F}_s^W$).

Exercise 3.5 Let X be a Markov process with state space E and transition function $(P_t)_{t\geq 0}$. Show that for every bounded, measurable function f on E and for all $t \geq 0$, the process $(P_{t-s}f(X_s))_{s\in[0,t]}$ is a martingale.

Exercise 3.6 Prove that $\mu_{t_1,\ldots,t_n}$ defined in the proof of Corollary 3.2.2 are probability measures that form a consistent system. Hint: for showing that they are probability measures, look at the proof of the Fubini theorem.

Exercise 3.7 Work out the details of the proof of Lemma 3.2.3.

Exercise 3.8 Show for the Poisson process X with initial distribution $\nu = \delta_x$ in Example 3.1.5, that X is a Markov process w.r.t. the natural filtration, with the transition function specified in the example.

Exercise 3.9 Show Corollary 3.3.6 that canonical Brownian motion has the strong Markov property.

Exercise 3.10 Prove Corollary 3.3.7 and Lemma 3.3.8.

Exercise 3.11 Prove Lemma 3.3.11. See hint in the 'proof'.

Exercise 3.12 Let X be a canonical, right-continuous Markov process with values in a Polish state space E, equipped with Borel-σ-algebra $\mathcal{E}$. Assume $t \mapsto \mathsf{E}_{X_t} f(X_s)$ right-continuous everywhere for each bounded continuous function $f : E \to \mathbf{R}$. For $x \in E$ consider the random time $\sigma_x = \inf\{t > 0 \mid X_t \neq x\}$.

i) Is σ_x a stopping time or an optional time? Using the Markov property, show that for every $x \in E$

$$\mathsf{P}_x\{\sigma_x > t + s\} = \mathsf{P}_x\{\sigma_x > t\}\mathsf{P}_x\{\sigma_x > s\},$$

for all $s, t \geq 0$.

ii) Conclude that there exists an $a \in [0, \infty]$, possibly depending on x, such that

$$\mathsf{P}_x(\sigma_x > t) = e^{-at}.$$

Remark: this leads to a classification of the points in the state space of a right-continuous canonical Markov process. A point for which $a = 0$ is called an *absorption point* or a *trap*. If $a \in (0, \infty)$, the point is called a *holding point*. Points for which $a = \infty$ are called *regular*.

iii) Determine a for the Markov jump process (in terms of λ and the stochastic matrix P) (cf. Example 3.2.1) and for the Poisson process. Hint: compute $\mathsf{E}_x \sigma_x$.

iv) Given that the process starts in state x, what is the probability that the new state is y after time σ_x for Markov jump process?

Exercise 3.13 Consider the situation of Exercise 3.12. Suppose that $x \in E$ is a holding point, i.e. a point for which $a \in (0, \infty)$.

i) Observe that $\sigma_x < \infty$, P_x-a.s. and that $\{X_{\sigma_x} = x, \sigma_x < \infty\} \subseteq \{\sigma_x \circ \theta_{\sigma_x} = 0, \sigma_x < \infty\}$.

ii) Using the strong Markov property, show that

$$\mathsf{P}_x\{X_{\sigma_x} = x, \sigma_x < \infty\} = \mathsf{P}_x\{X_{\sigma_x} = x, \sigma_x < \infty\}\mathsf{P}_x\{\sigma_x = 0\}.$$

iii) Conclude that $\mathsf{P}_x\{X_{\sigma_x} = x, \sigma_x < \infty\} = 0$, i.e. a canonical Markov process with right-continuous paths, satisfying the strong Markov property can only leave a holding point by a jump.

Exercise 3.14 Prove Lemma 3.3.12 (ii). Hint: suppose that Z is an $\mathcal{F}_t^X$-measurable random variable and σ an $(\mathcal{F}_t^X)_t$-optional time. Show that $Z \circ \theta_\sigma$ is $\mathcal{F}_{(\sigma+t)+}^X$-measurable. First show this for Z an indicator, then use the appropriate monotone class argument.

Exercise 3.15 Prove Lemma 3.4.3 and Theorem 3.4.4.

Exercise 3.16 Show for Example 3.4.1 that X is a Markov process, and show the validity of the assertions stated. Explain which condition of Theorem 3.3.4 fails in this example.

Exercise 3.17 Show that the maps ϕ and ψ in the proof of Theorem 3.4.1 are Borel measurable.

Exercise 3.18 **i)** Derive the expression for the joint density of BM and its running maximum given in Corollary 3.4.2.

ii) Let W be a standard BM and S_t its running maximum. Show that for all $t \geq 0$ and $x > 0$

$$\mathsf{P}\{S_t \geq x\} = \mathsf{P}\{\tau_x \leq t\} = 2\mathsf{P}\{W_t \geq x\} = \mathsf{P}\{|W_t| \geq x\}.$$

Exercise 3.19 Prove Corollary 3.4.5.

Exercise 3.20 Consider the construction of the binary tree in §3.4.4. Construct the binary tree for a random variable X, that has a uniform distribution on $[0, 1]$.

Consider Theorem 3.4.9. Show how the validity of statements (i,ii) follow for a general integrable random variable given that these statements are true for non-negative or non-positive integrable random variables. Show (iii).

Exercise 3.21 Prove Theorem 3.4.10.

Chapter 4

Generator of a Markov process with countable state space

4.1 The generator

In the case of a discrete Markov chain, the transition kernel with $t = 1$ (which is a stochastic matrix) completely determines the finite dimensional distribution of the process and hence the distribution of the process. The question arises whether the situation for continuous time processes is analogous: is the distribution of the process determined by one operator? In general the answer is no, but under certain conditions it will be yes.

A clue to the answer to this problem lies in the results discussed in Exercises 3.12, 3.13. Let X be a canonical, right-continuous Markov process with values in a Polish state space E, equipped with Borel-σ-algebra $\mathcal{E}$. Assume $t \mapsto \mathsf{E}_{X_t} f(X_s)$ right-continuous everywhere for each bounded continuous function $f : E \to \mathbf{R}$. For $x \in E$ there exists $q_x \in [0, \infty]$ such that

$$\mathsf{P}_x(\sigma_x > t) = e^{-q_x t},$$

with $\sigma_x = \inf\{t \geq 0 \mid X_t \neq x\}$ is the holding or sojourn time at x. A point x with $q_x = \infty$ is called regular, a non-regular point x, i.e. $q_x < \infty$, is called stable. If x is stable, then $X_{\sigma_x} \neq x$ a.s.

Definition 4.1.1 X is stable if x is stable for all $x \in E$.

It would seem that a stable Markov process X is completely determined by q_x and the distribution of X_{σ_x}. Is this true? How can one obtain q_x and the distribution of X_{σ_x} from the transition function $(P_t)_t$?

We will study this problem first in the case that E is a countable space, equipped with the discrete topology, and $\mathcal{E} = 2^E$ the σ-algebra generated by the one-point sets. The transition kernel P_t is an $E \times E$ matrix with elements denoted by $P_t(x, y)$, $x, y \in E$, $t \geq 0$, where now $P_t(x, B) = \sum_{y \in B} P_t(x, y)$.

Each real-valued function is continuous, and the Markov process is trivially strong Markov. There are counter-examples to the counterintuitive result that X need not be stable. We quote some definitions and results.

Definition 4.1.2 X is called *standard* if $\lim_{t \geq 0} P_t(x, y) = \mathbf{1}_{\{x\}}(y)$.

The following result is stated without proof.

Lemma 4.1.3 *Let X be standard and stable. Then $t \mapsto P_t(x,y)$ is continuously differentiable. Define $Q = P_t'|_{t=0}$ to be the generator of X. Then $Q = (q_{xy})_{x,y \in E}$ satisfies*

i) $q_{xx} = -q_x$;

ii) $0 \le q_{xy} < \infty$ *for* $y \ne x$;

iii) $\sum_y q_{xy} \le 0$.

iv) $\sum_y P_t'(x,y) = 0$.

This gives us the parameters of the holding times of the Markov process. (i) can be proved by an application of Exercise 3.12. Notice that necessarily $q_{xy} \ge 0$ if $y \ne 0$.

It makes sense to require that $\sum_y q_{xy} = (d/dt) \sum_y P_t(x,y)\Big|_{t=0} = 0$ (i.o.w. we may interchange sum and limit). If this is the case, then X is called *conservative.*

From now on assume that X is standard, conservative, stable and has right-continuous paths.

Continuation of Example 3.1.5 Poisson process The transition kernel of the Poisson process with parameter λ is given by $P_t(x,y) = P\{N_t = y - x\} = e^{-\lambda t}(\lambda t)^{y-x}/(y-x)!$, for $y \ge x$, $x, y \in \mathbf{Z}_+$. $t \mapsto P_t(x,y)$ is a continuously differentiable function. Differentiation yields

$$q_{x,y} = \begin{cases} -\lambda, & y = x \\ \lambda, & x+1 \\ 0, & \text{otherwise.} \end{cases}$$

Hence the Poisson process is standard, stable and conservative.

In general the situation is quite complicated, even for countable state space. First we will restrict to the case of bounded rates.

4.2 Bounded rates: $\sup_x q_x < \infty$.

Lemma 4.2.1 *The Kolmogorov forward and backward equations hold: $P_t' = P_t Q = Q P_t$.*

Proof. Use the Fatou lemma on

$$\frac{P_{t+h}(x,y) - P_t(x,y)}{h} = \sum_{k \ne x} \frac{P_h(x,k)}{h} P_t(k,y) - \frac{1 - P_h(x,x)}{h} P_t(x,y)$$

and

$$\frac{P_{t+h}(x,y) - P_t(x,y)}{h} = \sum_{k \ne x} P_t(x,k) \frac{P_h(k,y)}{h} - P_t(x,y) \frac{1 - P_h(x,y)}{h}.$$

This gives that $P_t'(x,y) \ge \sum_k q_{xk} P_t(k,y)$ and $P_t'(x,y) \ge \sum_k P_t(x,k) q_{ky}$. Taking the summation over all states in the first equation gives

$$0 \le \sum_y \Big(P_t'(x,y) - \sum_k q_{xk} P_t(k,y)\Big)$$

$$\begin{aligned} &= \sum_y (P_t'(x,y) - q_{xx}P_t(x,y) - \sum_{k\neq x} q_{xk}P_t(x,y)) \\ &= 0 - q_x \sum_y P_t(x,y) - \sum_{k\neq x} q_{xk} \sum_y P_t(x,y) \\ &= 0 - \sum_k q_{xk} = 0, \end{aligned}$$

hence $P_t'(x,y) = \sum_k q_{xk}P_t(k,y)$. In the one-but-last inequality we have used Fubini 's theorem and Lemma 4.1.3 (iv). This yields the backward equation. The forward equation is more problematic: we have to deal with the term $\sum_y \sum_k P_t(x,k)q_{ky}$. However, since $0 \leq q_{ky} \leq q_k \leq \sup_k q_k < \infty$, the rates are bounded. Hence $\sum_y \sum_k P_t(x,k)|q_{ky}| < \infty$ and so by Fubini's theorem we may interchange the order of summation. QED

The Kolmogorov forward equation implies that the following integral equation holds:

$$\begin{aligned} \mathsf{E}_x \mathbf{1}_{\{y\}}(X_t) = P_t(x,y) &= \mathbf{1}_{\{x\}}(y) + \int_0^t \sum_k P_t(x,k)q_{ky}ds \\ &= \mathbf{1}_{\{x\}}(y) + \int_0^t P_t(Q\mathbf{1}_{\{y\}})(x)ds \\ &= \mathbf{1}_{\{x\}}(y) + \mathsf{E}_x \int_0^t Q\mathbf{1}_{\{y\}}(X_s)ds, \end{aligned}$$

where we have used boundedness of the function $x \mapsto Q\mathbf{1}_{\{y\}}(x)$ and Fubini's theorem to allow for the interchange of integral and summation. Using Fubini's theorem in the same manner, one can prove analogously for each bounded function f that

$$\mathsf{E}_x f(X_t) = f(x) + \mathsf{E}_x \int_0^t Qf(X_s)ds. \tag{4.2.1}$$

Lemma 4.2.2 *Let X be defined on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t)$. For every bounded function $f : E \to \mathbf{R}$ and initial measure ν, the process*

$$M_t = f(X_t) - f(X_0) - \int_0^t Qf(X_s)ds, \quad t \geq 0$$

is a P_ν-martingale.

Proof. Exercise 4.1. QED

As a consequence the celebrated Dynkin's formula holds.

Corollary 4.2.3 Dynkin's formula *Under the conditions of Lemma 4.2.2, for any $(\mathcal{F}_t)_t$-stopping time with $\mathsf{E}_x\tau < \infty$, it holds that*

$$\mathsf{E}_x f(X_\tau) = f(x) + \mathsf{E}_x \int_0^\tau Qf(X_s)ds.$$

Proof. For $\tau \wedge n$ the optional sampling theorem applies, hence

$$\mathsf{E}_x f(X_{\tau\wedge n}) = f(x) + \mathsf{E}_x \int_0^{\tau\wedge n} Qf(X_s)ds.$$

Since $X_{\tau\wedge n} \to X_\tau$, by dominated convergence the left-hand side converges to $\mathsf{E}_x f(X_\tau)$. For the right-hand side, note that

$$|\int_0^{\tau\wedge n} Qf(X_s)ds| \leq \tau \cdot \max_y |Qf(y)| < \infty.$$

Use dominated convergence to complete the proof. QED

This can be applied to compute the distribution of X_{σ_x}. If $q_x = 0$ then x is an absorbing state, and $\sigma_x = \infty$, and so $X_{\sigma_x} = x$. Suppose that $q_x > 0$, then σ_x is a. a.s. finite optional time with $\mathsf{E}_x = q_x^{-1}$, hence it is a stopping time w.r.t. $(\mathcal{F}_{t+})_t$. Since X is also Markov w.r.t. $(\mathcal{F}_{t+})_t$, $(M_t)_t$ is an $(\mathcal{F}_{t+})_t$-martingale. One has for $y \neq x$

$$\mathsf{E}_x \mathbf{1}_{\{y\}}(X_{\sigma_x}) = \mathbf{1}_{\{y\}}(x) + \mathsf{E}_x \int_0^{\sigma_x} \mathbf{1}_{\{y\}}(X_s)ds = \mathsf{E}_x \int_0^{\sigma_x} q_{X_s y}ds = \mathsf{E}_x \sigma_x \cdot q_{xy} = \frac{q_{xy}}{q_x}.$$

In Exercise 3.12 you have in fact been asked to compute the generator Q of a Markov jump process (cf. Example 3.2.1).

4.3 Construction of Markov processes with given generator Q

It is convenient to define jump times of the Markov process, before discussing how to construct a Markov process from the generator. To this end, let $J_0 = 0$, and define recursively

$$J_{n+1} = \inf\{t \geq J_n \mid X_t \neq \lim_{s\uparrow t} X_s\},$$

where $J_{n+1} = \infty$ if X_{J_n} is an absorbing state.

For the construction, we are given an $E \times E$ matrix Q, with the properties

i) $q_{xx} \in (-\infty, 0]$ for all x (i.e. Q is stable);

ii) $q_{xy} \geq 0$, $y \neq x$ for all x;

iii) $\sum_y q_{xy} = 0$ for all x (i.e. Q is conservative).

First, *check for T an* $\exp(1)$ *distributed random variable, that T/c has an* $\exp(c)$ *distribution, for any constant $c \in [0, \infty)$.*

Let T_n, $n = 0, \ldots$, be a sequence of i.i.d. $\exp(1)$ distributed random variables, and, independently, let Y be a discrete time Markov chain with transition matrix P_J, defined by $P_J(x, y) = q_{xy}/q_x$, $y \neq x$, all defined on the same space $(\Omega, \mathcal{F}, \mathsf{P})$. The consecutive jump times are equal to: $J_0 = 0$ and

$$J_n = \sum_{k=0}^{n-1} T_k/q_{Y_k}, \quad n \geq 1.$$

Then

$$X_t = x \iff Y_n = x, \quad \text{and } J_n \leq t < J_{n+1}, \tag{4.3.1}$$

where $\sum_{k=0}^{-1} \cdots = 0$. I.o.w. X resides an $\exp(q_{Y_0})$ amount of time in Y_0, then jumps to state Y_1 with probability $q_{Y_0 Y_1}/q_{Y_0}$, etc.

The construction of the Markov chain Y can be included in the construction of X, in the following way. First, we need the following fact.

Lemma 4.3.1 *If $Z_1, \ldots, Z_n$ are independent random variables, with exponential distributions with successive parameters $c_1, \ldots, c_n$. Then $\min(Z_1, \ldots, Z_n)$ has an $\exp(\sum_1^n c_i)$ distribution and $\mathsf{P}\{Z_i = \min(Z_1, \ldots, Z_n)\} = c_i / \sum_{k=1}^n c_k$.*

Let now be given $\{T_{n,z}\}_{n=0,1,\ldots,z\in E}$ be i.i.d. exp(1) distributed random variables, and independently an E-valued random variable Y_0. Let

$$\begin{aligned} J_1 &= \min\{\frac{T_{0,z}}{q_{Y_0 z}} | z \in E, q_{Y_0 z} > 0\} \\ Y_1 &= y \quad \text{if } \frac{T_{0,y}}{q_{Y_0 y}} = \min\{\frac{T_{0,z}}{q_{Y_0 z}} | z \in E, q_{Y_0 z} > 0\}. \end{aligned}$$

Iteratively, let

$$\begin{aligned} J_{n+1} &= J_n + \min\{\frac{T_{n,z}}{q_{Y_n z}} | z \in E, q_{Y_n z} > 0\} \\ Y_{n+1} &= y \quad \text{if } \frac{T_{n,y}}{q_{Y_n y}} = \min\{\frac{T_{n,z}}{q_{Y_n z}} | z \in E, q_{Y_n z} > 0\}. \end{aligned}$$

In view of Lemma 4.3.1, we have guaranteed that Y is a Markov chain with transition matrix P_J. Define X_t now as in (4.3.1).

It is by no means guaranteed that it is a stochastic process! In fact X can be substochastic, so that probability mass is 'lost'. However, one can check that X is Markov, i.o.w. there exists a possibly substochastic transition function $\{f_t\}_t$ determining the fdd's of X, and X has the Markov property (follows from the construction). Note that we can make a substochastic transition function stochastic, by extending E with an *absorbing coffin state* to which all disappearing probability mass is directed.

The transition function of the constructed Markov process can be obtained by the procedure described below. Let

$$f_t^{(n)}(x,y) = \begin{cases} \mathbf{1}_{\{x\}}(y) e^{-q_x t}, & n = 0 \\ f_t^{(0)}(x,y) + \int_0^t e^{-q_x s} \sum_{k\neq x} q_{xk} f_{t-s}^{(n-1)}(k,y) ds, & n \geq 1. \end{cases} \tag{4.3.2}$$

$f_t^{(n)}(x,y)$ can be interpreted as the probability that the process X, given that it starts in x, is in state y at time t, after having made *at most* n jumps. The sequence $\{f_t^{(n)}(x,y)\}_n$ is monotonically non-decreasing, it is bounded by 1, and therefore has a limit $\{f_t(x,y)\}_{x,y}$. We have that $\sum_y f_t(x,y) \leq 1$. The transition function of the constructed process is equal to $(f_t)_{t\geq 0}$. Precisely this transition function defines the so-called minimal process.

Definition 4.3.2 X is said to be the minimal process, if the transition function of X is $(f_t)_t$.

Theorem 4.3.3 *$\{(f_t = (f_t(x,y))_{x,y}\}_t$ is the minimal solution to the Kolmogorov forward and backward equations, in the sense that if $(P_t)_t$ is a solution to either the Kolmogorov forward or the Kolmogorov backward equations, then $P_t(x,y) \geq f_t(x,y)$, $x, y \in E$, $t \geq 0$. It is the unique solution if $\sum_y f_t(x,y) = 1$ for all $t \geq 0$. In the latter case, the generator Q uniquely defines the transition function for which Q is the derivative at time 0.*

Lemma 4.3.4 *$\sum_y f_t(x,y) = 1$, for $x \in E$ and $t \geq 0$.*

This allows to derive a specifically bounded jump related expression for the transition function $(P_t)_t$.

Corollary 4.3.5 *Let $\tau \geq \sup_x q_x$. Let $P = \mathbf{I} + Q/\tau$. Then*

$$P_t = e^{-\tau t} \sum_{n\geq 0} \frac{(\tau t)^n}{n!} P^n, \quad t \geq 0,$$

I.o.w. X is a Markov jump process (see Example 3.2.1).

Proof. Show that $P_t'|_{t=0} = Q$. QED

Proof of Theorem 4.3.3. We give a partial proof. By taking the (monotonic) limit in (4.3.2), we get the following integral equation for f_t (where we interchanged $t-s$ and s)

$$f_t(x,y) = \mathbf{1}_{\{x\}}(y)e^{-q_x t} + \int_0^t e^{-q_x(t-s)} \sum_{k\neq x} q_{xk} f_s(k,y)ds, \quad x,y \in E, t \geq 0. \tag{4.3.3}$$

By boundedness of the integrand, it directly follows that $t \mapsto f_t(x,y)$ is continuous. The dominated convergence theorem and the fact that $f_t(x,y) \leq 1$ imply that $t \mapsto \sum_{k\neq x} q_{xk} f_t(k,y)$ is continuous. Hence, we can differentiate the right hand-side with respect to t, and get that $(f_t)_t$ solves the Kolmogorov backward equation.

Instead of considering the time of the first jump before t, we can also consider the time of the last jump. Then we get the following recursive sequence: $F_t^0(x,y) = \mathbf{1}_{\{x\}}(y)e^{-q_x t}$,

$$F_t^{n+1}(x,y) = F_t^0(x,y) + \int_0^t \sum_{k\neq x} F_{t-s}^n(x,k) q_{ky} e^{-q_y s} ds.$$

$(F_t^n)_t$ converges monotonically to a limit $(F_t)_t$, which satisfies the forward integral equation

$$F_t(x,y) = \mathbf{1}_{\{x\}}(y)e^{-q_x t} + \int_0^t \sum_{k\neq x} F_{t-s}(x,z) q_{ky} e^{-q_y s} ds.$$

Clearly $f_t^0 = F_t^0$ for $t \geq 0$. One can inductively prove that $f_t^n = F_t^n$, $t \geq 0$. By differentiation (check that this is possible) we get that $(f_t)_t$ is a solution to the forward equations.

Suppose that $(P_t)_t$ is another solution to the Kolmogorov backward equations. Then $P_t'(x,y) + q_x P_t(x,y) = \sum_{k\neq x} q_{xk} P_t(k,y)$. Multiply both sides with the integrating factor $e^{q_x t}$. Then we find

$$(e^{q_x t} P_t(x,y))' = e^{q_x t} P_t'(x,y) + q_x e^{q_x t} P_t(x,y) = e^{q_x t} \sum_{k\neq x} q_{xk} P_t(k,y).$$

Integrating both sides yields (4.3.2) with f_t replaced by P_t. Now we will iteratively prove that $P_t(x,y) \geq f_t^n(x,y)$ for $n = 0, \ldots$. First, note that $P_t(x,y) \geq \mathbf{1}_{\{x\}}(ye^{-q_x t} = f_t^0(x,y)$. Hence for all $x,y \in E$ and $t \geq 0$

$$\begin{aligned} P_t(x,y) &= \mathbf{1}_{\{x\}}(y)e^{-q_x t} + \int_0^t e^{-q_x(t-s)} \sum_{k\neq x} q_{xk} P_s(k,y)ds \\ &\geq \mathbf{1}_{\{x\}}(y)e^{-q_x t} + \int_0^t e^{-q_x(t-s)} \sum_{k\neq x} q_{xk} f_t^0(k,y)ds = f_t^1(x,y). \end{aligned}$$

Iterating, yields $P_t(x,y) \geq f_t(x,y)$ for $x,y \in E$ and $t \geq 0$.

A similar argument applies if $(P_t)_t$ solves the forward equations. QED

Proof of Lemma 4.3.4. Next we show that $(f_t)_t$ is stochastic. Notice that $\sum_y f_t^n(x,y) = \mathsf{P}_x\{J_{n+1} > t\}$. $\{J_n\}_n$ is a non-decreasing sequence, and so it has a limit, J_∞ say. The event $\{J_{n+1} > t\} = \cup_{k=0}^n \{J_k \leq t\}$. Hence $\{J_\infty > t\} = \lim_{n\to\infty}\{J_{n+1} > t\} = \cup_{n=0}^\infty \{J_k \leq t\}$, this is the event that there are only finitely many jumps in $[0,t]$. By the monotone convergence theorem, taking the limit $n \to \infty$, yields $\sum_u f_t(x,y) = \mathsf{P}_x\{J_\infty > t\}$. Consequently

$$1 - \sum_y f_t(x,y) = \mathsf{P}_x\{J_\infty \leq t\}.$$

Next, by the boundedness of jump rates, there exists a constant c, such that $q_x \leq c$ for all $x \in E$. Therefore

$$J_{n+1} = \frac{T_0}{q_{Y_0}} + \cdots + \frac{T_n}{q_{Y_n}} \geq \frac{1}{c}\sum_{k=0}^n T_k \to \infty, \qquad \text{a.s.}$$

by the Law of Large Numbers, since $\mathsf{E}T_n = 1$.

It follows that $f_t = P_t$ for all $t \geq 0$, and so $(f_t)_t$ is a solution to the Kolmogorov forward equations. QED

This allows to show the validity of (4.2.1) for a larger class of functions.

Lemma 4.3.6 *Suppose that $f : E \to \mathbf{R}$ satisfies $P_t|f| < \infty$ for all $t \geq 0$. Then f satisfies (4.2.1).*

Proof. See Exercise 4.2. QED

The validity of Dynkin's lemma is more involved, since one needs to bound the integrand in a suitable manner.

4.4 Unbounded rates

We next assume that $\sup_x q_x = \infty$.

In this case the proof of Lemma 4.3.4 breaks down, as the following example shows.

Example 4.4.1 Suppose Q is a $\mathbf{Z}_+ \times \mathbf{Z}_+$ matrix with elements

$$q_{x,x+1} = 2^x = -q_{x,x}, x \in \mathbf{Z}_+,$$

all other elements are zero. As in the previous paragraph, we can construct the minimal process X from this. Given that $X_0 = 0$, $J_{n+1} - J_n$ has an $\exp(2^n)$ distribution. Hence $\mathsf{E}J_{n+1} = \sum_{k=0}^n 2^{-k} \leq 2$ for all n. Since J_∞ is the a.s. limit, and $\{J_n\}_n$ is a non-decreasing, non-negative sequence of random variables, the monotone convergence theorem yields $\mathsf{E}J_\infty \leq 2$, and so $J_\infty < \infty$, a.s. This process is said to *explode in finite time.*

Recall that $1 - \sum_y f_t(x,y) = \mathsf{P}_x\{J_\infty \leq t\}$. It is immediate that $\sum_y f_t(x,y) < 1$ for all t sufficiently large. This results in non-unicity of transition functions with a given generator Q.

Constructively, one can add a coffin state to the state space, say Δ. We say that $X_{J_\infty} = \Delta$. From Δ we lead the process immediately back into the original state space (hence $\sigma_\Delta = 0$), and with probability the next state is x with probability p_x, $\sum_x p_x = 1$. The probability distribution $\{p_x\}_x$ can be chosen arbitrarily. However, for every choice, we obtain the same generator (restricted to the state of the space E) and a solution to the Kolmogorov backward equations.

Another problem arises as well. In the proof of Theorem 4.3.3 we did not use the assumption that $\sup_x q_x < \infty$. Hence $(f_t)_t$ is a solution to the Kolmogorov forward equations. However, (4.2.1) does *not* hold in general for bounded functions. Take $f(x) = 1$ for all x and suppose that (4.2.1) holds. Then

$$1 > \sum_y f_t(x,y) = \mathsf{E} f(X_t) = f(x) + \mathsf{E}_x \int_0^t Qf(X_s)ds = 1 + 0,$$

since $Qf = 0$. A contradiction.

This turns out to be generic: explosiveness strongly limits the class of functions for which the Kolmogorov forward integral equations (4.2.1) hold. A class of functions for which it is guaranteed to hold, can be obtained through a criterion described below.

What does hold? It turns out that the result of Theorem 4.3.3 is still valid. Let us properly define explosiveness.

Definition 4.4.1 X is said to be explosive if for $J_\infty = \lim_{n\to\infty} J_n$ it holds that

$$\mathsf{P}_x\{J_\infty < \infty\} > 0, \qquad \text{for some } x \in E.$$

If X is non-explosive, then $\sum_y f_t(x,y) = 1$ for all $x \in E$ and $t \geq 0$, where $(f_t)_t$ is obtained as a limit of $(f_t^n)_t$. This follows by inspection of the proof of Lemma 4.3.4. We summarise this below.

Theorem 4.4.2 *Let X be a stable, conservative, standard Markov process. If X is non-explosive, then $(P_t)_t$ is the unique solution to the Kolmogorov forward and backward equations. In particular, X is the minimal process, and $P_t = f_t$, $t \geq 0$, and Q uniquely defines $(P_t)_t$.*

An important question now become: how can one check whether X is non-explosive? The second question is: for what functions is the Kolmogorov forward integral equation valid?

For the answer to the first question we need to introduce the concept of a moment function.

Definition 4.4.3 $V : E \to \mathbf{R}_+$ is called a moment function, if there exists an increasing sequence $\{K_n\}_n \subset E$ if finite sets with $\lim_{n\to\infty} K_n = E$, such that $\lim_{n\to\infty} \inf\{V(x) \,|\, x \notin K_n\} = \infty$.

The following result holds.

Lemma 4.4.4 *The three following statements are equivalent for a minimal, stable, standard, conservative Markov process.*

i) *X is non-explosive;*

ii) *there exist a moment function V and a constant c such that $QV(x) \leq cV(x)$ for all $x \in E$, i.o.w. $QV \leq cV$;*

iii) *there is no bounded function $f \in b\mathcal{E}$ and $\lambda > 0$, such that $Qf = \lambda f$.*

The sufficiency has been known a long time. The necessity has been recently proved by the lecturer. The following result is also recent.

Theorem 4.4.5 *Under the conditions of Theorem 4.4.2 the Kolmogorov forward integral equation (4.2.1) for any bounded function $f : E \to \mathbf{R}$ for which $\mathsf{E}_x|(Qf)(X_s)|ds < \infty$, $x \in E$, $t \geq 0$. For the latter condition to be satisfied it is sufficient that there exists a function $f : E \to (0, \infty)$ and a constant c such that*

$$\sum_y q_{xy} f(y) + q_x \leq cf(x), \quad x \in E.$$

What can one do for unbounded functions? Also here the condition $QV \leq cV$ plays a crucial role.

4.4.1 V-transformation

Suppose that $QV \leq cV$, for some constant $c \geq 0$, where $V : E \to (0, \infty)$. We extend append a state Δ to E and obtain $E_\Delta = E \cup \{\Delta\}$. Define a new conservative and stable generator Q^V, which is a transformation of Q, as follows

$$q^V_{xy} = \begin{cases} \frac{q_{xy}V(y)}{V(x)} & y \neq x, y, x \in E \\ -c - q_{xx}, & x \in E \\ c - \sum_{z \in E} \frac{q_{xz}V(z)}{V(x)}, & y = \Delta \\ 0, & x = \Delta, y \in E_\Delta. \end{cases}$$

The state Δ has been appended to make Q^V conservative. It has been absorbing, so as not to interfere with the transitions between the states in E. If X is non-explosive, it is a the minimal process. Suppose that the minimal process X^V, obtained by construction from Q^V, is non-explosive as well. Then, by using the fact that we can obtain their respective transition functions by a limiting procedure, it is now difficult to obtain that

$$\frac{P_t(x, y)V(y)}{V(x)} = e^{ct} P^V_t(x, y), \quad x, y \in E, \tag{4.4.1}$$

where now $(P^V_t)_t$ is the transition function of X^V. Non-explosiveness of X^V can be checked directly on Q by virtue of Lemma 4.4.4. The existence of a moment function W for Q^V, is equivalent to the existence of a function $F : E \to (0, \infty)$ and an increasing sequence $\{K_n\}_n \subset E$ of finite sets, with $\lim_{n\to\infty} K_n = E$, such that $\lim_{n\to\infty} \inf\{F(x)/V(x) \,|\, x \notin K_n\} = \infty$. Such a function F will be called a V-moment function.

Theorem 4.4.5 can then applied to the transformed chain X^V. Using (4.4.1), the following result can be deduced.

Theorem 4.4.6 *Assume the conditions of Theorem 4.4.2. Suppose that $QV \leq cV$ for some function $V : E \to (0, \infty)$ and $c > 0$. Suppose that there exists a V-moment function $F : E \to (0, \infty)$. Then (4.2.1) applies to any function f, with $\sup_x |f(x)|/V(x) < \infty$ and $\mathsf{E}_x \int |Qf(X_s)|ds < \infty$, $x \in E$, $t \geq 0$. The latter condition is satisfied, if there exists a function $F : E \to (0, \infty)$ such that*

$$\sum_y q_{xy} F(y) + q_x V(x) \leq cF(x),$$

for some constant c.

4.5 Exercises

Exercise 4.1 Prove Lemma 4.2.2.

Exercise 4.2 Prove Lemma 4.3.6. First consider $f \geq 0$. Then show first $t \mapsto P_t f(x)$ is continuous for each x. Proceed to show that $(P_t f)_t$ satisfies the Kolmogorov backward integral by separating negative and positive terms. Then finish the proof.

Exercise 4.3 Branching model in continuous time Let $E = \mathbf{Z}_+ = \{0, 1, 2, \ldots\}$ be equipped with the discrete topology and let $\mathcal{E} = 2^E$ be the collection of all subsets of E. Let $\lambda, \mu > 0$.

Cells in a certain population either split or die (independently of other cells in the population) after an exponentially distributed time with parameter $\lambda + \mu$. With probability $\lambda/(\lambda + \mu)$ the cell then splits, and with probability $\mu/(\lambda + \mu)$ it dies. Denote by X_t the number of living cells at time t. This is an $(E, \mathcal{E})$-valued stochastic process. Assume that it is a right-continuous, standard Markov process.

i) Show that the generator Q is given by

$$Q(i,j) = \begin{cases} \lambda i & j = i+1 \\ -(\lambda+\mu)i, & j = i \\ \mu i, & j = i-1, i > 0. \end{cases}$$

You may use the result of Lemma 4.3.1.

ii) Suppose $X_0 = 1$ a.s. We would like to compute the generating function

$$G(z,t) = \sum_j z^j \mathsf{P}_1\{X_t = j\}.$$

Show (using the Kolmogorov forward equations, that you may assume to hold) that G satisfies the partial differential equation

$$\frac{\partial G}{\partial t} = (\lambda z - \mu)(z-1)\frac{\partial G}{\partial z},$$

with boundary condition $G(z, 0) = z$. Show that this PDE has solution

$$G(z,t) = \begin{cases} \frac{\lambda t(1-z)+z}{\lambda t(1-z)+1}, & \mu = \lambda \\ \frac{\mu(1-z)e^{-\mu t} - (\mu - \lambda z)e^{-\lambda t}}{\lambda(1-z)e^{-\mu t} - (\mu-\lambda z)e^{-\lambda t}}, & \mu \neq \lambda \end{cases}$$

iii) Compute $\mathsf{E}_1 X_t$ by differentiating G appropriately. Compute $\lim_{t\to\infty} \mathsf{E}_1 X_t$.

iv) Compute the extinction probability $\mathsf{P}_1\{X_t = 0\}$, as well as $\lim_{t\to\infty} \mathsf{P}_1\{X_t = 0\}$ (use G). What conditions on λ and μ ensure that the cell population dies out a.s.?

Chapter 5

Feller-Dynkin processes

The general problem of the existence of a derivative of a transition function is more complicated in the general state space case. It is not clear what is meant by derivative of a probability measure. The usual approach is functional-analytic one, where derivatives of $t \mapsto \mathsf{E}_x f(X_t)$ are considered. The goal is seek for conditions under which functional versions of the Kolmogorov forward and backward equations hold. In particular, under wich conditions does it hold that

$$\tfrac{d}{dt}(P_t f) = P_t(Qf) = Q(P_t f), \quad t \geq 0?$$

If this relation holds, then the integral form (4.2.1) holds and we have seen that this is not generally true, even for the simpler case of a countable state Markov process.

5.1 Semi-groups

As the question is mainly analytic, this section will be analytically oriented. The starting point is a transition function $\{P_t\}_{t\geq 0}$ on the measurable space $(E, \mathcal{E})$, where again E is Polish and $\mathcal{E}$ the Borel-σ-algebra on E. Let $\mathcal{S}$ be a Banach space of real-valued measurable functions on E, and let $\|\cdot\|$ denote the corresponding norm. By virtue of the Chapman-Kolmogorov equations, $\{P_t\}_t$ is a so-called *semigroup*.

Definition 5.1.1 The semigroup $\{P_t\}_t$ is a strongly continuous semigroup on $\mathcal{S}$ (shorthand notation: SCSG($\mathcal{S}$)), if

i) $P_t : \mathcal{S} \to \mathcal{S}$ is a *bounded linear operator* for each $t \geq 0$. I.o.w. $\|P_t\| := \sup_{f\in\mathcal{S}} \frac{\|P_t f\|}{\|f\|} < \infty$ for $t \geq 0$.

ii) $\lim_{t\downarrow 0} \|P_t f - f\| = 0$ for each $f \in \mathcal{S}$.

A main notion that we will use is *closedness*.

Let $B : \mathcal{D} \to \mathcal{S}$ be a linear operator defined on $\mathcal{D} \subset \mathcal{S}$, with $\mathcal{D}$ a linear subspace. $\mathcal{D}$ is called the *domain* of B. The set $\mathcal{G}(B) = \{(f, Bf) \mid f \in \mathcal{D}\} \subset \mathcal{S} \times \mathcal{S}$ is called the graph of B. Note that $\mathcal{S} \times \mathcal{S}$ is a Banach space with norm $\|(f, g)\| = \|f\| + \|g\|$. Then we call B is *closed* iff $\mathcal{G}(B) = \overline{\mathcal{G}(B)}$.

If $\{P_t\}_t$ is a SCSG($\mathcal{S}$), then $f \mapsto P_t f$ is continuous, and so P_t is a closed linear operator for each $t \geq 0$.

Example 5.1.1 Consider Example 3.1.2 (A, B) Brownian motion and the Ornstein-Uhlenbeck process. Let

$$\mathcal{S} = C_0(\mathbf{R}) = \{f : \mathbf{R} \to \mathbf{R} \,|\, f \text{ continuous with } \lim_{x\to\pm\infty} f(x) = 0\},$$

and let $\|f\| = \sup_{x\in\mathbf{R}} |f(x)|$. One can show that the associated transition functions are strongly continuous semigroups on $C_0(\mathbf{R})$ (cf. Section 5.3).

Example 5.1.2 Consider the Markov jump process in Example 3.2.1. Suppose that there exists a function $F : E \to \mathbf{R}_+$ and a constant $c > 0$ such that $PF(x) \leq cF(x)$, for all $x \in E$. Let

$$\mathcal{S} = \{f : E \to \mathbf{R} \,|\, \|f\|_F := \sup_x \frac{|f(x)|}{F(x)} < \infty\}.$$

Then the associated transition function is a strongly continuous semigroup on $\mathcal{S}$, with

$$\|P_t\| \leq \sum_{n=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^n}{n!} c^n = e^{(c-1)\lambda t}.$$

This norm is a *weighted supremum norm* and the Banach space $\mathcal{S}$ is used in much of modern Markov chain theory with applications in queueing and control. The choice $F \equiv 1$ often applies.

Example 5.1.3 Let X be a countable state Markov process, that is minimal, stable, standard and conservative with $P_t'|_{t=0} = Q$. Suppose that there exists a V-moment function and a constant such that $QV \leq cV$. Let

$$C_0(E,V) = \left\{ f : E \to \mathbf{R} \;\middle|\; \begin{array}{l} \text{for each } \epsilon > 0 \text{ there exists a finite set } K = K(\epsilon, f), \\ \text{such that } \frac{|f(x)|}{V(x)} \leq \epsilon, \text{ for } x \notin K \end{array} \right\}$$

$C_0(E,V)$ equipped with the norm $\|\cdot\|_V$ is a Banach space, and $(P_t)_t$ is a strongly continuous semigroup on this space.

The norm of an SCSG($\mathcal{S}$) cannot grow quicker than exponentially. This follows from the following lemma.

Lemma 5.1.2 *Let $\{P_t\}_{t\geq 0}$ be a SCSG($\mathcal{S}$). There are constants $M \geq 1$, $\alpha \geq 0$, such that $\|P_t\| \leq Me^{\alpha t}$.*

Proof. Note first that there exists constants $M \geq 1$ and $t_0 > 0$, such that $\|P_t\| \leq M$, $t \leq t_0$.

Suppose that not. Then there exists a sequence $t_n \downarrow 0$, $n \to \infty$, such that $\|P_{t_n}\| \to \infty$. The Banach-Steinhaus theorem (cf. BN Theorem 10.4) then implies that $\sup_n \|P_{t_n} f\| = \infty$ for some $f \in \mathcal{S}$. This contradicts strong continuity. Hence there exists a constant $M \geq 1$, such that $\|P_t\| \leq M$ for $t \leq t_0$.

Finally, put $\alpha = (\log M)/t_0$. Let $t \in [0,\infty)$. Then with $k = \lfloor t/t_0 \rfloor$, we get

$$\|P_t\| = \|P_{kt_0} P_{t-kt_0}\| \leq \|P_{t_0}\|^k \|P_{t-kt_0}\| \leq e^{\alpha t} \cdot M.$$

QED

Corollary 5.1.3 *$t \mapsto P_t f$ is continuous (i.o.w. $\lim_{s\to t} \|P_t f - P_s f\| = 0$) for all $f \in \mathcal{S}$, and $t \geq 0$).*

Proof. We will only prove right-continuity. Let $h > 0$. Then

$$\|P_{t+h} f - P_t f\| \leq \|P_t\| \|P_h f - f\| \to 0, \quad h \downarrow 0.$$

QED

Generator Let next

$$\mathcal{D} = \{f \in \mathcal{S} \mid \exists g \in \mathcal{S} \text{ such that } \lim_{t \downarrow 0} \| \frac{P_t f - f}{t} - g\| = 0\}.$$

A priori it is not clear whether $\mathcal{D}$ is even non-empty! For each $f \in \mathcal{D}$ we write

$$Af = g = \lim_{t \downarrow 0} \frac{P_t f - f}{t}.$$

$\mathsf{A} : \mathcal{D} \to \mathcal{S}$ is a (generally unbounded) linear operator, with domain $\mathcal{D}(\mathsf{A}) = \mathcal{D}$, A is called the *generator*.

From the definition we immediately see that for $f \in \mathcal{D}(\mathsf{A})$

$$\mathsf{E}_\nu(f(X_{t+h}) - f(X_t) \mid \mathcal{F}_t^X) = h\mathsf{A}f(X_t) + o(h), \quad \mathsf{P}_\nu - \text{a.s.},$$

as $h \downarrow 0$. In this sense the generator describes the motion in an infinitesimal time-interval.

Example 5.1.4 Brownian motion has $\mathcal{D}(A) = \{f \in C_0(\mathbf{R}) \mid f', f'' \in C_0(\mathbf{R})\}$. It holds that $\mathsf{A}f = f''/2$ for $f \in \mathcal{D}(\mathsf{A})$. The proof is given in §5.3. Notice that this implies that $\mathbf{1}_{\{B\}}$, $B \in \mathcal{B}$, are not even contained in the domain, and so these functions do not satisfy the Kolmogorov forward and backward equations!

Example 5.1.5 (Ornstein-Uhlenbeck process) Consider the Orstein-Uhlenbeck process in Example 3.1.2 (B). The generator is given by

$$\mathsf{A}f(x) = \tfrac{1}{2}\sigma^2 f''(x) - \alpha x f'(x), \quad x \in \mathbf{R},$$

if $f \in \{g \in C_0(\mathbf{R}) \mid g', g'' \in C_0(\mathbf{R})\}$(cf. Exercise 5.1).

Recall that we introduced Brownian motion as a model for the position of a particle. The problem however is that Brownian motion paths are nowhere differentiable, whereas the derivative of the position of a particle is its velocity, hence it should be differentiable. It appears that the Ornstein-Uhlenbeck process is a model for the velocity of a particle, and then its position at time t is given by

$$S_t = \int_0^t X_u du.$$

It can be shown that $\alpha S_{nt}/\sqrt{n} \to W_t$ in distribution, as $n \to \infty$. Hence, for large time scales, Brownian motion may be accepted as a model for particle motion.

Example 5.1.6 Consider the Geometric Brownian motion in Example 3.1.2 (C). The generator is given by

$$\mathsf{A}f(x) = \mu x f'(x) + \tfrac{1}{2}\sigma^2 x^2 f''(x),$$

for $f \in \{g \in C_0(\mathbf{R}) \,|\, g', g'' \in C_0(\mathbf{R})\}$.

The full description of the domain $\mathcal{D}(A)$ is very difficult in general. Lateron we provide some tools that might help for its specification.

We next derive the important Kolmogorov forward and backward equations. This requires integrating $\mathcal{S}$-valued functions of t.

Denote by $C_{\mathcal{S}}(a,b) = \{u : [a,b] \to \mathcal{S} \,|\, u \text{ is continuous}\}$, $a, b \in [-\infty, \infty]$. A function $u : [a,b] \to \mathcal{S}$ is said to be (Rieman) integrable over $[a,b]$ if $\lim_{h\to 0} \sum_{k=1}^n u(s_k)(t_k - t_{k-1})$ exists, where $a = t_0 \leq s_1 \leq t_1 \leq \cdots \leq t_{n-1} \leq s_n \leq t_n = b$ and $h = \max_k(t_k - t_{k-1})$, and the limit is independent of the particular sequence $t_0, s_1, \ldots, t_k$. It is then denoted by $\int_a^b u(t)dt$. If a and/or $b = \infty$, the integral is defined as an improper integral.

Furthermore, by $\mathbf{I}$ we mean the identity operator. The following result holds.

Integration Lemma

a) If $u \in C_{\mathcal{S}}(a,b)$ and $\int_a^b \|u(t)\|dt < \infty$, then u is integrable over $[a,b]$ and

$$\| \int_a^b u(t)dt \| \leq \int_a^b \|u(t)\|dt.$$

If a, b are finite then every function in $C_{\mathcal{S}}(a,b)$ is integrable over $[a,b]$.

b) Let B be a closed linear operator on $\mathcal{S}$. Suppose that $u \in C_{\mathcal{S}}(a,b)$, $u(t) \in \mathcal{D}(B)$ for all $t \in [a,b]$, and both u, Bu are integrable over $[a,b]$. Then $\int_a^b u(t)dt \in \mathcal{D}(B)$ and

$$B \int_a^b u(t)dt = \int_a^b Bu(t)dt.$$

c) If $u \in C_{\mathcal{S}}[a,b]$ and u continuously differentiable on $[a,b]$ then

$$\int_a^b \frac{d}{dt}u(t)dt = u(b) - u(a).$$

Proof. See Exercise 5.2. QED

The consequence is that we can interchange of integral and closed linear operators. By Corollary 5.1.3 $s \mapsto P_s f$ is continuous, hence integrable over $[0,t]$.

The following theorem holds.

Theorem 5.1.4 *Let $\{P_t\}_t$ be an SCSG($\mathcal{S}$).*

i) *Let $f \in \mathcal{S}$, $t \geq 0$. Then $\int_0^t P_s f ds \in \mathcal{D}(\mathsf{A})$ and*

$$P_t f - f = \mathsf{A} \int_0^t P_s f ds.$$

ii) *Let $f \in \mathcal{D}(\mathsf{A})$ and $t \geq 0$. Then $P_t f \in \mathcal{D}(\mathsf{A})$. The function $t \mapsto P_t f$ is differentiable in $\mathcal{S}$ and the Kolmogorov backward and forward equations hold:*

$$\frac{d}{dt} P_t f = \mathsf{A} P_t f = P_t \mathsf{A} f.$$

More precisely,

$$\lim_{h \downarrow 0} \left\| \frac{P_{t+h} f - P_t f}{h} - P_t \mathsf{A} f \right\| = \lim_{h \downarrow 0} \left\| \frac{P_{t+h} f - P_t f}{h} - \mathsf{A} P_t f \right\| = 0.$$

iii) *Let $f \in \mathcal{D}(\mathsf{A})$, $t \geq 0$. Then*

$$P_t f - f = \int_0^t P_s \mathsf{A} f ds = \int_0^t \mathsf{A} P_s f ds.$$

Proof. For the proof of (i) note that

$$\begin{aligned} \frac{1}{h}(P_h - \mathbf{I}) \int_0^t P_s f ds &= \frac{1}{h} \int_0^t \Big(P_{s+h} f - P_s f \Big) ds \\ &= \frac{1}{h} \int_H^{t+h} P_s f ds - \frac{1}{h} \int_0^t P_s f ds. \end{aligned}$$

For the second term we get

$$\left\| \frac{1}{h} \int_0^t P_s f ds - f \right\| \leq \frac{1}{h} \int_0^h \| P_s f - f \| ds \to 0, \quad h \downarrow 0.$$

Similarly, for the first term

$$\left\| \frac{1}{h} \int_t^{t+h} P_s f ds - P_t f \right\| \leq \frac{\| P_t \|}{h} \int_0^h \| P_s f - f \| ds \to 0, \quad h \downarrow 0.$$

The result follows. For (ii) note that

$$\left\| \frac{P_{t+h} f - P_t f}{h} - P_t \mathsf{A} f \right\| = \left\| P_t \Big(\frac{P_h f - f}{h} - \mathsf{A} f \Big) \right\| \leq \| P_t \| \left\| \frac{P_h f - f}{h} - \mathsf{A} f \right\|.$$

Taking the limit $h \downarrow 0$ yields that

$$\lim_{h \downarrow 0} \left\| \frac{P_{t+h} f - P_t f}{h} - P_t \mathsf{A} f \right\| = 0. \tag{5.1.1}$$

Since $\mathsf{A} f \in \mathcal{S}$, $g = P_t \mathsf{A} f \in \mathcal{S}$. Rewriting (5.1.1) gives

$$\lim_{h \downarrow 0} \left\| \frac{P_h(P_t f) - P_t f}{h} - g \right\| = 0.$$

Hence $g = \mathsf{A} P_t f$. Consequently $P_t \mathsf{A} f = g = \mathsf{A} P_t f = (d^+/dt) P_t f$ (d^+/dt stands for the right-derivative). To see that the left derivative exists and equals the right-derivative, observe for $h > 0$ that

$$\begin{aligned} \left\| \frac{P_t f - P_{t-h} f}{h} - P_t \mathsf{A} f \right\| &\leq \left\| \frac{P_t f - P_{t-h} f}{h} - P_{t-h} \mathsf{A} f \right\| + \left\| P_{t-h} \mathsf{A} f - P_t \mathsf{A} f \right\| \\ &\leq \| P_{t-h} \| \left\| \frac{P_h f - f}{h} - \mathsf{A} f \right\| + \left\| P_{t-h} \mathsf{A} f - P_t \mathsf{A} f \right\| \to 0, \quad h \downarrow 0, \end{aligned}$$

where we have used strong continuity and the fact that $\mathsf{A}f \in \mathcal{S}$. For (iii) note that $(d/dt)P_t f = P_t\mathsf{A}f$ is a continuous function of t by Corollary 5.1.3. It is therefore integrable, and so

$$P_t f - f = \int_0^t \frac{d}{ds} P_s f ds = \int_0^t \mathsf{A} P_s f ds = \int_0^t P_s \mathsf{A} f ds.$$

QED

The previous theorem (i) shows that $\mathcal{D}(\mathsf{A})$ is non-empty. In fact it is dense in $\mathcal{S}$ and A is a so-called closed operator.

Corollary 5.1.5 *Let $\{P_t\}_t$ be an SCSG($\mathcal{S}$). Then $\overline{\mathcal{D}(\mathsf{A})} = \mathcal{S}$ and A is a closed operator.*

Proof. Theorem 5.1.4 (i) and the fact that $\| \int_0^t P_s f ds/t - f\| \to 0$, $t \downarrow 0$ immediately imply that $\overline{\mathcal{D}(\mathsf{A})} = \mathcal{S}$.

Let $\{f_n\}_n \subset \mathcal{D}(\mathsf{A})$ be any sequence with the property that there exist $f, g \in \mathcal{S}$ such that $f_n \to f$ and $\mathsf{A} f_n \to g$ as $n \to \infty$. We need to show that $g = \mathsf{A} f$.

To this end, note that $P_t f_n - f_n = \int_0^t P_s(\mathsf{A} f_n) ds$, for all $t > 0$, by virtue of Theorem 5.1.4 (iii). Since $\|(P_t f_n - f_n) - (P f - f)\| \to 0$ and $\| \int_0^t P_s \mathsf{A} f_n ds - \int_0^t P_s g ds\| \to 0$ as $n \to \infty$, necessarily $P_t f - f = \int_0^t P_s g ds$, for all $t > 0$. Hence

$$\left\| \frac{P_t f - f}{t} - \frac{\int_0^t P_s g ds}{t} \right\| = 0, \quad \forall t > 0.$$

It follows that

$$\lim_{t \downarrow 0} \left\| \frac{P_t f - f}{t} - g \right\| \le \lim_{t \downarrow 0} \left\| \frac{P_t f - f}{t} - \frac{\int_0^t P_s g ds}{t} \right\| + \lim_{t \downarrow 0} \left\| \frac{\int_0^t P_s g ds}{t} - g \right\| = 0,$$

so that $g = \mathsf{A} f$. QED

5.2 The generator determines the semi-group: the Hille-Yosida theorem

By virtue of Corollary 5.1.3, the map $t \to P_t f$ is continuous for each $f \in \mathcal{S}$. Recall that $\|P_t\| \le M e^{\alpha t}$ for some constants $M \ge 1$ and $\alpha \ge 0$. By the Integration Lemma, for all $\lambda > \alpha$ we may define

$$R_\lambda f(x) = \int_0^\infty e^{-\lambda t} P_t f(x) dt.$$

R_λ is simply the Laplace transform of the semigroup calculated at the 'frequency' λ. The next lemma collects preliminary properties of the operators R_λ. In particular, it states that for all $\lambda > 0$, R_λ is in fact an operator that maps $\mathcal{S}$ into itself. It is called the *resolvent of order* λ.

Lemma 5.2.1 *Let $\{P_t\}_t$ be a SCSG($\boldsymbol{S}$).*

i) $\|R_\lambda\| \le M/(\lambda - \alpha)$.

ii) *The resolvent equation*

$$R_\mu - R_\lambda + (\mu - \lambda) R_\mu R_\lambda = 0$$

holds for all $\lambda, \mu > \alpha$.

Proof. The first part is straightforward. To prove the resolvent equation, note that

$$e^{-\mu t} - e^{-\lambda t} = (\lambda - \mu) e^{-\lambda t} \int_0^t e^{(\lambda - \mu)s} ds.$$

Hence,

$$\begin{aligned} R_\mu f(x) - R_\lambda f(x) &= \int_0^\infty (e^{-\mu t} - e^{-\lambda t}) P_t f(x) dt \\ &= (\lambda - \mu) \int_0^\infty e^{-\lambda t} \Big(\int_0^t e^{(\lambda-\mu)s} P_t f(x) ds \Big) dt \\ &= (\lambda - \mu) \int_0^\infty e^{-\mu s} \Big(\int_s^\infty e^{-\lambda(t-s)} P_t f(x) dt \Big) ds, \end{aligned}$$

by the integration Lemma. A change of variables, the semigroup property of the transition function and another application of Integration Lemma show that the inner integral equals

$$\begin{aligned} \int_0^\infty e^{-\lambda u} P_{s+u} f(x) du &= \int_0^\infty e^{-\lambda u} P_s P_u f(x) du \\ &= \int_0^\infty e^{-\lambda u} \Big(\int_E P_u f(y) P_s(x, dy) \Big) du \\ &= \int_E \Big(\int_0^\infty e^{-\lambda u} P_u f(y) du \Big) P_s(x, dy) \\ &= P_s R_\lambda f(x). \end{aligned}$$

Inserting this in the preceding equation yields the resolvent equation. QED

The following important connection between resolvent and generator is easily derived.

Theorem 5.2.2 *Let* $\{P_t\}_t$ *be a SCSG(*$\boldsymbol{S}$*) with* $\|P_t\| \leq M \cdot e^{\alpha t}$*. For all* $\lambda > \alpha$ *the following hold.*

i) $R_\lambda \mathcal{S} = \mathcal{D}(\mathsf{A})$.

ii) $\lambda \mathbf{I} - \mathsf{A} : \mathcal{D}(\mathsf{A}) \to \mathcal{S}$ *is a 1-1 linear operator with* $(\lambda \mathbf{I} - \mathsf{A}) \mathcal{D}(\mathsf{A}) = \mathcal{S}$.

iii) $(\lambda \mathbf{I} - \mathsf{A})^{-1} : \mathcal{S} \to \mathcal{D}(\mathsf{A})$ *exists as a bounded linear operator. In particular* $(\lambda \mathbf{I} - \mathsf{A})^{-1} = R_\lambda$.

iv) $R_\lambda (\lambda \mathbf{I} - \mathsf{A}) f = f$ *for all* $f \in \mathcal{D}(\mathsf{A})$.

v) $(\lambda \mathbf{I} - \mathsf{A}) R_\lambda g = g$ *for all* $g \in \mathcal{S}$.

Proof. The proof consists of 2 main steps: *Step 1* Proof of (v); and *Step 2* Proof of (iv). As a consequence, (iv) implies (i) and the first part of (ii); (v) implies the second part of (ii). (iii) then follows by combining (iv,v).

Proof of Step 1. Let $g \in \mathcal{S}$. By the Integration Lemma we may write

$$P_h R_\lambda g = \int_0^\infty e^{-\lambda t} P_{t+h} g dt.$$

Hence

$$\begin{aligned}\frac{P_h R_\lambda g - R_\lambda g}{h} &= \frac{1}{h}\Big[\int_0^\infty e^{-\lambda t} P_{t+h} g dt - \int_0^\infty e^{-\lambda t} P_t g dt\Big] \\ &= \frac{e^{\lambda h}-1}{h} R_\lambda g - \frac{1}{h} e^{\lambda h} \int_0^h e^{-\lambda t} P_t g dt.\end{aligned}$$

The right-hand side converges to $\lambda R_\lambda g - g$. It follows that

$$\Big\| \frac{P_h R_\lambda g - R_\lambda g}{h} - (\lambda R_\lambda g - g) \Big\| \to 0, \quad h \downarrow 0.$$

By definition, $R_\lambda g \in \mathcal{D}(\mathsf{A})$ and

$$\mathsf{A} R_\lambda g = \lambda R_\lambda g - g. \tag{5.2.1}$$

The result follows by rewriting.

Proof of Step 2. Let $f \in \mathcal{D}(\mathsf{A})$, then by definition $\mathsf{A} f \in \mathcal{S}$. We have

$$R_\lambda[\mathsf{A} f] = \int_0^\infty e^{-\lambda t} P_t[\mathsf{A} f] dt = \int_0^\infty e^{-\lambda t} \mathsf{A}[P_t f] dt = \mathsf{A} \int_0^\infty e^{-\lambda t} P_t f dt = \mathsf{A} R_\lambda f. \tag{5.2.2}$$

The last equality follows from the Integration Lemma by using that A is closed. The second follows from Theorem 5.1.4 (ii). The rest follows by inserting (5.2.2) into (5.2.1) and rewriting. QED

Due to the importance of resolvents, we will explicitly compute these for two examples.

Example 5.2.1 Consider the BM-process from Example 3.1.2 (A). Its resolvents are given by

$$R_\lambda f(x) = \int_{\mathrm{R}} f(y) r_\lambda(x, y) dy,$$

where $r_\lambda(x, y) = \exp\{-\sqrt{2\lambda}|x - y|\}/\sqrt{2\lambda}$ (see Exercise 5.3).

Example 5.2.2 Let X be the Markov jump process from Example 5.1.2. The resolvent is given by (cf. Exercise 5.4).

$$R_\mu f = \frac{1}{\lambda+\mu} \sum_{n \geq 0} \Big(\frac{\lambda}{\lambda+\mu}\Big)^n P^n f = ((\lambda+\mu)\mathbf{I} - \lambda P)^{-1} f \quad f \in \mathcal{S},$$

for $\mu > (c-1)\lambda$. This is by a direct computation, and the fact that (proved using induction)

$$\int_0^\infty t^n e^{-\xi t} dt = \frac{n!}{\xi^{n+1}}.$$

We have proved one part of the Hille-Yosida theorem specifying the precise relation between generator and semigroup. It will however be convenient to restrict to so-called contraction SCSG's: the transition function $\{P_t\}_t$ is a *strongly continuous contracting semigroup on the Banach space* $\mathcal{S}$ (SCCSG($\mathcal{S}$)) if it is a SCSG($\mathcal{S}$) with $\|P_t\| \leq 1$ for all $t \geq 0$.

This is no restriction. Suppose that $\|P_t\| \leq M \cdot e^{\alpha t}$ for constants $M \geq 1$ and $\alpha \geq 0$. Then $P_t e^{-\alpha t}$ is a SCSG($\mathcal{S}$) with $\|P_t e^{-\alpha t}\| \leq M$. Define a new norm $\|\cdot\|^*$ by

$$\|f\|^* = \sup_{t \geq 0} \|P_t e^{-\alpha t} f\|,$$

then $\|f\| \leq \|f\|^* \leq M\|f\|$. Hence $\|\cdot\|$ and $\|\cdot\|^*$ are equivalent norms and $\mathcal{S}$ is a Banach space with respect to $\|\cdot\|^*$. It easily follows that $\{P_t e^{-\alpha t}\}_t$ is a SCCSG($\mathcal{S}$) with respect to the new norm.

We need the notion of dissipativeness. The linear operator $B : V \to \mathcal{S}$, V a linear subspace of $\mathcal{S}$, is (α, M)-*dissipative* if

$$\|(\lambda \mathbf{I} - B)f\| \geq \frac{\lambda - \alpha}{M}\|f\|, \quad \forall f \in V, \lambda > 0.$$

Theorem 5.2.3 (Hille-Yosida Theorem) *Suppose that there exists a linear operator* $\mathsf{A} : \mathcal{D}(\mathsf{A}) \to \mathcal{S}$, *where* $\mathcal{D}(\mathsf{A})$ *is a linear subspace of* $\mathcal{S}$. *Then* A *is the generator of a SCCSG(*$\mathcal{S}$*) semigroup if and only if the following three properties hold:*

i) $\overline{\mathcal{D}(\mathsf{A})} = \mathcal{S}$;

ii) A *is* $(0,1)$*-dissipative. In other words:*

$$\|\lambda f - \mathsf{A}f\| \geq \lambda\|f\|, \quad \forall f \in \mathcal{D}(\mathsf{A}), \lambda > 0;$$

iii) $(\lambda \mathbf{I} - \mathsf{A})\mathcal{D}(\mathsf{A}) = \mathcal{S}$ *for some* $\lambda > 0$.

Proof of "⇒". The only thing left to prove is (ii) (why?). By Theorem 5.2.2 (i) there exists $g \in \mathcal{S}$, such that $f = R_\lambda g$. By the same theorem (v) $(\lambda \mathbf{I} - \mathsf{A})f = g$. We have

$$\|f\| = \|R_\lambda g\| \leq \frac{1}{\lambda}\|g\| = \frac{1}{\lambda}\|\lambda f - \mathsf{A}f\|.$$

QED

For proving '⇐', we need to derive a number of lemmas. We will formulate these for A being (α, M)-dissipative and return to the $(0,1)$-dissipative case at the moment we really need it.

Lemma 5.2.4 *Let* $\mathsf{A} : \mathcal{D}(\mathsf{A}) \to \mathcal{S}$, *be an* (α, M) *dissipative linear operator, with* $\mathcal{D}(\mathsf{A}) \subset \mathcal{S}$ *a linear subspace, and* $M \geq 1$, $\alpha \geq 0$. *Then* A *is a closed operator if and only if* $(\lambda \mathbf{I} - \mathsf{A})\mathcal{D}(\mathsf{A})$ *is a closed set (in* $\mathcal{S}$*), for some* $\lambda > \alpha$. *Under either condition* $(\lambda \mathbf{I} - \mathsf{A})\mathcal{D}(\mathsf{A})$ *is a closed set (in* $\mathcal{S}$*) for all* $\lambda > \alpha$.

Proof. Assume that A is closed. Let $\lambda > \alpha$. Let $\{f_n\}_n \subset \mathcal{D}(\mathsf{A})$ be a sequence such that $(\lambda\mathbf{I} - \mathsf{A})f_n \to h$, $n \to \infty$, for some $h \in \mathcal{S}$. We have prove that $h \in (\lambda\mathbf{I} - \mathsf{A})\mathcal{D}(\mathsf{A})$.

By (α, M)-dissipativeness of A we have

$$\|(\lambda\mathbf{I} - \mathsf{A})(f_n - f_m)\| \geq \frac{\lambda - \alpha}{M}\|f_n - f_m)\|.$$

Hence $\{f_n\}_n$ is a Cauchy-sequence in $\mathcal{S}$, and so it has a limit $f \in \mathcal{S}$. It holds that

$$\mathsf{A}f_n = -(\lambda\mathbf{I} - \mathsf{A})f_n + \lambda f_n \to -h + \lambda f, \quad n \to \infty.$$

A is closed, hence $f \in \mathcal{D}(\mathsf{A})$, $\mathsf{A}f = -h + \lambda f$. Therefore $h = (\lambda\mathbf{I} - \mathsf{A})f$. The conclusion is that $(\lambda\mathbf{I} - \mathsf{A})\mathcal{D}(\mathsf{A})$ is closed.

Next we assume that $(\lambda\mathbf{I} - \mathsf{A})\mathcal{D}(\mathsf{A})$ is closed in $\mathcal{S}$ for some $\lambda > \alpha$. Let $\{f_n\}_n \subset \mathcal{D}(\mathsf{A})$, with $f_n \to f$, $\mathsf{A}f_n \to g$ for some $f, g \in \mathcal{S}$. Then

$$(\lambda\mathbf{I} - \mathsf{A})f_n \to \lambda f - g, \quad n \to \infty,$$

and so $\lambda f - g \in (\lambda\mathbf{I} - \mathsf{A})\mathcal{D}(\mathsf{A})$. Hence, there exists $h \in \mathcal{D}(\mathsf{A})$, such that $\lambda h - \mathsf{A}h = \lambda f - g$. A is (α, M)-dissipative, so that

$$\|\lambda\mathbf{I} - \mathsf{A})(f_n - h)\| = \|(\lambda(f_n - h) - \mathsf{A}(f_n - h)\| \geq \frac{\lambda - \alpha}{M}\|f_n - h\|.$$

The left-hand side converges to 0, as $n \to \infty$. Hence $f_n \to h$ and so $f = h \in \mathcal{D}(\mathsf{A})$ and $g = \mathsf{A}f$. This shows that A is closed. QED

Lemma 5.2.5 *Let $\mathsf{A} : \mathcal{D}(\mathsf{A}) \to \mathcal{S}$, be an (α, M) dissipative, closed linear operator, with $\mathcal{D}(\mathsf{A}) \subset \mathcal{S}$ a linear subspace, and $M \geq 1$, $\alpha \geq 0$. Let*

$$\lambda(\mathsf{A}) = \left\{ \lambda > \alpha \;\middle|\; \begin{array}{l} (\lambda\mathbf{I} - \mathsf{A}) \textit{ is 1-1}, \\ (\lambda\mathbf{I} - \mathsf{A})\mathcal{D}(\mathsf{A}) = \mathcal{S}, \\ (\lambda\mathbf{I} - \mathsf{A})^{-1} \textit{exists as a bounded linear operator on } \mathcal{S} \end{array} \right\}$$

Then $\lambda(\mathsf{A}) \neq \emptyset \Rightarrow \lambda(\mathsf{A}) = (\alpha, \infty)$.

Proof. It is sufficient to show that $\lambda(\mathsf{A})$ is both open and closed in (α, ∞). First we will show that $\lambda(\mathsf{A})$ is open in (α, ∞). To this end, let $\lambda \in \lambda(\mathsf{A})$. Let

$$B = \{\mu \in (\alpha, \infty) \,|\, |\lambda - \mu| < \|(\lambda\mathbf{I} - \mathsf{A})^{-1}\|^{-1}\}.$$

B is open in (α, ∞). We will show that $B \subset \lambda(\mathsf{A})$.

Let $\mu \in B$. Then

$$C = \sum_{n=0}^{\infty} (\lambda - \mu)^n \Big((\lambda\mathbf{I} - \mathsf{A})^{-1}\Big)^{n+1} : \mathcal{S} \to \mathcal{S}$$

is a bounded linear operator. We claim that $C = (\mu\mathbf{I} - \mathsf{A})^{-1}$. Use $\mu\mathbf{I} - \mathsf{A} = \lambda\mathbf{I} - \mathsf{A} + (\mu - \lambda)\mathbf{I}$ to show that $C(\mu\mathbf{I} - \mathsf{A}) = (\mu\mathbf{I} - \mathsf{A})C = \mathbf{I}$. Then it easily follows that $\mu \in \lambda(\mathsf{A})$.

We next show that $\lambda(\mathsf{A})$ is closed in (α,∞). To this end, let $\{\lambda_n\}_n \subset \lambda(\mathsf{A})$, with $\lambda_n \to \lambda$, $n \to \infty$, for some $\lambda \in \mathbf{R}$, $\lambda > \alpha$. We have to show that $\lambda \in \lambda(\mathsf{A})$.

The proof consists of two steps: Step 1 $(\lambda\mathbf{I}-\mathsf{A})\mathcal{D}(\mathsf{A}) = \mathcal{S}$; Step 2 $(\lambda\mathbf{I}-\mathsf{A})$ is 1-1. Steps 1 and 2 then imply that for each $g \in \mathcal{S}$ there exists precisely one element $f \in \mathcal{D}(\mathsf{A})$ with $g = (\lambda\mathbf{I}-\mathsf{A})f$. This means that the inverse $(\lambda\mathbf{I}-\mathsf{A})^{-1}$ exists. By (α, M)-dissipativeness

$$\|(\lambda\mathbf{I}-\mathsf{A})^{-1}g\| = \|f\| \le \frac{M}{\lambda-\alpha}\|\lambda f - \mathsf{A}f\| = \frac{M}{\lambda-\alpha}\|g\|.$$

Since $g \in \mathcal{S}$ was arbitrary, $\|(\lambda\mathbf{I}-\mathsf{A})^{-1}\| \le M/(\lambda-\alpha) < \infty$. This shows that $\lambda \in \lambda(\mathsf{A})$.

Proof of Step 1. Let $g \in \mathcal{S}$. Put $g_n = (\lambda\mathbf{I}-\mathsf{A})(\lambda_n\mathbf{I}-\mathsf{A})^{-1}g$, for all n. Clearly $g = (\lambda\mathbf{I}-\mathsf{A})(\lambda_n\mathbf{I}-\mathsf{A})^{-1}g$ and so $g_n - g = (\lambda-\lambda_n)(\lambda_n\mathbf{I}-\mathsf{A})^{-1}g$. A is (α, M)-dissipative, and so

$$\|g_n - g\| \le |\lambda_n - \lambda| \|(\lambda_n\mathbf{I}-\mathsf{A})^{-1}g\| \le |\lambda_n-\lambda|\frac{M}{\lambda-\alpha}\|g\| \to 0, \quad n\to\infty.$$

It follows that $\overline{(\lambda\mathbf{I}-\mathsf{A})\mathcal{D}(\mathsf{A})} = \mathcal{S}$. Since $(\lambda\mathbf{I}-\mathsf{A})\mathcal{D}(\mathsf{A})$ is closed by Lemma 5.2.4, $\mathcal{S} = (\lambda\mathbf{I}-\mathsf{A})\mathcal{D}(\mathsf{A})$.

Proof of Step 2. This follows immediately from (α, M)-dissipativeness. QED

Lemma 5.2.6 *Let $\mathsf{A} : \mathcal{D}(\mathsf{A}) \to \mathcal{S}$, be an (α, M) dissipative, closed linear operator, with $\mathcal{D}(\mathsf{A}) \subset \mathcal{S}$ a linear subspace, and $M \ge 1$, $\alpha \ge 0$. Suppose that $\overline{\mathcal{D}(\mathsf{A})} = \mathcal{S}$ and $\lambda(\mathsf{A}) = (\alpha,\infty)$. Define the Yosida-approximation $\mathsf{A}_\lambda = \lambda\mathsf{A}(\lambda\mathbf{I}-\mathsf{A})^{-1}$, $\lambda > \alpha$. Then A_λ, $\lambda > \alpha$, have the following properties.*

a) *A_λ is a bounded linear operator $\mathcal{S} \to \mathcal{S}$ and $e^{t\mathsf{A}_\lambda} = \sum_{n=0}^\infty t^n\mathsf{A}_\lambda^n/n!$ is a SCSG($\mathcal{S}$) with generator A_λ.*

b) *$\mathsf{A}_\lambda\mathsf{A}_\mu = \mathsf{A}_\mu\mathsf{A}_\lambda$.*

c) *$\|\mathsf{A}f - \mathsf{A}_\lambda f\| \to 0$, $\lambda \to \infty$, for all $f \in \mathcal{D}(\mathsf{A})$.*

Proof. For all $\lambda > \alpha$ write $U_\lambda = (\lambda\mathbf{I}-\mathsf{A})^{-1}$. In the proof of Lemma 5.2.5 we have seen that $\|U_\lambda\| \le M/(\lambda-\alpha)$. Further $U_\lambda U_\mu = U_\mu U_\lambda$. This follows by a straightforward computation. Hence

$$(\lambda\mathbf{I}-\mathsf{A})U_\lambda = \mathbf{I}, \quad \text{on } \mathcal{S} \tag{5.2.3}$$
$$U_\lambda(\lambda\mathbf{I}-\mathsf{A}) = \mathbf{I}, \quad \text{on } \mathcal{D}(\mathsf{A}). \tag{5.2.4}$$

We first prove (a). Using the above, we may rewrite $\mathsf{A}_\lambda = \lambda\mathsf{A}U_\lambda$ by

$$\mathsf{A}_\lambda = \lambda^2 U_\lambda - \lambda\mathbf{I}, \quad \text{on } \mathcal{S} \tag{5.2.5}$$
$$= \lambda U_\lambda \mathsf{A}, \quad \text{on } \mathcal{D}(\mathsf{A}). \tag{5.2.6}$$

(5.2.5) implies that A_λ is bounded, with

$$\|e^{t\mathsf{A}_\lambda}\| \le e^{-t\lambda}e^{t\lambda^2\|U_\lambda\|} \le e^{t\lambda^2 M/(\lambda-\alpha)-t\lambda}. \tag{5.2.7}$$

Further, for all $f \in \mathcal{S}$

$$\|e^{t\mathsf{A}_\lambda} f - f\| \leq \sum_{n=1}^{\infty} t^n \|\mathsf{A}_\lambda\|^n \|f\|/n! \to 0, \quad t \downarrow 0.$$

Hence $\{e^{t\mathsf{A}_\lambda}\}_t$ is a SCSG($\mathcal{S}$). In a similar way, one can prove that A_λ is the generator.

The proof of (b) follows by using the expression (5.2.5) for A_λ on $\mathcal{S}$ and the fact that U_λ and U_μ commute. We will finally prove (c). First we show that $\|\lambda U_\lambda f - f\| \to 0$, $\lambda \to \infty$, for all $f \in \mathcal{S}$.

For $f \in \mathcal{D}(\mathsf{A})$, we use (5.2.4) to obtain

$$\|\lambda U_\lambda f - f\| = \|U_\lambda \mathsf{A} f\| \leq \frac{M}{\lambda - \alpha}\|\mathsf{A} f\| \to 0, \quad \lambda \to \infty.$$

Let $f \in \mathcal{S}$ and let $\{f_n\}_n \subset \mathcal{D}(\mathsf{A})$ converge to f. Then for all n

$$\limsup_{\lambda\to\infty} \|\lambda U_\lambda f - f\| \leq \limsup_{\lambda\to\infty} \Big[\|\lambda U_\lambda f_n - f_n\| + \frac{M\lambda}{\lambda-\alpha}\|f_n - f\| + \|f_n - f\|\Big]$$

Let $\lambda \to \infty$. The first term on the right-hand side converges to 0, the second converges to $M\|f_n - f\|$ and the third equals $\|f_n - f\|$. Since the left-hand side is independent of n, we can take the limit $n \to \infty$ and obtain that the left-hand side must equal 0. (c) follows by combining with (5.2.6). QED

Lemma 5.2.7 *Suppose that B, C are bounded linear operators on $\mathcal{S}$ with $\|e^{tB}\|, \|e^{tC}\| \leq 1$, that commute: $BC = CB$. Then*

$$\|e^{tB} f - e^{tC} f\| \leq t\|Bf - Cf\|$$

for every $f \in \mathcal{S}$ and $t \geq 0$.

Proof. Use the identity

$$\begin{aligned} e^{tB} f - e^{tC} f &= \int_0^t \frac{d}{ds}[e^{sB} e^{(t-s)C} f ds = \int_0^t e^{sB}(B - C) e^{(t-s)C} ds \\ &= \int_0^t e^{sB} e^{(t-s)C} (B - C) f ds. \end{aligned}$$

QED

Continuation of the proof of the Hille-Yosida Theorem: '$\Leftarrow$' The idea is the define a strongly continuous contraction semigroup $\{P_t\}_t$ and then show that it has generator A.

Conditions (ii), (iii) and Lemma 5.2.4 imply that A is closed. By inspection of the arguments for proving closedness of $\lambda(\mathsf{A})$ in Lemma 5.2.5, we can deduce that $\lambda(\mathsf{A}) \neq \emptyset$. Lemma 5.2.5 implies $\lambda(\mathsf{A}) = (0, \infty)$.

Using the notation in Lemma 5.2.6, define for each $\lambda > 0$ the SCCSG($\mathcal{S}$) $\{e^{t\mathsf{A}_\lambda}\}_t$. Let us check that it is indeed a SCCSG($\mathcal{S}$).

By virtue of (5.2.7), $\|e^{t\mathsf{A}_\lambda}\| \leq 1$. By virtue of Lemmas 5.2.6(b) and 5.2.7

$$\|e^{t\mathsf{A}_\lambda} f - e^{t\mathsf{A}_\mu} f\| \leq t\|\mathsf{A}_\lambda f - \mathsf{A}_\mu f\|,$$

for all $t \geq 0$, and $f \in \mathcal{S}$. By virtue of Lemma 5.2.6 (c), $\lim_{\lambda\to\infty} e^{t\mathsf{A}_\lambda} f$ exists for all $t \geq 0$, uniformly in $t \in [0,T]$, for each $T > 0$, for all $f \in \mathcal{D}(\mathsf{A})$. Define $P_t f = \lim_{\lambda\to\infty} e^{t\mathsf{A}_\lambda} f$, for all $f \in \mathcal{D}(\mathsf{A})$. By uniform convergence, $t \to P_t f$ is continuous for each $f \in \mathcal{D}(\mathsf{A})$.

The fact that $\overline{\mathcal{D}(\mathsf{A})} = \mathcal{S}$ allows to define $P_t f$ on all of $\mathcal{S}$, for all $t \geq 0$.

Next it holds that

$$P_{t+s} f - P_t P_s f \;=\; P_{t+s} f - e^{(t+s)\mathsf{A}_\lambda} f + e^{t\mathsf{A}_\lambda}(e^{s\mathsf{A}_\lambda} f - P_s f) + (e^{s\mathsf{A}_\lambda} - P_s) P_t f.$$

This allows conclude that the Chapman-Kolmogorov equations apply. We may similarly prove that strong continuity holds and that P_t is a bounded linear operator with norm at most 1, $t \geq 0$. We may then conclude that $\{P_t\}_t$ is a SCCSG($\mathcal{S}$).

Finally we will show that this SCSG has generator A. By Theorem 5.1.4 (iii)

$$e^{t\mathsf{A}_\lambda} f - f = \int_0^t e^{s\mathsf{A}_\lambda} \mathsf{A}_\lambda f ds, \tag{5.2.8}$$

for all $f \in \mathcal{S}, t \geq 0, \lambda > 0$. For all $f \in \mathcal{D}(\mathsf{A})$ and $t \geq 0$

$$e^{s\mathsf{A}_\lambda} \mathsf{A}_\lambda f - P_s f = e^{s\mathsf{A}_\lambda}(\mathsf{A}_\lambda f - \mathsf{A} f) + (e^{s\mathsf{A}_\lambda} - P_s) f.$$

By virtue of Lemma 5.2.6 (iii) this implies that $\|e^{s\mathsf{A}_\lambda}\mathsf{A}_\lambda f - P_s \mathsf{A} f\| \to 0, \quad \lambda \to \infty$, uniformly in $s \in [0,t]$. Combining with (5.2.8) yields $P_t f - f = \int_0^t P_s \mathsf{A} f ds$, for all $f \in \mathcal{D}(\mathsf{A})$ and $t \geq 0$.

Suppose that $\{P_t\}_t$ has generator B, with domain $\mathcal{D}(\mathsf{B})$. The above implies that $\mathcal{D}(\mathsf{B}) \supset \mathcal{D}(\mathsf{A})$ and $\mathsf{B} = \mathsf{A}$ on $\mathcal{D}(\mathsf{A})$. Hence B extends A.

By Theorem 5.2.2 $\lambda\mathbf{I} - \mathsf{B}$ is 1-1 for $\lambda > 0$. Since $\mathcal{S} = (\lambda\mathbf{I} - \mathsf{A})\mathcal{D}(\mathsf{A}) = (\lambda\mathbf{I} - \mathsf{B})\mathcal{D}(\mathsf{A})$, $\mathcal{D}(\mathsf{B}) \setminus \mathcal{D}(\mathsf{A}) = \emptyset$, otherwise we would get a contradiction with the fact that $\lambda\mathbf{I} - \mathsf{B}$ is 1-1. QED

It is generally hard to determine the domain of a generator. The following lemma may be of use. Notice the connection with Lemma 4.4.4.

Lemma 5.2.8 *Let $\{P_t\}_t$ be a SCSG($\mathcal{S}$) with $\|P_t\| \leq Me^{\alpha t}$. Suppose that the linear operator B is an extension of the generator A. In other words, $B : \mathcal{D} \to \mathcal{S}$ is a linear operator with $\mathcal{D} \supset \mathcal{D}(\mathsf{A})$, and $Bf = \mathsf{A}f$ for $f \in \mathcal{D}(\mathsf{A})$. If $\lambda\mathbf{I} - B$ is 1-1 for some value $\lambda > \alpha$, then B equals A, that is, $\mathcal{D} = \mathcal{D}(\mathsf{A})$. A sufficient condition for $\lambda\mathbf{I} - B$ to be 1-1 is that there is no non-trivial $f \in \mathcal{D}$ with $Bf = \lambda f$.*

Proof. Suppose that $f \in \mathcal{D}$. Put $g = \lambda f - Bf$. Then $h = R_\lambda g \in \mathcal{D}(\mathsf{A})$ and so

$$\lambda f - Bf = g = \lambda h - \mathsf{A}h = \lambda h - Bh, \tag{5.2.9}$$

since $\mathsf{A} = B$ on $\mathcal{D}(\mathsf{A})$. Hence $f = h \in \mathcal{D}(\mathsf{A})$, if $\lambda\mathbf{I} - B$ is 1-1.

Suppose that we only know that there does not exist any function non-zero $F \in \mathcal{S}$ with $BF = \lambda F$. Then substracting the right side in (5.2.9) from the left, we obtain that $f = h$, and thus $\lambda\mathbf{I} - B$ is 1-1. QED

We may finally ask ourselves whether the transition function is *uniquely* determined by the generator. The answer is again yes in the case of a SCSG($\mathcal{S}$). In general this need not be true.

In the case of special Banach spaces, more interesting properties prevail.

5.3 Feller-Dynkin transition functions

From now on consider semigroups with additional properties. For simplicity, the state space E is assumed to be a closed or open subset of $\mathbf{R}^d$ with $\mathcal{E}$ its Borel-σ-algebra, or of $\mathbf{Z}^d$ with the discrete topology and with the $\mathcal{E}$ the σ-algebra generated by the one-point sets.

By $C_0(E)$ we denote the space of real-valued functions that vanish at infinity. $C_0(E)$ functions are bounded, and so we can endow the space with the supremum norm defined by

$$\|f\| = \sup_{x\in E} |f(x)|.$$

In these notes, we can formally describe $C_0(E)$ by

$$C_0(E) = \left\{ f: E \to \mathbf{R} \;\middle|\; \begin{array}{l} f \text{ continuous and} \\ \text{for each } \epsilon > 0 \text{ there exists a compact set } K = K(\epsilon, f), \\ \text{such that } |f(x)| \le \epsilon, \text{ for } x \notin K \end{array} \right\}$$

Note that $C_0(E)$ is a subset of the space of $b\mathcal{E}$ of bounded, measurable functions on E, so we can consider the restriction of the transition operators $(P_t)_t$ to $C_0(E)$. It is also a Banach space.

We will now introduce a seemingly weaker condition on the semigroup than strong continuity. This notion does not have a unique name in the literature: sometimes it is called the Feller property.

Definition 5.3.1 The transition function $(P_t)_{t\ge 0}$ is called a *Feller-Dynkin transition function* if

i) $P_t C_0(E) \subseteq C_0(E)$, for all $t \ge 0$;

ii) $P_t f(x) \to f(x)$, $t \downarrow 0$, for every $f \in C_0(E)$ and $x \in E$.

A Markov process with Feller-Dynkin transition function is called a Feller-Dynkin process.

Note that the operators P_t are *contractions* on $C_0(E)$, i.e. for every $f \in C_0(E)$ we have

$$\|P_t f\| = \sup_{x\in E} \left| \int_E f(y) P_t(x, dy) \right| \le \|f\|.$$

So, for all $t \ge 0$ we have $\|P_t\|_\infty \le 1$, where $\|P_t\|_\infty$ is the norm of P_t as a linear operator on the normed linear space $C_0(E)$, endowed with the supremum norm (see Appendix B LN, or BN section 11).

If $f \in C_0(E)$, then $P_t f \in C_0(E)$ by part (i) of Definition 5.3.1. By the semigroup property and part (ii) it follows that

$$P_{t+h} f(x) = P_h(P_t f)(x) \to P_t f(x), \quad h \downarrow 0.$$

In other words, the map $t \mapsto P_t f(x)$ is right-continuous for all $f \in C_0(E)$ and $x \in E$.

Right-continuous functions $f: \mathbf{R} \to \mathbf{R}$ are $\mathcal{B}(\mathbf{R})/\mathcal{B}(\mathbf{R})$-measurable. See BN§3.

In particular, this map is measurable, and so for all $\lambda > 0$ we may define the resolvent R_λ. Also in this weaker case, it is a bounded linear operator with norm $\|R_\lambda\| \leq 1$. We will now show that in the case of the space $C_0(E)$, pointwise continuity implies strong continuity.

Theorem 5.3.2 *Suppose that $\{P_t\}_t$ is a Feller-Dynkin transition function. Then $\overline{R_\lambda C_0(E)} = C_0(E)$, and $\{P_t\}_t$ is a SCCSG($C_0(E)$).*

Proof. In order that a Feller-Dynkin transition function be a SCCSG($C_0(E)$), we only need to show strong continuity. We will first show that this is easily checked, provided that $R_\lambda C_0(E)$ is dense in $C_0(E)$.

Since $P_t R_\lambda f(x) = e^{\lambda t} \int_{t=0}^{\infty} e^{-\lambda s} P_s f(x) ds$ by the Integration Lemma,

$$P_t R_\lambda f(x) - R_\lambda f(x) = (e^{\lambda t} - 1) \int_t^\infty e^{-\lambda s} P_s f(x) ds - \int_0^t e^{-\lambda s} P_s f(x) ds.$$

Therefore

$$\|P_t R_\lambda f - R_\lambda f\|_\infty \leq (e^{\lambda t} - 1) \|R_\lambda |f|\|_\infty + t\|f\|_\infty.$$

Since the right-hand side tends to 0 as $t \downarrow 0$, this shows desired norm continuity for functions in the dense subset $R_\lambda C_0(E)$ of $C_0(E)$. Now let $f \in C_0(E)$ be arbitrary. Then for every $g \in R_\lambda C_0(E)$ it holds that

$$\begin{aligned} \|P_t f - f\|_\infty &\leq \|P_t f - P_t g\|_\infty + \|P_t g - g\|_\infty + \|g - f\|_\infty \\ &\leq \|P_t g - g\|_\infty + 2\|g - f\|_\infty. \end{aligned}$$

Taking the $\limsup_{t\downarrow 0}$ and using the first part of the proof, we get

$$\limsup_{t\downarrow 0} \|P_t f - f\|_\infty \leq 2\|g - f\|_\infty,$$

for every $g \in R_\lambda C_0(E)$. The right-hand side can be made arbitrarily small, since $R_\lambda C_0(E)$ is dense in $C_0(E)$. Hence $\lim_{t\downarrow 0} \|P_t f - f\|_\infty = 0$.

Next we will show that $R_\lambda C_0(E)$ is dense in $C_0(E)$. Suppose that this is not true and that $\overline{R_\lambda C_0(E)} \subsetneq C_0(E)$. By the Hahn-Banach theorem (BN §9, Corollary 11.2) there exists a non-trivial bounded linear functional B on $C_0(E)$ that vanishes on $R_\lambda(C_0(E))$. By the Riesz representation theorem (BN §9 Theorem 11.3) there exist finite Borel measures ν and ν' on E such that

$$B(f) = \int_E f d\nu - \int_E f d\nu' = \int_E f d(\nu - \nu'),$$

for every $f \in C_0(E)$. By part (ii) of Definition 5.3.1 and dominated convergence, for every $x \in E$

$$\lambda R_\lambda f(x) = \int_0^\infty \lambda e^{-\lambda t} P_t f(x) dt = \int_0^\infty e^{-s} P_{s/\lambda} f(x) ds \to f(x), \quad \lambda \to \infty. \tag{5.3.1}$$

Note that $\|\lambda R_\lambda f\|_\infty \leq \|f\|_\infty$. Then dominated convergence implies

$$0 = B(\lambda R_\lambda f) = \int_E \lambda R_\lambda f(x)(\nu - \nu')(dx) \to \int_E f(x) d(\nu - \nu')(dx) = B(f), \quad \lambda \to \infty.$$

We conclude that the functional B vanishes on the entire space $C_0(E)$ and so B is trivial. A contradiction. QED

Example 5.3.1 Brownian motion, the Ornstein-Uhlenbeck process en Geometric Brownian motion from Example 3.1.2 (A, B, C) are Feller-Dynkin processes. Hence their semigroups are SCCSG($C_0(\mathbf{R})$) (cf. Example 5.1.1).

We will show that the BM transition function has the Feller-Dynkin property. First we check continuity of $x \mapsto P_t f(x)$ for each $t \geq 0$ and $f \in C_0(\mathbf{R})$. To this end let $x \in \mathbf{R}$ and let $\{x_n\}_n \subset \mathbf{R}$ be a converging sequence with limit x and let $f \in C_0(\mathbf{R})$. Then

$$P_t f(x_n) = \int_{\mathrm{R}} f(x_n - u) \frac{1}{\sqrt{2\pi t}} e^{-u^2/2t} du.$$

Define the functions f_n by $f_n(u) = f(x_n - u)$, $n = 1, \ldots$. We have $f_n \in C_0(\mathbf{R})$, and $\sup_n \|f_n\| = \|f\| < \infty$. Note that $f_n(u) \to f(x-u)$, $n \to \infty$, by continuity of f. By dominated convergence it follows that

$$\lim_{n\to\infty} P_t f(x_n) = \lim_{n\to\infty} \int_{\mathrm{R}} f_n(u) \frac{1}{\sqrt{2\pi t}} e^{-u^2/2t} du = \int_{\mathrm{R}} f(x-u) \frac{1}{\sqrt{2\pi t}} e^{-u^2/2t} du = P_t f(x).$$

The proof that $P_t f(x) \to 0$ as $|x| \to \infty$, is proved similarly.

We will prove pointwise continuity of the function $t \mapsto P_t f(x)$ as $t \downarrow 0$, $x \in \mathbf{R}$. Note that this amounts to proving that $\mathsf{E}_x f(X_t) \to \mathsf{E}_x f(X_0)$ for each $f \in C_0(\mathbf{R})$. First, by sample path continuity $X_t \to X_0$, $t \downarrow 0$, P_x-a.s. Hence by BN Lemma 5.4 $X_t \stackrel{\mathcal{D}}{\to} X_0$, $t \downarrow 0$. The BN Portmanteau theorem 5.3 implies desired convergence.

Example 5.3.2 New result Consider a minimal, standard, stable and right-continuous Markov process with values in the countable state space E equipped with $\mathcal{E} = 2^E$. Suppose there exist a moment function $V : E \to (0, \infty)$ and a constant $\alpha > 0$ such that $QV \leq \alpha V$, where $Q = P_t'|_{t=0}$. Then the V-transformation (see section 4.4.1) is a Feller-Dynkin process.

It follows that X is a Feller-Dynkin process with respect to the space

$$C_0(E, V) = \left\{ f : E \to \mathbf{R} \;\middle|\; \begin{array}{l} \text{for each } \epsilon > 0 \text{ there exists a finite set } K = K(\epsilon, f), \\ \text{such that } \frac{|f(x)|}{V(x)} \leq \epsilon, \text{ for } x \notin K \end{array} \right\}$$

(cf. Exercise 5.5).

5.3.1 Computation of the generator

Let us first look compute the generator of Brownian motion in a straightforward manner.

Example 5.3.3 (cf. Example 5.1.4). We claim that for the BM process we have $\mathcal{D}(\mathsf{A}) = \{f \in C^2(\mathbf{R}) \mid f, f'' \in C_0(\mathbf{R})\}$. One can show that $f, f'' \in C_0(\mathbf{R})$ implies that $f' \in C_0(\mathbf{R})$. Furthermore, for $f \in \mathcal{D}(\mathsf{A})$ we have $\mathsf{A}f = f''/2$.

The procedure to prove this claim, is by showing for $h \in \mathcal{D}(\mathsf{A})$ that

A1 $h \in \mathsf{C}_0^2(\mathbf{R})$; and

A2 $\lambda h - \frac{1}{2}h'' = \lambda h - \mathsf{A}h$.

It then follows that $\mathcal{D}(\mathsf{A}) \subseteq \{f \in C^2(\mathbf{R}) \,|\, f, f'' \in C_0(\mathbf{R})\}$. Hence A' defined by $\mathsf{A}'h = \frac{1}{2}h''$, $h \in \{f \in C^2(\mathbf{R}) \,|\, f, f'' \in C_0(\mathbf{R})\}$ is an extension of A.

We will use Lemma 5.2.8 to show that $\mathsf{A}' = \mathsf{A}$. The lemma implies that this is true if, given $\lambda > 0$, say $\lambda = 1$, there exists no function $h \in \mathcal{D}(\mathsf{A}')$ with $h''/2 = h$.

Let us assume that such a function exists. Then there exists $x \in \mathbf{R}$ with $h(x) \geq h(y)$ for all $y \in \mathbf{R}$. This implies that $h'(x) = 0$ and $h''(x) \leq 0$ by a second order Taylor expansion. Hence $h(x) = h''(x)/2 \leq 0$. Consequently $0 \geq h(y) = h''(y)/2$ for all $y \in \mathbf{R}$. It follows that $h'(y) \leq 0$ for $y \geq x$.

By the assumption that h is non-trivial, there must be some $x' \geq x$ with $h(x') < 0$. Since $h \in C_0(\mathbf{R})$, $\lim_{y\to\infty} h(y) = 0$. Hence, there exists $y > x$ with $h'(y) > 0$. Contradiction.

We are left to show that **A1** and **A2** hold. Let $h \in \mathcal{D}(\mathsf{A})$. Then there exists $f \in \mathsf{C}_0(\mathbf{R})$ such that $h = \mathsf{R}_\lambda f$, i.e.

$$\begin{aligned} h(x) = R_\lambda f(x) &= \int_{\mathrm{R}} f(y) r_\lambda(x,y)dy \\ &= \int_{\mathrm{R}} f(y)\frac{1}{\sqrt{2\lambda}}e^{-\sqrt{2\lambda}|y-x|}dy, \end{aligned}$$

which integral is bounded and differentiable to x. Hence,

$$h'(x) = \int_{\mathrm{R}} f(y)\sqrt{2\lambda}\, r_\lambda(x,y)\, \mathrm{sgn}(y-x)dy.$$

The integrand is not continuous in $y = x$! We have to show that h' is differentiable. For $\delta > 0$ we have

$$\begin{aligned} h'(x+\delta) - h'(x) &= \int_{y<x} f(y)\sqrt{2\lambda}\Big(\frac{r_\lambda(x,y) - r_\lambda(x+\delta,y)}{\delta}\Big)dy \\ &\quad + \int_{y>x+\delta} f(y)\sqrt{2\lambda}\Big(\frac{r_\lambda(x+\delta,y) - r_\lambda(x,y)}{\delta}\Big)dy \\ &\quad - \int_{x\leq y\leq x+\delta} f(y)\sqrt{2\lambda}\Big(\frac{r_\lambda(x+\delta,y) + r_\lambda(x,y)}{\delta}\Big)dy \end{aligned}$$

Clearly,

$$\int_{y<x} \cdots + \int_{y>x+\delta} \cdots \to 2\lambda \int_{\mathrm{R}} f(y) r_\lambda(x,y)dy = 2\lambda h(x), \quad \delta \to 0.$$

Further,

$$\begin{aligned} \int_x^{x+\delta} \cdots &= \int_{u=0}^{\delta} \Big(\frac{\exp^{\sqrt{2\lambda}(u-\delta)} + \exp^{-\sqrt{2\lambda}u}}{\delta}\Big) f(x+u)du \\ &= \int_{u=0}^{1} \mathbf{1}_{\{u\leq\delta\}}\Big(\frac{\exp^{\sqrt{2\lambda}(u-\delta)} + \exp^{-\sqrt{2\lambda}u}}{\delta}\Big) f(x+u)du \\ &\to 2f(x), \quad \delta \downarrow 0, \end{aligned}$$

by dominated convergence. Combining yields

$$\frac{h'(x+\delta) - h'(x)}{\delta} \to 2\lambda h(x) - 2f(x), \quad \delta \downarrow 0.$$

The same holds for $\delta \uparrow 0$, and so we find that h is twice differentiable with

$$h'' = 2\lambda h - 2f.$$

Hence,

$$\lambda h - \tfrac{1}{2}h'' = f = (\lambda \mathbf{I} - \mathsf{A})h = \lambda h - \mathsf{A}h,$$

and so **A1,2** hold.

The first observation is that this is a nasty computation that is hardly amenable to use for more complicated situations. The second is that we derive the domain directly from Lemma 5.2.8. What we implicitly use is the so-called *maximum principle* that we will not further discuss. We believe that the essential feature validating an extension B to equal the generator A, is the fact that $\lambda \mathbf{I} - B$ is 1-1.

Before further investigating ways to compute the generator, we first generalise Lemma 3.1.4, showing conditions under which a function ϕ of a Feller-Dynkin process is Feller-Dynkin, and it provides a relation between the corresponding generators.

Lemma 5.3.3 *Let X be a Feller-Dynkin process with state space $(E, \mathcal{E})$, initial distribution ν and transition function $(P_t)_t$ defined on an underlying filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathsf{P}_\nu)$. Suppose that $(E', \mathcal{E}')$ is a measurable space. Let $\phi : E \to E'$ be continuous and onto, and such that $\|\phi(x_n)\| \to \infty$ if and only if $\|x_n\| \to \infty$.*

Suppose that $(Q_t)_t$ is a collection of transition kernels, such that $P_t(f \circ \phi) = (Q_t f) \circ \phi$ for all $f \in b\mathcal{E}'$. Then $Y = \phi(X)$ is a Feller-Dynkin process with state space $(E', \mathcal{E}')$, initial measure ν', with $\nu'(B') = \nu(\phi^{-1}(B'))$, $B' \in \mathcal{E}'$, and transition function $(Q_t)_t$. The generator B of Y satisfies $\mathcal{D}(\mathsf{B}) = \{f \in C_0(E') \mid f \circ \phi \in \mathcal{D}(\mathsf{A})\}$ and $\mathsf{A}(f \circ \phi) = (\mathsf{B}f) \circ \phi$ for $f \in \mathcal{D}(\mathsf{B})$.

Example 5.3.4 In Example 3.1.6 we have seen that W_t^2 is a Markov process. W_t^2 is also a Feller-Dynkin process with generator $\mathsf{B}f(x) = 2xf''(x) + f'(x)$, $f \in \mathcal{D}(\mathsf{B}) = C_0^2(\mathbf{R}_+)$. See Exercise 5.6.

5.3.2 Applications of the generator and alternative computation

Generators provide an important link between Feller-Dynkin processes and martingales.

Theorem 5.3.4 *Let X be a Feller-Dynkin process, defined on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_t)$. For every $f \in \mathcal{D}(\mathsf{A})$ and initial probability measure ν, the process*

$$M_t^f = f(X_t) - f(X_0) - \int_0^t \mathsf{A}f(X_s)ds,$$

is a P_ν-martingale.

The proof is straightforward, and similarly to the countable state space case, implies the validity of Dynkin's formula.

Corollary 5.3.5 (Dynkin's formula) *Let X be a right-continuous Feller-Dynkin process. For every $f \in \mathcal{D}(\mathsf{A})$ and every $\{\mathcal{F}_t^X\}_t$-stopping time τ with $\mathsf{E}_x\tau < \infty$, we have*

$$\mathsf{E}_x f(X_\tau) = f(x) + \mathsf{E}_x \int_0^\tau \mathsf{A}f(X_s)ds, \quad x \in E.$$

Proof. By Theorem 5.3.4 and the optional sampling theorem, we have

$$\mathsf{E}_x f(X_{\tau\wedge n}) = f(x) + \mathsf{E}_x \int_0^{\tau\wedge n} \mathsf{A}f(X_s)ds,$$

for every $n \in \mathbf{Z}_+$. The left-handside converges to $\mathsf{E}_x f(X_\tau)$, $n \to \infty$ (why, since we do not assume left limits?). Since $\mathsf{A}f \in C_0(E)$, we have $\|\mathsf{A}f\|_\infty < \infty$ and so

$$\Big| \int_0^{\tau\wedge n} \mathsf{A}f(X_s)ds \Big| \leq \|\mathsf{A}f\|\tau.$$

By the fact that $\mathsf{E}_x\tau < \infty$ and by dominated convergence, the integral on the right-handside converges to

$$\mathsf{E}_x \int_0^{\tau} \mathsf{A}f(X_s)ds.$$

QED

This lemma is particularly useful.

Example 5.3.5 Consider the (canonical continuous) BM process $X_t = X_0 + W_t$, $t \geq 0$, where X_0 and $(W_t)_t$ are independent, and $(W_t)_t$ a standard BM.

Let $(a, b) \subset \mathbf{R}$, $a < b$. The problem is to determine a function $f : [a, b] \to \mathbf{R}$, $f \in C_0^2[a, b]$, with $f''(x) = 0$, $x \in (a, b)$ and $f(a) = c_1$, $f(b) = c_2$ for given constants c_1, c_2. Clearly, in dimension 1 this is a simple problem - f is a linear function. However, we would like to use it as an illustration of our theory.

Suppose such a function f exists. Then f can be extended as a $C_0^2(\mathbf{R})$ function. Then $f \in \mathcal{D}(\mathsf{A})$ for our process X. Let $\nu = \delta_x$ be the initial distribution of X for some $x \in (a, b)$. We have seen that $\tau_{a,b} = \inf\{t > 0 \,|\, X_t \in \{a, b\}\}$ is a finite stopping time with finite expectation. Consequently, Dynkin's formula applies and so

$$\mathsf{E}_x f(X_{\tau_{a,b}}) = f(x) + \mathsf{E}_x \int_0^{\tau_{a,b}} \mathsf{A}f(X_s)ds = f(x) + \mathsf{E}_x \int_0^{\tau_{a,b}} \tfrac{1}{2} f''(X_s)ds = f(x).$$

The left-handside equals

$$c_1 \frac{b - x}{b - a} + c_2 \frac{x - a}{b - a}$$

(cf. Exercise 2.29).

Characteristic operator We will now give a probabilistic interpretation of the generator.

Call a point $x \in E$ *absorbing* if for all $t \geq 0$ it holds that $P_t(x, \{x\}) = 1$. This means that if the process starts at an absorbing point x, it never leaves x (cf. Exercise 3.12).

Let X be a right-continuous, Feller-Dynkin process For $r > 0$, define the $\{\mathcal{F}_t\}_t$-stopping time

$$\eta_r = \inf\{t \geq 0 \,|\, \|X_t - X_0\| \geq r\}. \tag{5.3.2}$$

If x is absorbing, then P_x-a.s. we have $\eta_r = \infty$ for all $r > 0$. For non-absorbing points however, the *escape time* η_r is a.s. finite and has finite mean provided r is small enough.

Lemma 5.3.6 *If $x \in E$ is not absorbing, then $\mathsf{E}_x\eta_r < \infty$ for all $r > 0$ sufficiently small.*

Proof. Let $B_x(\epsilon) = \{y \mid \|y - x\| \le \epsilon\}$ be the closed ball of radius ϵ around the point x. If x is not absorbing, then $P_t(x, B_x(\epsilon)) < p < 1$ for some $t, \epsilon > 0$.

By the Feller-Dynkin property of the semi-group P_t we have that $P_t(y,\cdot) \xrightarrow{w} P_t(x,\cdot)$ as $y \to x$. Hence, the Portmanteau theorem, and the fact that $B_x(\epsilon)$ is closed imply that

$$\limsup_{y\to x} P_t(y, B_x(\epsilon)) \le P_t(x, B_x(\epsilon)).$$

Let $\hat p \in (p, 1)$. It follows that for all y sufficiently close to x, say $y \in B_x(r)$ for some $r \in (0, \epsilon)$, we have $P_t(y, B_x(r)) \le \hat p$. Using the Markov property it is easy to show (cf. Exercise 5.7) that $\mathsf{P}_x(\eta_r > nt\} \le \hat p^n$, $n = 0, 1, \ldots$. Hence,

$$\mathsf{E}_x\eta_r = \int_0^\infty \mathsf{P}_x\{\eta_r \ge s\}ds \le t\sum_{n=0}^\infty \mathsf{P}_x(\eta_r \ge nt\} \le \frac{t}{1-\hat p} < \infty.$$

This completes the proof. QED

We can now prove the following alternative description of the generator.

Theorem 5.3.7 *Let X be a right-continuous Feller-Dynkin process. For $f \in \mathcal{D}(\mathsf{A})$ we have $\mathsf{A}f(x) = 0$ if x is absorbing, and otherwise*

$$\mathsf{A}f(x) = \lim_{r\downarrow 0} \frac{\mathsf{E}_x f(X_{\eta_r}) - f(x)}{\mathsf{E}_x\eta_r}, \tag{5.3.3}$$

pointwise!

Proof. If x is absorbing, we have $P_t f(x) = f(x)$ for all $t \ge 0$ and so $\mathsf{A}f(x) = 0$. For non-absorbing $x \in E$ the stopping time η_r has finite mean for sufficiently small r. Dynkin's formula imples

$$\mathsf{E}_x f(X_{\eta_r}) = f(x) + \mathsf{E}_x \int_0^{\eta_r} \mathsf{A}f(X_s)ds.$$

It follows that

$$\begin{aligned}\Big|\frac{\mathsf{E}_x f(X_{\eta_r}) - f(x)}{\mathsf{E}_x\eta_r} - \mathsf{A}f(x)\Big| &\le \frac{\mathsf{E}_x\int_0^{\eta_r}|\mathsf{A}f(X_s) - \mathsf{A}f(x)|ds}{\mathsf{E}_x\eta_r}\\ &\le \sup_{\|y-x\|\le r}|\mathsf{A}f(y) - \mathsf{A}f(x)|.\end{aligned}$$

This completes the proof, since $\mathsf{A}f \in C_0(E)$. QED

The operator defined by the right-handside of (5.3.3) is called the *characteristic operator* of the Markov process X. Its domain is simply the collection of all functions $f \in C_0(E)$ for which the limit in (5.3.3) exists as a $C_0(E)$-function. The theorem states that for right-continuous, canonical Feller-Dynkin processes the characteristic operator extends the infinitesimal generator. We will check the conditions of Lemma 5.2.8 to show that the characteristic operator is the generator. To this end, denote the characteristic operator by B. Suppose that generator and characteristic operator are not equal. Then there exists $f \in C_0(E)$, with $\lambda f = Bf$. We may assume $\lambda = 1$. Then this implies that

$$f(x) = \lim_{r\downarrow 0} \frac{\mathsf{E}_x f(X_{\eta_r}) - f(x)}{\mathsf{E}_x\eta_r}.$$

Suppose that f has a maximum at x: $f(x) \geq f(y)$ for all $y \in E$. Then the above implies that $f(x) \leq 0$ and hence $f(y) \leq 0$ for all $y \in E$. On the other hand, $g = -f$ satisfies $Bg = g$. Let x' be a point where g is maximum. Then, similarly, $g(x') \leq 0$ and so $g(y) \leq 0$ for all $y \in E$. Consequently $f = -f \equiv 0$. Hence, *it is sufficient to computer the characteristic operator and its domain, for the computation of the generator and its domain.*

5.4 Killed Feller-Dynkin processes

NOT FINISHED YET

In this section we consider a Feller-Dynkin cadlag process X with values in $(E, \mathcal{E})$, where E is an open or closed subset of $\mathbf{R}^d$ or $\mathbf{Z}^d$ and $\mathcal{E}$ the Borel-σ-algebra of E.

Upto this point we have always assumed that the transition function $(P_t)_t$ satisfies $P_t(x, E) = 1$, i.e. $P_t(x, \cdot)$ are probability measures for all $x \in E$ and $t \geq 0$. It is sometimes useful to consider transitions functions for which $P_t(x, E) < 1$, for some $x \in E$ and $t \geq 0$. We have seen this in the context of explosive countable state Markov processes. We call the transition function *sub-stochastic* if this is the case.

Intuitively, a sub-stochastic transition function describes the motion of a particle that can disappear from the state space E, or die, in finite time. A sub-stochastic transition function can be turned into a stochastic one (as has already been mentioned) by adjoining a new point Δ to E, called the *coffin state*. Put $E_\Delta = E \cup \{\Delta\}$. Extend the topology of E to E_Δ in such a way that E_Δ is the one-point compactification of E, if E is not compact, and Δ is an isolated point otherwise. Then put $\mathcal{E} = \sigma(\mathcal{E}, \{\Delta\})$. Define a new transition function $(P_t^\Delta)_t$ by putting

$$P_t^\Delta(x, A) = \begin{cases} P_t(x, A), & A \in \mathcal{E}, x \in E \\ 1 - P_t(x, E), & A = \{\Delta\}, x \in E \\ 1, & x = \Delta, A = \{\Delta\}. \end{cases}$$

By construction the point Δ is absorbing for the new process.

By convention, all functions on E are extended to E_Δ by putting $f(\Delta) = 0$. This is consistent in the one-point compactification case. A function $f \in C_0(E_\Delta)$ therefore satisfies $f(\Delta) = 0$. By Corollary 3.2.2 for each probability measure ν on $(E_\delta, \mathcal{E}_\delta)$ there exists a probability measure P_ν on the canonical space $(\Omega, \mathcal{F}) = (E_\delta^{\mathrm{R}+}, \mathcal{E}_\delta^{\mathrm{R}+})$, such that under P_ν the canonical process X is a Markov process with respect to the natural filtration $(\mathcal{F}_t^X)$, with transition function $(P_t^\delta)_t$ and initial distribution ν. Then the process on the extended space is still a cadlag Feller-Dynkin process.

In the sequel we will not distinguish between P_t and P^Δ in our notation, and denote by X the extended process.

We now define the killing time by

$$\zeta = \inf\{t \geq 0 \,|\, X_{t-} = \Delta \text{ or } X_t = \Delta\}$$

Clearly $\zeta < \infty$ with positive probability if X is sub-stochastic. Since $\{\Delta\}$ is both closed and open, ζ is a stopping time w.r.t. the filtration w.r.t. which X is Markov.

Lemma 5.4.1 *For every $\lambda > 0$ and every nonnegative function $f \in C_0(E_\delta)$, the process*

$$e^{-\lambda t} R_\lambda f(X_t)$$

is a P_ν-supermartingale with respect to the filtration $(\mathcal{F}_t^X)$, for every initial distribution ν.

Proof. By virtue of the Markov property we have

$$\mathsf{E}_\nu(e^{-\lambda t}R_\lambda f(X_t)\,|\,\mathcal{F}_s^X) = e^{-\lambda t}P_{t-s}R_\lambda f(X_s) \quad \mathsf{P}_\nu-\text{a.s.}$$

(see Theorem 3.2.4). Hence, to prove the statement of the lemma it suffices to prove that

$$e^{-\lambda t}P_{t-s}R_\lambda f(x) \le e^{-\lambda}R_\lambda f(x), \quad x\in E. \tag{5.4.1}$$

This is a straightforward calculation (cf. Exercise 5.12). QED

Finally we turn to the regularisation problem for Feller-Dynkin processes.

5.5 Regularisation of Feller-Dynkin processes

5.5.1 Construction of canonical, cadlag version

In this section we consider a Feller-Dynkin transition function P_t on $(E,\mathcal{E})$, with E an open or closed subset of $\mathbf{R}^d$ or $\mathbf{Z}^d$ and $\mathcal{E}$ the Borel-σ-algebra of E. For constructing a cadlag modification, we need to add a *coffin state*, δ say, to our state space E: $E_\delta = E\cup\delta$, such that E_δ is compact, metrisable. δ represents the point at infinity in the one-point compactification of E. Then $\mathcal{E}_\delta = \sigma(\mathcal{E},\{\delta\})$ and we extend the transition function by putting

$$P_t^\delta(x,B) = \begin{cases} P_t(x,B), & x\in E, B\in\mathcal{E}\\ 1_\delta(B), & x=\delta, B\in\mathcal{E}_\delta.\end{cases}$$

Then P_t^δ is a Feller-Dynkin transition function on $(E_\delta,\mathcal{E}_\delta)$. Note that $f\in C_0(E_\delta)$ if and only if the restriction of $f-f(\delta)$ to E belongs to $C_0(E)$.

I plan to include a formal proof of this statement in BN, and I will discuss some topological issues.

By Corollary 3.2.2 for each probability measure ν on $(E_\delta,\mathcal{E}_\delta)$ there exists a probability measure P_ν on the canonical space $(\Omega,\mathcal{F}) = (E_\delta^{\mathrm{R}_+},\mathcal{E}_\delta^{\mathrm{R}_+})$, such that under P_ν the canonical process X is a Markov process with respect to the natural filtration $(\mathcal{F}_t^X)$, with transition function $(P_t^\delta)_t$ and initial distribution ν.

Lemma 5.4.1 allows to use the regularisation results for supermartingales from the preceding chapter.

Theorem 5.5.1 *The canonical Feller-Dynkin process X admits a cadlag modification. More precisely, there exists a cadlag process Y on the canonical space $(\Omega,\mathcal{F})$ such that for all $t\ge 0$ and every initial distribution ν on $(E_\delta,\mathcal{E}_\delta)$ we have $X_t = Y_t$, P_ν-a.s.*

Proof. Fix an arbitrary initial distribution ν on $(E_\delta,\mathcal{E}_\delta)$. Let $\mathcal{H}$ be a countable, dense subset of the space $C_0^+(E)$. Then $\mathcal{H}$ separates the points of E_δ (see Exercise 5.13). By the second statement of Corollary **??**, the class

$$\mathcal{H}' = \{nR_nh\,|\,h\in\mathcal{H}, n\in\mathbf{Z}_+\}$$

has the same property. The proof of Theorem 2.3.2 can be adapted to show that the set

$$\Omega_{h'} = \{\omega\,|\,\lim_{q\downarrow t}h'(X_r)(\omega), \lim_{q\uparrow t}h'(X_r)(\omega) \text{ exist as finite limits for all } t>0\} \tag{5.5.1}$$

is $\mathcal{F}_\infty^X$-measurable. By virtue of Lemma 5.4.1 and Theorem 2.3.2 $\mathsf{P}_\nu(\Omega_{h'}) = 1$ for all $h' \in \mathcal{H}$ and initial measures ν. Take $\Omega' = \bigcap_{h'} \Omega_{h'}$. Then $\Omega' \in \mathcal{F}_\infty^X$ and $\mathsf{P}_\nu(\Omega') = 1$.

In view of Exercise 5.14, it follows that on Ω' the limits

$$\lim_{q \downarrow t} X_q(\omega), \quad \lim_{q \uparrow t} X_q(\omega)$$

exist in E_δ, for all $t \geq 0$, $\omega \in \Omega'$.

Now fix an arbitrary point $x_0 \in E$ and define a new process $Y = (Y_t)$ as follows. For $\omega \notin \Omega'$, put $Y_t(\omega) = x_0$. For $\omega \in \Omega'$ and $t \geq 0$ define

$$Y_t(\omega) = \lim_{q \downarrow t} X_q(\omega).$$

We claim that for every initial distribution ν and $t \geq 0$, we have $X_t = Y_t$ P_ν-a.s. To prove this, let f and g be two functions on $C_0(E_\delta)$. By dominated convergence, and the Markov property

$$\begin{aligned} \mathsf{E}_\nu f(X_t) g(Y_t) &= \lim_{q \downarrow t} \mathsf{E}_\nu f(X_t) g(X_q) \\ &= \lim_{q \downarrow t} \mathsf{E}_\nu \mathsf{E}_\nu (f(X_t) g(X_q) \mid \mathcal{F}_t^X) \\ &= \lim_{q \downarrow t} \mathsf{E}_\nu f(X_t) P_{q-t} g(X_t). \end{aligned}$$

By strong continuity, $P_{q-t} g(X_t) \to g(X_t)$, $q \downarrow t$, P_ν-a.s.. By dominated convergence, it follows that $\mathsf{E}_\nu f(X_t) g(Y_t) = \mathsf{E}_\nu f(X_t) g(X_t)$. By Exercise 5.15 we indeed have that $X_t = Y_t$, P_ν-a.s.

The process Y is right-continuous by construction, and we have shown that Y is a modification of X. It remains to prove that for every initial distribution ν, Y has left limit with P_ν-probability 1. To this end, note that for all $h \in \mathcal{H}'$, the process $h(Y)$ is a right-continuous martingale. By Corollary 2.3.3 this implies that $h(Y)$ has left limits with P_ν-probability 1. In view of Exercise 5.14, it follows that Y has left limits with P_ν-probability 1. QED

Note that Y has the Markov property w.r.t the natural filtration. This follows from the fact that X and Y have the same fdd's and from Characterisation lemma 3.1.5.

By convention we extend each $\omega \in \Omega$ to a map $\omega : [0, \infty] \to E_\delta$ by setting $\omega_\infty = \delta$. We do not assume that the limit of Y_t for $t \to \infty$ exists, but by the above convention $Y_\infty = \delta$.
The formal setup at this point (after redefining) is the *canonical cadlag Feller-Dynkin process X with values in $(E_\delta, \mathcal{E}_\delta)$ and transition function $(P_t^\delta)_t$*. It is defined on the measure space $(\Omega, \mathcal{F})$, where Ω is the set of extended cadlag paths, $\mathcal{F} = \mathcal{E}_\delta^{\mathbb{R}_+} \cap \Omega$ the induced σ-algebra. The associated filtration is the natural filtration $(\mathcal{F}_t^X)_t$. With each initial distribution ν on $(E_\delta, \mathcal{E}_\delta)$, X has induced distribution P_ν (through the outer measure, see Ch.1 Lemma $*$).

5.5.2 Augmented filtration and strong Markov property

Let X be the canonical, cadlag version of a Feller-Dynkin process with state space E_δ (with E a closed or open subset of $\mathbf{R}^d$ or $\mathbf{Z}_+^d$) equipped with the Borel-σ-algebra $\mathcal{E}_\delta$ and Feller-Dynkin transition function P_t^δ. So far, we have been working with the natural filtration $(\mathcal{F}_t^X)$. In general this filtration is neither complete nor right-continuous. We would like to replace it with a larger filtration that satisfies the usual conditions (see Definition 1.6.3) and with respect to which the process X is still a Markov process.

We will first construct a new filtration for every fixed initial distribution ν. Let $\mathcal{F}_\infty^\nu$ be the completion of $\mathcal{F}_\infty^X$ w.r.t. P_ν (cf. BN p. 4) and extend P_ν to this larger σ-algebra.

Denote by $\mathcal{N}^\nu$ the P_ν-negligible sets in $\mathcal{F}_\infty^\nu$, i.e. the sets of zero P_ν-probability. Define the filtration $\mathcal{F}_t^\nu$ by

$$\mathcal{F}_t^\nu = \sigma(\mathcal{F}_t^X, \mathcal{N}^\nu), \quad t \geq 0.$$

Finally, we define the filtration $(\mathcal{F}_t)$ by

$$\mathcal{F}_t = \bigcap_\nu \mathcal{F}_t^\nu$$

where the intersection is taken over all probability measures on the space $(E_\delta, \mathcal{E}_\delta)$. We call $(\mathcal{F}_t)_t$ the *usual augmentation* of the natural filtration $(\mathcal{F}_t^X)_t$. Remarkably, it turns out the we have made the filtration right-continuous!

For a characterisation of the augmented σ-algebras see BN§10.

Theorem 5.5.2 *The filtrations $(\mathcal{F}_t)_t$ and $(\mathcal{F}_t^\nu)_t$ are right-continuous.*

Proof. First note that right-continuity of $(\mathcal{F}_t^\nu)_t$ for all ν implies right-continuity of $(\mathcal{F}_t)_t$. It suffices to show right-continuity of $(\mathcal{F}_t^\nu)_t$.

To this end we will show that $B \in \mathcal{F}_{t+}^\nu$ implies $B \in \mathcal{F}_t^\nu$. So, let $B \in \mathcal{F}_{t+}^\nu$. Then $B \in \mathcal{F}_\infty^\nu$. Hence, there exists a set $B' \in \mathcal{F}_\infty^X$ such that $\mathsf{P}_\nu(B' \triangle B) = 0$. We have

$$\mathbf{1}_{\{B\}} = \mathsf{E}_\nu(\mathbf{1}_{\{B\}} \,|\, \mathcal{F}_{t+}^\nu) \overset{\mathsf{P}_\nu-\text{a.s.}}{=} \mathsf{E}_\nu(\mathbf{1}_{\{B'\}} \,|\, \mathcal{F}_{t+}^\nu).$$

It therefore suffices to show (explain!) that

$$\mathsf{E}_\nu(\mathbf{1}_{\{B'\}} \,|\, \mathcal{F}_{t+}^\nu) = \mathsf{E}_\nu(\mathbf{1}_{\{B'\}} \,|\, \mathcal{F}_t^\nu), \mathsf{P}_\nu - \text{a.s.}$$

To this end, define

$$\mathcal{S} = \{A \in \mathcal{F}_\infty^X \,|\, \mathsf{E}_\nu(\mathbf{1}_{\{A\}} \,|\, \mathcal{F}_{t+}^\nu) = \mathsf{E}_\nu(\mathbf{1}_{\{A\}} \,|\, \mathcal{F}_t^\nu), \mathsf{P}_\nu - \text{a.s.}\}.$$

This is a d-system, and so by BN Lemma 3.7 it suffices to show that $\mathcal{S}$ contains a π-system generating $\mathcal{F}_\infty^X$. The appropriate π-system is the collection of finite-dimensional rectangles. Let A be a finite-dimensional rectangle, i.e.

$$A = \{X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n\},$$

for n, $0 \leq t_1 < \cdots < t_n$, $A_k \in \mathcal{E}_\delta$, $k = 1, \ldots, n$. Then

$$\mathbf{1}_{\{A\}} = \prod_{k=1}^n \mathbf{1}_{\{A_k\}}(X_{t_k}). \tag{5.5.2}$$

By the Feller-Dynkin properties, we need to consider $C_0(E_\delta)$ functions instead of indicator functions. To this end, we will prove for $Z = \prod_{k=1}^n f_k(X_{t_k})$, $f_1, \ldots, f_n \in C_0(E_\delta)$, that

$$\mathsf{E}(Z \mid \mathcal{F}_t^\nu) = \mathsf{E}(Z \mid \mathcal{F}_{t_+}^\nu), \quad \mathsf{P}_\nu - \text{a.s.} \tag{5.5.3}$$

The proof will then be finished by an approximation argument.

Suppose that $t_{k-1} \le t < t_k$ (the case that $t \le t_1$ or $t > t_n$ is similar). Let $h < t_k - t$. Note that $\mathcal{F}_{t+h}^\nu$ and $\mathcal{F}_{t+h}^X$ differ only by P_ν-null sets. Hence

$$\mathsf{E}_\nu(Z \mid \mathcal{F}_{t+h}^\nu) = \mathsf{E}_\nu(Z \mid \mathcal{F}_{t+h}^X), \quad \mathsf{P}_\nu - \text{a.s.}$$

For completeness we will elaborate this. Let $Y_1 = \mathsf{E}_\nu(Z \mid \mathcal{F}_{t+h}^\nu)$ and $Y_2 = \mathsf{E}_\nu(Z \mid \mathcal{F}_{t+h}^X)$. Note that $\mathcal{F}_{t+h}^\nu \supseteq \mathcal{F}_{t+h}^X$. Then Y_1 and Y_2 are both $\mathcal{F}_{t+h}^\nu$-measurable. Then $\{Y_1 > Y_2 + 1/n\} \in \mathcal{F}_{t+h}^\nu$ and so there exists $A_1, A_2 \in \mathcal{F}_{t+h}^X$, with $A_1 \subseteq \{Y_1 > Y_2 + 1/n\} \subseteq A_2$ and $\mathsf{P}_\nu\{A_2 \setminus A_1\} = 0$. Since $A_1 \in \mathcal{F}_{t+h}^X \subseteq \mathcal{F}_{t+h}^\nu$

$$\int_{A_1} (Y_1 - Y_2) d\mathsf{P}_\nu = \int_{A_1} (Z - Z) d\mathsf{P}_\nu = 0,$$

but also

$$\int_{A_1} (Y_1 - Y_2) d\mathsf{P}_\nu \ge \mathsf{P}_\nu\{A_1\}/n.$$

Hence $\mathsf{P}_\nu\{A_1\} = \mathsf{P}_\nu\{A_2\} = 0$, so that also $\mathsf{P}_\nu\{Y_1 > Y_2 + 1/n\} = 0$, for each n. Using that $\{Y_1 > Y_2\} = \cup_n \{Y_1 > Y_2 + 1/n\}$, it follows that $\mathsf{P}_\nu\{Y_1 > Y_2\} = 0$. The reverse is proved similarly.

Use the Markov property to obtain that

$$\mathsf{E}_\nu(Z \mid \mathcal{F}_{t+h}^X) = \prod_{i=1}^{k-1} f_i(X_{t_i}) \mathsf{E}(\prod_{i=k}^n f_i(X_{t_i}) \mid \mathcal{F}_{t+h}^X) = \prod_{i=1}^{k-1} f_i(X_{t_i}) g^h(X_{t+h}), \quad \mathsf{P}_\nu - \text{a.s.}$$

with

$$g^h(x) = P_{t_k - (t+h)} f_k P_{t_{k+1} - t_k} f_{k+1} \cdots P_{t_n - t_{n-1}} f_n(x).$$

By strong continuity of P_t, $\|g^h - g^0\|_\infty \to 0$, as $h \downarrow 0$. By right-continuity of X, $X_{t+h} \to X_t$, P_ν-a.s., so that $g^h(X_{t+h}) \to g^0(X_t)$, $h \downarrow 0$, P_ν-a.s. It follows that

$$\mathsf{E}_\nu(Z \mid \mathcal{F}_{t+h}^\nu) = \mathsf{E}(Z \mid \mathcal{F}_{t+h}^X) = \prod_{i=1}^{k-1} f_i(X_{t_i}) g^h(X_{t+h}) \to$$

$$\to \prod_{i=1}^{k-1} f_i(X_{t_i}) g^0(X_t) = \mathsf{E}_\nu(Z \mid \mathcal{F}_t^X) = \mathsf{E}_\nu(Z \mid \mathcal{F}_t^\nu), \quad \mathsf{P}_\nu - \text{a.s.}$$

On the other hand, by virtue of the Lévy-Doob downward theorem 2.2.16

$$\mathsf{E}_\nu(Z \mid \mathcal{F}_{t+h}^\nu) \to \mathsf{E}_\nu(Z \mid \mathcal{F}_{t+}^\nu), \quad \mathsf{P}_\nu - \text{a.s.}.$$

This implies (5.5.3).

The only thing left to prove is that we can replace Z by $\mathbf{1}_{\{A\}}$ from (5.5.2). Define

$$f_i^m(x) = 1 - m \cdot \min\{\tfrac{1}{m}, d(x, A_i)\},$$

where d is a metric on E_δ consistent with the topology. Clearly, $f_i^m \in C_0(E_\delta)$, and $f_i^m \downarrow f_i$, as $m \to \infty$. For $Z = \prod_{i=1}^n f_i^m(X_{t_i})$ (5.5.3) holds. Use monotone convergence, to obtain that (5.5.3) holds for $Z = \mathbf{1}_{\{A\}}$ given in (5.5.2). Hence, the d-system $\mathcal{S}$ contains all finite-dimensional rectangles, and consequently $\mathcal{F}_\infty^X$. This is precisely what we wanted to prove. QED

Our next aim is to prove that the generalised Markov property from Theorem 3.2.4 remains true if we replace the natural filtration $(\mathcal{F}_t^X)_t$ by its usual augmentation. This will imply that X is still a Markov process in the sense of the old definition 3.1.3.

First we will have to address some measurability issues. We begin by considering the completion of the Borel-σ-algebra $\mathcal{E}_\delta$ on E_δ. If μ is a probability measure on $(E_\delta, \mathcal{E}_\delta)$, we denote by $\mathcal{E}_\delta^\mu$ the completion of $\mathcal{E}_\delta$ w.r.t μ. We then define

$$\mathcal{E}^* = \bigcap_\mu \mathcal{E}_\delta^\mu,$$

where the intersection is taken over all probability measures on $(E_\delta, \mathcal{E}_\delta)$. The σ-algebra $\mathcal{E}^*$ is called the σ-algebra of *universally measurable sets.*

Lemma 5.5.3 *If Z is a bounded or non-negative, $\mathcal{F}_\infty$-measurable random variable, then the map $x \mapsto \mathsf{E}_x Z$ is $\mathcal{E}^*$-measurable, and*

$$\mathsf{E}_\nu Z = \int_x \mathsf{E}_x Z \, \nu(dx),$$

for every initial distribution ν.

Proof. Fix ν. Note that $\mathcal{F}_\infty \subseteq \mathcal{F}_\infty^\nu$. By definition of $\mathcal{F}_\infty^\nu$, there exist two $\mathcal{F}_\infty^X$ random variables Z_1, Z_2, such that $Z_1 \leq Z \leq Z_2$ and $\mathsf{E}_\nu(Z_2 - Z_1) = 0$. It follows for every $x \in E$ that $\mathsf{E}_x Z_1 \leq \mathsf{E}_x Z \leq \mathsf{E}_x Z_2$. Moreover, the maps $x \mapsto \mathsf{E}_x Z_i$ are $\mathcal{E}_\delta$-measurable by Lemma 3.2.3 and

$$\int (\mathsf{E}_x Z_2 - \mathsf{E}_x Z_1)\nu(dx) = \mathsf{E}_\nu(Z_2 - Z_1) = 0.$$

By definition of $\mathcal{E}_\delta^\nu$ this shows that $x \mapsto \mathsf{E}_x Z$ is $\mathcal{E}_\delta^\nu$-measurable and that

$$\mathsf{E}_\nu Z = \mathsf{E}_\nu Z_1 = \int \mathsf{E}_x Z_1 \, \nu(dx) = \int \mathsf{E}_x Z \, \nu(dx).$$

Since ν is arbitrary it follows that $x \mapsto \mathsf{E}_x Z$ is in fact $\mathcal{E}^*$-measurable. For a detailed argumentation go through the standard machinery. QED

Lemma 5.5.4 *For all $t \geq 0$, the random variable X_t is measurable as a map from $(\Omega, \mathcal{F}_t)$ to $(E_\delta, \mathcal{E}^*)$.*

Proof. Take $A \in \mathcal{E}^*$, and fix an initial distribution ν on $(E_\delta, \mathcal{E}_\delta)$. Denote the distribution of X_t on $(E_\delta, \mathcal{E}_\delta)$ under P_ν by μ. Since $\mathcal{E}^* \subseteq \mathcal{E}_\delta^\mu$, there exist $A_1, A_2 \in \mathcal{E}_\delta$, such that $A_1 \subseteq A \subseteq A_2$ and $\mu(A_2 \setminus A_1) = 0$. Consequently, $X_t^{-1}(A_1) \subseteq X_t^{-1}(A) \subseteq X_t^{-1}(A_2)$. Since $X_t^{-1}(A_1), X_t^{-1}(A_2) \in \mathcal{F}_t^X$ and

$$\mathsf{P}_\nu\{X_t^{-1}(A_1) \setminus X_t^{-1}(A_2)\} = \mathsf{P}_\nu(X_t^{-1}(A_2 \setminus A_1) = \mu(A_2 \setminus A_1) = 0,$$

the set $X_t^{-1}(A)$ is contained in the P_ν-completion of $\mathcal{F}_t^X$. But ν is arbitrary, and so the proof is complete. QED

Corollary 5.5.5 *Let Z be an $\mathcal{F}_\infty$-measurable random variable, bounded or non-negative. Let ν be any initial distribution and let μ denote the P_ν-distribution of X_t. Then $\mathsf{E}_\nu \mathsf{E}_{X_t} Z = \mathsf{E}_\mu Z$.*

We can now prove that the generalised Markov property, formulated in terms of shift operators, is still valid for the usual augmentation $(\mathcal{F}_t)$ of the natural filtration of the Feller-Dynkin process.

Theorem 5.5.6 (Generalised Markov property) *Let Z be a $\mathcal{F}_\infty$-measurable random variable, non-negative or bounded. Then for every $t > 0$ and initial distribution ν,*

$$\mathsf{E}_\nu(Z \circ \theta_t \,|\, \mathcal{F}_t) = \mathsf{E}_{X_t} Z, \quad \mathsf{P}_\nu - \text{a.s.}$$

In particular, X is an $(E_\delta, \mathcal{E}^*)$-valued Markov process w.r.t. $(\mathcal{F}_t)_t$.

Proof. We will only prove the first statement. Lemmas 5.5.3 and 5.5.4 imply that $\mathsf{E}_{X_t} Z$ is $\mathcal{F}_t$-measurable. So we only have to prove for $A \in \mathcal{F}_t$ that

$$\int_A Z \circ \theta_t d\mathsf{P}_\nu = \int_A \mathsf{E}_{X_t} Z d\mathsf{P}_\nu. \tag{5.5.4}$$

Assume that Z is bounded, and denote the law of X_t under P_ν by μ. By definition of $\mathcal{F}_\infty$ there exist a $\mathcal{F}_\infty^X$-measurable random variable Z', such that $\{Z \neq Z'\} \subset \Gamma$, $\Gamma \in \mathcal{F}_\infty^X$ and $\mathsf{P}_\nu(\Gamma) = 0$ (use the standard machinery). We have that

$$\{Z \circ \theta_t \neq Z' \circ \theta_t\} = \theta^{-1}\{Z \neq Z'\} \subseteq \theta^{-1}(\Gamma).$$

By Theorem 3.2.4

$$\mathsf{P}_\nu\{\theta_t^{-1}(\Gamma)\} = \mathsf{E}_\nu(\mathbf{1}_{\{\Gamma\}} \circ \theta_t) = \mathsf{E}_\nu \mathsf{E}_\nu(\mathbf{1}_{\{\Gamma\}} \circ \theta_t \,|\, \mathcal{F}_t^X) = \mathsf{E}_\nu \mathsf{E}_{X_t} \mathbf{1}_{\{\Gamma\}} = \int \mathsf{E}_x \mathbf{1}_{\{\Gamma\}} \mu(dx) = \mathsf{P}_\mu \mathbf{1}_{\{\Gamma\}} = 0,$$

since the distribution of X_t under P_ν is given by μ. This shows that we may replace the left-handside of (5.5.4) by $\int_A Z' \circ \theta_t d\mathsf{P}_\nu$. Further, we have used that the two probability measures $B \mapsto \mathsf{E}_\nu \mathsf{E}_{X_t} \mathbf{1}_{\{B\}}$ and $B \mapsto \mathsf{P}_\mu(B)$ coincide for $B \in \mathcal{F}_\infty$. Since $\mathsf{P}_\mu\{Z \neq Z'\} \leq \mathsf{P}_\mu\{\Gamma\} = 0$

$$\mathsf{E}_\nu |\mathsf{E}_{X_t} Z - \mathsf{E}_{X_t} Z'| \leq \mathsf{E}_\nu \mathsf{E}_{X_t} |Z - Z'| = \mathsf{E}_\mu |Z - Z'| = 0.$$

It follows that $\mathsf{E}_{X_t} Z = \mathsf{E}_{X_t} Z'$, P_ν-a.s. In the right-handside of (5.5.4) we may replace Z by Z' as well. Since Z' is $\mathcal{F}_\infty^X$-measurable, the statement now follows from Theorem 3.2.4 (we have to use that a set $A \in \mathcal{F}_t$ can be replaced by a set $A' \in \mathcal{F}_t^X$). QED

We consider again a Feller-Dynkin canonical cadlag process X with state space $(E_\delta, \mathcal{E}_\delta)$, where $E \subseteq \mathbf{R}^d, \mathbf{Z}_+^d$. This is a Markov process with respect to the usual augmentation $(\mathcal{F}_t)_t$ of the natural filtration of the canonical process on the compactified state space E_δ. As before, we denote shift operators by θ_t.

In this section we will prove that for Feller-Dynkin processes the Markov property of Theorem 5.5.6 does not only hold for deterministic times t, but also for $(\mathcal{F}_t)_t$-stopping times.

This is called the *strong Markov property.* Recall that for deterministic $t \geq 0$ the shift operator θ_t on the canonical space Ω maps a path $s \mapsto \omega_s$ to the path $s \mapsto \omega_{t+s}$. Likewise, for a random time τ we now define θ_τ as the operator that maps the path $s \mapsto \omega_s$ to the path $s \mapsto \omega_{\tau(\omega)+s}$. If τ equals the deterministic time t, then $\tau(\omega) = t$ for all ω and so θ_τ equals the old operator θ_t.

Since the canonical process X is just the identity on the space Ω, we have for instance that $(X_t \circ \theta_\tau)(\omega) = X_t(\theta_\tau(\omega)) = (\theta_\tau)(\omega))_t = \omega_{\tau(\omega)+t} = X_{\tau(\omega)+t}(\omega)$, in other words $X_t \circ \theta_\tau = X_{\tau+t}$. So the operators θ_τ can still be viewed as time shifts.

Theorem 5.5.7 (Strong Markov property) *Let Z be an $\mathcal{F}_\infty$-measurable random variable, non-negative or bounded. Then for every $(\mathcal{F}_t)$-stopping time τ and initial distribution ν, we have P_ν-a.s.*

$$\mathsf{E}_\nu(Z \circ \theta_\tau \,|\, \mathcal{F}_\tau) = \mathsf{E}_{X_\tau}(Z). \tag{5.5.5}$$

Note that on $\tau = \infty$ by convention $X_\tau = \delta$.

Proof. First, check that X_τ is $\mathcal{F}_\tau$-measurable (use arguments similar to Lemmas 1.6.16 and 1.6.17. Further, check that $\mathsf{E}_{X_\tau} Z$ is bounded or non-negative $\mathcal{F}_\tau$-measurable for all bounded or non-negative $\mathcal{F}_\infty$-measurable random variables Z.

Suppose that τ is a stopping time that takes values in a countable set $D \cup \{\infty\}$. Since θ_τ equals θ_d on the event $\{\tau = d\}$, we have (see Ch.1 Exercise 1.21) for every initial distribution ν

$$\begin{aligned}
\mathsf{E}_\nu(Z \circ \theta_\tau \,|\, \mathcal{F}_\tau) &= \sum_{d\in D} \mathbf{1}_{\{\tau=d\}} \mathsf{E}_\nu(Z \circ \theta_\tau \,|\, \mathcal{F}_\tau) \\
&= \sum_{d\in D} \mathbf{1}_{\{\tau=d\}} \mathsf{E}_\nu(Z \circ \theta_d \,|\, \mathcal{F}_d) \\
&= \sum_{d\in D} \mathbf{1}_{\{\tau=d\}} \mathsf{E}_{X_d} Z = \mathsf{E}_{X_\tau} Z,
\end{aligned}$$

P_ν-a.s. by the Markov property.

Let us consider a general stopping time τ. We will first show that (5.5.5) holds for Z an $\mathcal{F}^X_\infty$-measurable random variable. A similar reasoning as in the proof of Theorem 5.5.2 shows (check yourself) that we can restrict to showing (5.5.5) for Z of the form

$$Z = \prod_{i=1}^k f_i(X_{t_i}),$$

$t_1 < \ldots < t_k$, $f_1, \ldots, f_k \in C_0(E_\delta)$, $k \in \mathbf{Z}_+$. Define countably valued stopping times τ_n as follows:

$$\tau_n(\omega) = \sum_{k=0}^\infty \mathbf{1}_{\{k2^{-n} \leq \tau(\omega) < (k+1)\cdot 2^{-n}\}} \frac{k+1}{2^n} + \mathbf{1}_{\{\tau(\omega)=\infty\}} \cdot \infty.$$

Clearly $\tau_n(\omega) \downarrow \tau(\omega)$, and $\mathcal{F}_{\tau_n} \supseteq \mathcal{F}_{\tau_{n+1}} \supseteq \cdots \supseteq \mathcal{F}_\tau$ for all n by virtue of Exercise 1.18. By the preceding,

$$\mathsf{E}_\nu(\prod_i f_i(X_{t_i}) \circ \theta_{\tau_n} \,|\, \mathcal{F}_{\tau_n}) = \mathsf{E}_{X_{\tau_n}} \prod_i f_i(X_{t_i}) = \mathbf{1}_{\{\tau_n<\infty\}} g(X_{\tau_n}),$$

P_ν-a.s., where

$$g(x) = P^\delta_{t_1} f_1 P^\delta_{t_2-t_1} f_2 \cdots P^\delta_{t_k-t_{k-1}} f_k(x).$$

By right-continuity of paths, the right-hand side converges P_ν-a.s. to $g(X_\tau)$. By virtue of Corollary 2.2.17 the left-handside converges P_ν-a.s. to

$$\mathsf{E}_\nu(\prod_i f_i X_{t_i}) \circ \theta_\tau \mid \mathcal{F}_\tau),$$

provided that $\mathcal{F}_\tau = \cap_n \mathcal{F}_{\tau_n}$. Note that $\mathcal{F}_\tau \subseteq \cap_n \mathcal{F}_{\tau_n}$, and so we have to prove the reverse implication. We would like to point out that problems may arise, since $\{\tau_n \leq t\}$ need not increase to $\{\tau \leq t\}$. However, we do have $\{\tau \leq t\} = \cap_m \cup_n \{\tau_n \leq t + 1/m\}$.

Let $A \in \mathcal{F}_{\tau_n}$ for all n. Then $A \cap \{\tau_n \leq t + 1/m\} \in \mathcal{F}_{t+1/m}$ for all n. Hence $A \cap \cup_n\{\tau_n \leq t + 1/m\} \in \mathcal{F}_{t+1/m}$, and so $A \cap \{\tau \leq t\} = A \cap_m \cup_n \{\tau_n \leq t + 1/m\} \in \cap_m \mathcal{F}_{t+1/m} = \mathcal{F}_{t^+} = \mathcal{F}_t$.

This suffices to show (5.5.5) for Z $\mathcal{F}^X_\infty$-measurable. Let next Z be a $\mathcal{F}_\infty$-measurable random variable. We will now use a similar argument to the proof of Theorem 5.5.6.

Denote the distribution of X_τ under P_ν by μ. By construction, $\mathcal{F}_\infty$ is contained in the P_μ-completion of $\mathcal{F}^X_\infty$. Hence there exist two $\mathcal{F}^X_\infty$-measurable, bounded or non-negative random variables Z', Z'', with $Z' \leq Z \leq Z''$ and $\mathsf{E}_\mu(Z'' - Z') = 0$. It follows that $Z' \circ \theta_\tau \leq Z \circ \theta_\tau \leq Z'' \circ \theta_\tau$. By the preceding

$$\begin{aligned}
\mathsf{E}_\nu(Z'' \circ \theta_\tau - Z' \circ \theta_\tau) &= \mathsf{E}_\nu \mathsf{E}_\nu(Z'' \circ \theta_\tau - Z' \circ \theta_\tau \mid \mathcal{F}_\tau) \\
&= \mathsf{E}_\nu \mathsf{E}_{X_\tau}(Z'' - Z') \\
&= \int \mathsf{E}_x(Z'' - Z')\mu(dx) = \mathsf{E}_\mu(Z'' - Z') = 0.
\end{aligned}$$

It follows that $Z \circ \theta_\tau$ is measurable with respect to the P_ν-completion of $\mathcal{F}^X_\infty$. Since ν is arbitrary, we conclude that $Z \circ \theta_\tau$ is $\mathcal{F}_\infty$-measurable. Observe that P_ν-a.s.

$$\mathsf{E}(Z' \circ \theta_\tau \mid \mathcal{F}_\tau) \leq \mathsf{E}(Z \circ \theta_\tau \mid \mathcal{F}_\tau) \leq \mathsf{E}(Z'' \circ \theta_\tau \mid \mathcal{F}_\tau).$$

By the preceding, the outer terms P_ν-a.s. equal $\mathsf{E}_{X_\tau} Z'$ and $\mathsf{E}_{X_\tau} Z''$ respectively. These are P_ν-a.s. equal. Since $Z' \leq Z \leq Z''$ they are both P_ν-a.s. equal to $\mathsf{E}_{X_\tau} Z$. QED

5.6 Feller diffusions

5.7 Exercises

Exercise 5.1 Show for the generator A of the Ornstein-Uhlenbeck process (cf. Example 3.1.2 (B) and 5.1.5) that

$$\mathsf{A}f(x) = \tfrac{1}{2}\sigma^2 f''(x) - \alpha x f'(x), \quad x \in \mathbf{R}, f \in \{g : \mathbf{R} \to \mathbf{R} \,|\, g, g'', \hat{g} \in C_0(\mathbf{R}), \hat{g}(x) = xg'(x)\}.$$

You may use the expression for the generator of Brownian motion derived in Example 5.3.3. Hint: denote by P_t^X and P_t^W the transition functions of Ornstein-Uhlenbeck process and BM respectively. Show that $P_t^X f(x) = P_{g(t)}^W f(e^{-\alpha t}x)$ where $g(t) = \sigma^2(1 - e^{-2\alpha t})/2\alpha$.

Exercise 5.2 Prove the Integration Lemma.

Exercise 5.3 Prove the claim made in Example 5.2.1. Hint: to derive the explicit expression for the resolvent kernel it is needed to calculate integrals of the form

$$\int_0^\infty \frac{e^{-a^2 t - b^2/t}}{\sqrt{t}} dt.$$

To this end, first perform the substitution $t = (b/a)s^2$. Next, make a change of variables $u = s - 1/s$ and observe that $u(s) = s - 1/s$ is a continuously differentiable bijective function from $(0, \infty)$ to $\mathbf{R}$, the inverse $u^{-1} : \mathbf{R} \to (0, \infty)$ of which satisfies $u^{-1}(t) - u^{-1}(-t) = t$, whence $(u^{-1})'(t) + (u^{-1})'(t) = 1$.

Exercise 5.4 Prove the validity of the expression for the resolvent of the Markov jump process in Example 5.2.2.

Exercise 5.5 Show that the Markov process from Example 3.2.1 is a Feller-Dynkin process if $PC_0(E) \subset C_0(E)$. Give an example of a Markov jump process that is not a Feller-Dynkin process.

Exercise 5.6 Prove Lemma 5.3.3. That means that you may assume the validity of Lemma 3.1.4.

Exercise 5.7 In the proof of Lemma 5.3.6, show that $\mathsf{P}\{\eta_r > nt\} \le \hat{p}^n$ for $n = 0, 1, \ldots$. Hint: use the Markov property.

Exercise 5.8 Suppose that X is a real-valued canonical continuous Feller-Dynkin process, with generator

$$\mathsf{A}f(x) = \alpha(x)f'(x) + \tfrac{1}{2}f''(x), \quad x \in \mathbf{R},$$

for $f \in \mathcal{D} = \{g : \mathbf{R} \to \mathbf{R} \,|\, g, g', g'' \in C_0(\mathbf{R})\}$, where α is an arbitrary but fixed continuous, bounded function on $\mathbf{R}$. Suppose that there exists a function $f \in \mathcal{D}$, $f \not\equiv 0$, such that

$$\mathsf{A}f(x) = 0, \quad x \in \mathbf{R}. \tag{5.7.1}$$

Then the martingale M_t^f has a simpler structure, namely $M_t^f = f(X_t) - f(X_0)$.

i) Show that for $f \in \mathcal{D}(\mathsf{A})$, satisfying (5.7.1), Dynkin's formula holds, for all $x \in E$. Hence the requirement that $\mathsf{E}_x\tau < \infty$ is not necessary!

Let $(a,b) \subset \mathbf{R}$, $a < b$. Put $\tau = \inf\{t > 0 \,|\, X_t \in (-\infty, a] \cup [b, \infty)\}$. Define $p_x = \mathsf{P}_x\{X_\tau = b\}$.

ii) Assume that $\tau < \infty$, P_x-a.s. for all $x \in (a,b)$. Prove that $p_x = \frac{f(x)-f(a)}{f(b)-f(a)}$, $x \in (a,b)$, provided that $f(b) \neq f(a)$.

iii) Let X be a real-valued canonical, right-continuous Feller-Dynkin process, such that $X_t = X_0 + \beta t + \sigma W_t$, where X_0 and $(W_t)_t$ are independent, and $(W_t)_t$ a standard BM. Show for the generator A that $\mathcal{D}(\mathsf{A}) \supset \mathcal{D} = \{g : \mathbf{R} \to \mathbf{R} \,|\, f, f', f'' \in C_0(\mathbf{R})$ and is given by

$$\mathsf{A}f = \beta f' + \tfrac{1}{2}\sigma^2 f''$$

for $f \in \mathcal{D}$) (you may use the generator of BM).

Show that $\tau < \infty$, P_x-a.s., $x \in (a,b)$. Determine p_x for $x \in (a,b)$. Hint: you have to solve a simple differential equation to find f with $\beta f' + \sigma^2 f''/2 = 0$. This f is not a $C_0^2(\mathbf{R})$ function. Explain that this is no problem since X_t only lives on $[a,b]$ until the stopping time.

iv) Let X be the Ornstein-Uhlenbeck process (cf. Example 3.1.2 (B) and 3.3.14). Show that $\tau < \infty$, P_x-a.s. and determine p_x for $x \in (a,b)$. You may use the result of Exercise 5.1 on the generator of the Ornstein-Uhlenbeck process. See also hint of (iii). Notice that the solution can only be represented as an integral.

Exercise 5.9 Consider canonical Geometric Brownian Motion from Example 3.1.2 (C) with continuous paths. Let

$$\mathcal{D} = \{f \in C_0(\mathbf{R}_+) \,|\, x \mapsto f'(x), f''(x), xf'(x), x^2 f''(x), x \in \mathbf{R}_+ \in C_0(E)\}.$$

Geometric BM has generator

$$\mathsf{A}f(x) = \mu x f'(x) + \frac{1}{2}\sigma^2 x^2 f''(x), x \in \mathbf{R}_+, \tag{5.7.2}$$

on $f \in \mathcal{D} \subset \mathcal{D}(\mathsf{A})$.

i) Show that this is true for $\mu = \sigma^2/2$ and *characterise* all of $\mathcal{D}(\mathsf{A})$ in this case. You may use the results for BM (see Example 5.3.3).

ii) Give the main steps in the proof of (5.7.2) in the general case, for functions in $\mathcal{D}$.

Exercise 5.10 We want to construct a standard BM in $\mathbf{R}^d$ ($d < \infty$): this is an $\mathbf{R}^d$-valued process $W = (W^1, \ldots, W^d)$, where $W^1, \ldots, W^d$ are independent standard BMin $\mathbf{R}$.

i) Sketch how to construct d-dimensional BM.

ii) Show that W has stationary, independent increments.

iii) Show that W is a Feller-Dynkin process with respect to the natural filtration, with transition function

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

where $y = (y_1, \ldots, y_d)$, $x = (x_1, \ldots, x_d) \in \mathbf{R}^d$ and $\|y - x\| = \sqrt{\sum_{i=1}^d (y_i - x_i)^2}$ is the $L^2(\mathbf{R}^d)$-norm.

Exercise 5.11 (Continuation of Exercise 5.10) Let X be an $\mathbf{R}^d$-valued canonical continuous Feller-Dynkin process, such that $X_t = X_0 + W_t$, where X_0 is an $\mathbf{R}^d$-valued r.v. and $(W_t)_t$ a standard d-dimensional BM that is independent of X_0. Notice that X is strong Markov.

We would like to show that the generator is defined by

$$\mathsf{A}f(x) = \tfrac{1}{2}\Delta f(x), \qquad (5.7.3)$$

where $\Delta f(x) = \sum_{i=1}^d \frac{\partial^2}{\partial x_i^2} f(x)$ is the Laplacian of f, for $f \in \mathcal{D} = \{f : \mathbf{R}^d \to \mathbf{R} \mid f, \frac{\partial}{\partial x_i} f, \frac{\partial^2}{\partial x_i \partial x_j} f \in C_0(\mathbf{R}^d), i,j = 1,\ldots,d\}$. We again want to use the characteristic operator. To this end, define for $r > 0$

$$\tau_r = \inf\{t \geq 0 \mid \|X_t - X_0\| \geq r\}.$$

i) Argue that τ_r is a finite $(\mathcal{F}_t)_t$-stopping time. Show that $\mathsf{E}_x \tau_r = r^2/d$ (by using optional stopping). Argue that X_{τ_r} has the uniform distribution on $\{y \mid \|y - x\| = r\}$.

ii) Show the validity of (5.7.3) for $f \in \mathcal{D}$ (use the characteristic operator). Argue that this implies $\mathcal{D}(\mathsf{A}) \supset \mathcal{D}$.

iii) For $0 < a < \|x\| < b$, show that

$$\mathsf{P}_x\{T_a < T_b\} = \begin{cases} \frac{\log b - \log \|x\|}{\log b - \log a}, & d = 2 \\ \frac{\|x\|^{2-d} - b^{2-d}}{a^{2-d} - b^{2-d}}, & d \geq 3, \end{cases}$$

where $T_a = \inf\{t \geq 0 \mid \|X_t\| \leq a\}$ and $T_b = \inf\{t \geq 0 \mid \|X_t\| \geq b\}$. Hint: a similar procedure as in Exercise 5.8.

iv) Compute $\mathsf{P}_x\{T_a < \infty\}$ for x with $a < \|x\|$.

Exercise 5.12 Prove (5.4.1) in the proof of Lemma 5.4.1.

Exercise 5.13 Suppose that $E \subseteq \mathbf{R}^d$. Show that every countable, dense subset $\mathcal{H}$ of the space $C_0^+(E)$ of non-negative functions in $C_0(E)$ separates the points of E_δ. This means that for all $x \neq y$ in E there exists a function $h \in \mathcal{H}$, such that $h(x) \neq h(y)$, and for all $x \in E$ there exists a function $h \in \mathcal{H}$, such that $h(x) \neq h(\delta) = 0$.

Exercise 5.14 Let (X, d) be a compact metric space (with metric d). Let $\mathcal{H}$ be a class of non-negative, continuous functions on X that separates the points of X. Prove that $d(x_n, x) \to 0$ if and only if $h(x_n) \to h(x)$ for all $h \in \mathcal{H}$. Hint: suppose that $\mathcal{H} = \{h_1, h_2, \ldots\}$, endow $\mathbf{R}^\infty$ with the product topology and consider the map $A(x) = (h_1(x), h_2(x), \ldots)$.

Exercise 5.15 Let X, Y be two random variables defined on the same probability space, taking values in the Polish space E equipped with the Borel-σ-algebra. Show that $X = Y$ a.s. if and only if $\mathsf{E}f(X)g(Y) = \mathsf{E}f(X)g(X)$ for all $C_0(E)$ functions f and g on E. Hint: use the monotone class theorem (see BN) and consider the class $\mathcal{H} = \{h : E \times E \to \mathbf{R} \mid h\ \mathcal{E} \times \mathcal{E}-$measurable, $\|h\|_\infty < \infty, \mathsf{E}h(X,Y) = \mathsf{E}h(X,X)\}$.

Exercise 5.16 Let $(\mathcal{F}_t)_t$ be the usual augmentation of the natural filtration of a canonical, cadlag Feller-Dynkin process. Show that for every nonnegative, $\mathcal{F}_t$-measurable random variable Z and every finite stopping time τ, the random variable $Z \circ \tau$ is $\mathcal{F}_{\tau+t}$-measurable. Hint: first prove it for $Z = \mathbf{1}_{\{A\}}$, $A \in \mathcal{F}_t^X$. Next, prove it for $Z = \mathbf{1}_{\{A\}}$, $A \in \mathcal{F}_t$, and use the fact that $A \in \mathcal{F}_t^\nu$ if and only if there exists $B \in \mathcal{F}_t^X$ and $C, D \in N^\nu$, such that $B \setminus C \subseteq A \subseteq B \cup D$ (this follows from Problem 10.1 in BN). Finally prove it for arbitrary Z.

Exercise 5.17 Let X be a Feller-Dynkin canonical cadlag process and let $(\mathcal{F}_t)_t$ be the usual augmentation. Suppose that we have $(\mathcal{F}_t)_t$-stopping times $\tau_n \uparrow \tau$ a.s. Show that $\lim_n X_{\tau_n} = X_\tau$ a.s. on $\{\tau < \infty\}$. This is called the *quasi-left continuity* of Feller-Dynkin processes. Hint: first argue that it is sufficient to show the result for bounded τ. Next, put $Y = \lim_n X_{\tau_n}$ and explain why this limit exists. Use the strong Markov property to show for $f, g \in C_0(E_\delta)$ that

$$\mathsf{E}_x f(Y) g(X_\tau) = \lim_{t \downarrow 0} \lim_n \mathsf{E}_x f(X_{\tau_n}) g(X_{\tau_n + t}) = \mathsf{E}_x f(Y) g(Y).$$

The claim then follows from Exercise 5.15.