An Introduction to Stochastic Processes in Continuous Time

adaptation of the text by

Harry van Zanten

to be used at your own expense

May 16, 2010

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 tis 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 $(E,\mathcal{E})$. 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 $T=\mathbf{R}_+=[0,\infty)$. 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. For instance, a discrete-time process viewed as the continuous-time process described earlier, is a process with right-continuous sample paths.

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

Definition 1.1.2 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) $W_t - W_s$ is independent of $(W_u, u \leq s)$ for all $s \leq t$, that is, $\sigma(W_t - W_s)$ and $\sigma(W_u, u \leq s)$ are independent;

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 process with property (ii) is said to have *independent increments*. Property (iii) implies that the distribution of the increment $W_t - W_s$ only depends on $t - s$. This is called *stationarity of the increments*. 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 1.1.3 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 takes only values in $\mathbf{Z}_+$;
- **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) $\limsup_{s\downarrow t} N(s) \leq \liminf_{s\uparrow t} N_s + 1$, for all $t \geq 0$.

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

iii) (independence of increments) $\sigma(N_t - N_s)$ and $\sigma(N_u, u \leq s)$ are independent;

iv) (stationarity of increments) $N_t - N_s \stackrel{\mathsf{d}}{=} N_{t-s}$ for all $s \leq t$;

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

$$N(t) = \max\{n \mid S_n \leq t\}.$$

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 1.1.4 *There exists a set Ω^* with $\mathsf{P}\{\Omega^*\} = 1$, such that $N(t, \omega) < \infty$ for all $t \geq 0$, $\omega \in \Omega$, and $S_n(\omega) < S_{n+1}(\omega)$, $n = 0, \ldots$. Moreover, $N(t)$ is $\mathcal{F}/\mathcal{E}$-measurable.*

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 of these terms is measurable. The infinite sum is then measurable as well, being the monotone limit of measurable functions. QED

The following lemma allows to restrict to Ω^*. We will do so, but denote the restricted probability space again by $(\Omega, \mathcal{F}, \mathsf{P})$.

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 1.1.5 *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 $\mathcal{A} = \{A \cap F \mid F \in \mathcal{F}\}$.*

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

i) *N has properties (i,...,iv) from Definition 1.1.3. Moreover, $N(t)$ 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

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

First let $i,j>0$. 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,

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

For $i=0$, we get the same conclusion, and (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 $\omega\mapsto N(u,\omega)$ measurable. In Exercise **??** you will be asked to deduce its structure, and to show that 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

$$\mathsf{P}\{N(t_l)=i_l, l=0,\ldots,n, N(t)-N(s)=i\}=\mathsf{P}\{N(t_l)=i_l, l=0,\ldots,n\}\cdot\mathsf{P}\{N(t)-N(s)=i\}.$$

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 $\omega'_t \leq \liminf_{s\uparrow t} \omega'_s + 1$.

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. 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 $\mathcal{F}' = \mathcal{E}^{[0,\infty)} \cap \Omega'$.

Review BN §2 on measurability issues for a description of $\mathcal{E}^{[0,\infty)}$

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 $\mathsf{P}'\{A\} = \mathsf{P}\{\omega \mid N(\omega) \in A\}$.

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 1.1.2.

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

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

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 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 1.1.2.*

Proof. First 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

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

Next, 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 $(W_{t_0}, W_{t_1} - W_{t_0}, \ldots, W_{t_n} - W_{t_{n-1}}) \stackrel{\mathsf{d}}{=} \mathsf{N}(0, \mathbf{\Sigma})$. Finally, 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 $(W_{t_0}, W_{t_1}, \ldots, W_{t_n}) \stackrel{\mathsf{d}}{=} \mathsf{N}(0, \mathbf{\Sigma}_{t_0,\ldots,t_n})$. Then finish the proof. 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: $\{\sup_{t\geq 0} W_t \leq x\}$, $\inf\{t \geq 0 \mid W_t = x\}$.

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 $B_1, \ldots, B_n \in \mathcal{E}$

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

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. Exercise 1.2 for the second case). 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*.

Theorem 1.3.3 (Kolmogorov's continuity criterion) *Let $X = (X_t)_{t\in[0,T]}$ be an $\mathbf{R}^d$-valued 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 of X.

Note: $\beta > 0$ is needed for the continuous modification to exist. See Exercise 1.4.

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

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

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

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

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

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

It follows that

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

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 $0 < t - s < 2^{-N_\omega}$. Then $t - s \leq 2^{-(N_\omega+1)}$. Hence, there exists $n \geq N_\omega$, such that $2^{-(n+1)} \leq t - s < 2^{-n}$.

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

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

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 $s = k/2^{n+1}$ 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 $m \geq n+1$. We will show its validity for $m+1$.

Put $s' = \min\{u \in D_m \,|\, u \geq s\}$ and $t' = \max\{u \in D_m \,|\, u \leq t\}$. By construction $s \leq s' \leq t' \leq t$, and $s' - s, t - t' \leq 2^{-(m+1)}$. Then $0 < t' - s' \leq t - s < 2^{-n}$. 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

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

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 $2^{-(n+1)} \leq t - s < 2^{-n}$. Then there exists $m \geq n+1$ such that $t, s \in D_m$. Apply (1.3.5) to obtain

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

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 1.3). 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 1.4). 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 1.1.2. 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 1.6 you are asked to show that X is a BrowniFan motion process. QED

Lemma 1.1.5 allows us to restrict to continuous paths.

Kolmogorov's continuity criterion applied to 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.

Note that one can always construct the canonical process to have a desired distribution. 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.

Suppose 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 1.7. QED

Brownian motion is a special case of a Gaussian process. In particula 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 1.1.2. We have the following result.

Lemma 1.4.3 *A 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!

Theorem 1.4.4 *Let W be a* BM *process. Then the following are* BM *processes as well:*

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

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 $X_t = tW_{1/t}$, $t > 0$.*

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 1.8 you are asked to show this for the processes in (i,ii,iii). We will only prove (iv).

Clearly the processes X in (iv) is a stochastic process. We will show that almost all sample paths of X are continuous. A simple application of Lemma 1.4.2 finishes the proof.

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. 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.

Assume that W is a stochastic process defined on the underlying probability space $(\Omega, \mathcal{F}, \mathsf{P})$. Let $\Omega^* = \{\omega \in \Omega \,|\, (W_t(\omega))_{t\geq 0} \text{ continuous}\}$. By assumption $\mathsf{P}\{\Omega^*\} = 1$. Further, $(X_t)_{t>0}$ has continuous paths on Ω^*.

Let $\omega \in \Omega^*$. 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. It follows that (cf. Exercise 1.9) the probability of the latter equals

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

since $(W_t)_{t\geq 0}$ has continuous paths on Ω^*. Hence $\mathsf{P}\{\{\omega : \lim_{t\downarrow 0} X_t(\omega) = 0\} \cap \Omega^*\} = 1$, and so $\mathsf{P}\{\omega : \lim_{t\downarrow 0} X_t(\omega) = 0\} = 1$.

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 1.2) 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 $\mathsf{P}\{\sup_t W_t \leq 0\}$. 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 $\mathbf{R}$ with probability 1. 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*

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

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 tha A is not neacessarily 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

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

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

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

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,

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

Choose $\epsilon = n/2^n$, we see that $\mathsf{P}\{Y_n \leq n/2^n\} \to 0$, as $n \to \infty$. This implies that $\mathsf{P}\{B\} = \mathsf{P}\{\liminf_{n\to\infty}\{Y_n \leq n/2^n\}\} \leq \liminf_{n\to\infty} \mathsf{P}\{Y_n \leq n/2^n\} = 0$. 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 $s \leq t$. 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:

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

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. If X is a canonical process on the subspace $(\Gamma, \Gamma \cap \mathcal{E}^T)$ of the path space, then $\mathcal{F}_t^X = \gamma \cap \mathcal{E}^{[0,t]}$.

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 1.15). 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 $\mathcal{F}_\infty$.

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

Definition 1.6.4 A $[0,\infty]$-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 $\tau < \infty$ a.s., we call τ a *finite* stopping time.

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

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

(see Exercise 1.16). This should be viewed as the collection of all events that happen prior to the stopping time τ. Note that the notation causes no confusion, since a deterministic time $t \in T$ is clearly a stopping time and its associated σ-algebra is simply the σ-algebra $\mathcal{F}_t$.

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 1.22). For general filtrations we introduce the following class of random times.

Definition 1.6.5 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*.

Lemma 1.6.6 *τ is an optional time with respect to $(\mathcal{F}_t)$ if and only if it is a stopping time with respect to $(\mathcal{F}_{t+})$. Every stopping time is an optional time.*

Proof. See Exercise 1.23. QED

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.

Lemma 1.6.7 *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 the* first entrance time of B*, defined by*

$$\sigma_B = \inf\{t \geq 0 \,|\, X_t \in B\}$$

is an $(\mathcal{F}_t^X)$-stopping time.[1]

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

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

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 $s \leq t$. This means that

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

QED

Lemma 1.6.8 *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 an open set in E. Then the* hitting time of B*, defined by*

$$\tau_B = \inf\{t > 0 \,|\, X_t \in B\},$$

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

Proof. It holds that $\tau_B(\omega) < t$ if and only if there exist a sequence $t_n(\omega) \downarrow t$ such that $X_{t_n(\omega)}(\omega) \in B$. 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

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

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

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

Example 1.6.9 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 $x > 0$ and W is continuous, τ_x can be written as

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

By Lemma 1.6.7 this is an $(\mathcal{F}_t^W)$-stopping time. Next we will show that $\mathsf{P}\{\tau_x < \infty\} = 1$.

Note that $\{\tau_x < \infty\} = \cup_{n=1}^{\infty}\{\tau_x \leq n\}$ is a measurable set. Consider $A = \{\omega : \sup_{t\geq 0} W_t = \infty, \inf_{t\geq 0} W_t = -\infty\}$. 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 $W_{T_\omega} \geq T$, $W_{T'_\omega} \leq -T$. 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

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 1.6.10 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 1.6.11 *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 $t \geq 0$. We have to check that

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

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

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

This is a right-continuous process on $[0, t]$. It is also measurable, since

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

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.

Lemma 1.6.12 *Suppose that X is a progressively measurable process. Let τ be a finite stopping time. Then X_τ is an $\mathcal{F}_\tau$-measurable random variable.*

Proof. We have to show that $\{X_\tau \in B\} \cap \{\tau \le t\} \in \mathcal{F}_t$, for every $B \in \mathcal{E}$ and every $t \ge 0$. Now note that

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

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 $\{X_\tau \in B, \tau \le t\} \in \mathcal{F}_t$. Hence, it suffices to show that the map $\omega \to X_{\tau(\omega)\wedge t}(\omega)$ is $\mathcal{F}_t$-measurable. This map is the composition of the maps $\omega \to (\tau(\omega) \wedge t, \omega)$ from Ω to $[0,t] \times \Omega$, and $(s,\omega) \to X_s(\omega)$ from $[0,t] \times \Omega$ to E. The first map is measurable as a map from $(\Omega, \mathcal{F}_t)$ to $([0,t] \times \Omega, \mathcal{B}([0,t]) \times \mathcal{F}_t)$ (this is almost trivial, see Exercise 1.24). Since X is progressively measurable, the second map is measurable as a map from $([0,t] \times \Omega, \mathcal{B}([0,t]) \times \mathcal{F}_t)$ to $(E, \mathcal{E})$. This completes the proof, since the composition of measurable maps is measurable. QED

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

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

By Lemma 1.6.12 and Exercises 1.17 and 1.19, we have the following result.

Lemma 1.6.13 *If X is progressively measurable with respect to $(\mathcal{F}_t)$ and τ an $(\mathcal{F}_t)$-stopping time, then the stopped pocess X^τ is adapted to the filtrations $(\mathcal{F}_{\tau\wedge t})$ and $(\mathcal{F}_t)$.*

|n 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 finitely many values.

Lemma 1.6.14 *Let τ be a stopping time. Then there exist stopping times τ_n that only take finitely many values and such $\tau_n \downarrow \tau$.*

Proof. Define

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

Then τ_n is a stopping time and $\tau_n \downarrow \tau$ (see Exercise 1.25). QED

Using the notion of filtrations, we can extend the definition of BMas follows.

Definition 1.6.15 Suppose that on a probabillity space $(\Omega, \mathcal{F}, \mathsf{P})$ we have a filtration $(\mathcal{F}_t)_{t\ge 0}$ and an adapted stochastic process $W = (W_t)_{t\ge 0}$. 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) $W_t - W_s$ is independent of $\mathcal{F}_s$ for all $s \le t$;

iii) $W_t - W_s \stackrel{\mathsf{d}}{=} \mathcal{N}(0, t-s)$ 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 Complete the proof of Corollary 1.2.4. Give full details.

Exercise 1.2 Give an example of two processes that are versions of each other, but not modifications.

Exercise 1.3 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 1.4 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 $X_t = \mathbf{1}_{\{Y\leq t\}}$. Hence, $X_t(\omega) = \mathbf{1}_{\{Y(\omega)\leq t\}} = \mathbf{1}_{\{\omega\leq t\}}$.

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

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

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 $\Omega^* = \cap_{n\geq 1}\Omega_n$. Show that $\mathsf{P}\{\Omega^*\} = 1$. Then conclude that a continuous modification cannot exist.

Exercise 1.5 Suppose that X and Y are modifications of each other, and for both X and Y all sample paths are either left or right continuous. Show that

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

Exercise 1.6 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 1.7 Prove Lemma 1.4.2.

Exercise 1.8 Prove parts (i,ii,iii) of Theorem 1.4.4.

Exercise 1.9 Consider the proof of the time-inversion property of Theorem 1.4.4. Prove that

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

Exercise 1.10 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 1.11 Let W be a BM and fix $t > 0$. Define the process B by

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

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 1.12 **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 1.13 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 1.14 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 1.15 Show that if $(\mathcal{F}_t)$ is a filtration, then $(\mathcal{F}_{t+})$ is a filtration as well.

Exercise 1.16 Prove that the collection $\mathcal{F}_\tau$ associated with a stopping time τ is a σ-algebra.

Exercise 1.17 Show that if σ, τ are stopping times with $\sigma \leq \tau$, then $\mathcal{F}_\sigma \subseteq \mathcal{F}_\tau$.

Exercise 1.18 Let σ and τ be two $(\mathcal{F}_t)$-stopping times. Show that $\{\sigma \leq \tau\} \subset \mathcal{F}_\sigma \cap \mathcal{F}_\tau$.

Exercise 1.19 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 $A \cap \{\sigma \leq \tau\} \in \mathcal{F}_\tau$.

Exercise 1.20 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

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

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 1.21 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 $\mathbf{1}_{\{\tau=\sigma\}}\mathsf{E}(X\,|\,\mathcal{F}_\tau) \overset{\text{a.s.}}{=} \mathbf{1}_{\{\tau=\sigma\}}\mathsf{E}(X\,|\,\mathcal{F}_\sigma)$. Hint: show that $\mathbf{1}_{\{\tau=\sigma\}}\mathsf{E}(X\,|\,\mathcal{F}_\tau) = \mathbf{1}_{\{\tau=\sigma\}}\mathsf{E}(X\,|\,\mathcal{F}_\tau\cap\mathcal{F}_\sigma)$.

Exercise 1.22 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 1.23 Prove Lemma 1.6.6.

Exercise 1.24 Show that the map $\omega \to (\tau(\omega)\wedge t, \omega)$ in the proof of Lemma 1.6.12 is measurable as a map from $(\Omega, \mathcal{F}_t)$ to $([0,t]\times\Omega, \mathcal{B}([0,t])\times\mathcal{F}_t))$.

Exercise 1.25 Show that τ_n in the proof of Lemma 1.6.14 are indeed stopping times and that they converge to τ.

Exercise 1.26 Translate the definitions of §1.6 to the special case that time is discrete, i.e. $T = \mathbf{Z}_+$.

Exercise 1.27 Let W be a BM and let $Z = \{t \geq 0\,|\,W_t = 0\}$ be its zero set. Show that with probability 1 the set Z has Lebesgue measure 0, is closed and unbounded.

Exercise 1.28 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.

ii) Show that L_x is measurable for all x. Derive first that $\{L_x < t\} = \cap_{n>t}\{|W_s - x| > 0, t \leq s \leq n\}$.

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 1.6.9.

iv) Show that for almost all $\omega \in \Omega$ there exists a 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 1.29 Consider Brownian motion W, 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 1.28 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 1.30 Let W be a BM 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 1.31 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.

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) $\mathsf{E}(M_t \mid \mathcal{F}_s) \stackrel{\text{a.s.}}{=} M_s$ for all $s \leq t$.

If property (ii) holds with '$\geq$' (resp. '$\leq$') 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 $t \geq s$ 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') $\mathsf{E}\{M_{n+1} \mid \mathcal{F}_n\} = M_n(\geq, \leq)$ a.s. for $n \in \mathbf{Z}_+$.

Example 2.1.2 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 2.1.3 (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 2.1.4 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 2.1.5 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 2.1.6 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) $\mathbf{Z}_+$. Note that as a consequence, it only makes sense to consider $\overline{\mathbf{Z}}_+$-valued stopping time. In discrete time, τ is a stopping time with respect to the filtration $(\mathcal{F}_n)_{n\in\mathbf{Z}_+}$, if $\{\tau \le n\} \in \mathcal{F}_n$ 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 a 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 (sub-, super-)martingale and τ a stopping time. Then the stopped process M^τ is a (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 stopping theorem.* The general version will be discussed in section 2.2.5.

Theorem 2.2.5 *Let M be a submartingale 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 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 1.6.12).

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 for an 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

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

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

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

Now, note that $M_K^\tau = M_K$ 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 $\tau = n \wedge \inf\{k \mid M_k \geq \lambda\}$. This is a stopping time (see Lemma 1.6.7) 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 $n \in \mathbf{N}$*

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

provided M is in L^p.

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

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

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 $\mathsf{P}\{M^* \geq x\}$. Thus

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

Insert this in (2.2.6), then

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

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

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

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

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

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

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

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

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

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

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

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

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

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). Taking expectations in the latter expression shows that $\mathsf{E}A_n = \mathsf{E}X_n - \mathsf{E}X_0 \leq \mathsf{E}|X_n| + \mathsf{E}|X_0|$. 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 2.2.10 The number $U_n[a, b]$ is the largest value $k \in \mathbf{Z}_+$, such that there exis $0 \leq s_1 < t_1 < s_2 < \cdots < s_k < t_k \leq n$ 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 2.2.11 (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*

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

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

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

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

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

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 (2.22). 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 2.2.12 (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

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

Hence $\mathsf{P}\{A\} \leq \sum_{\substack{a,b \in \mathbb{Q}: \\ a<b}} \mathsf{P}\{U_\infty[a,b] = \infty\} = 0$. 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 2.2.13 *Let M be a supermartingale that is bounded in L^1. Then $M_n \xrightarrow{\mathsf{L}^1} M_\infty$, $n \to \infty$, 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*

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

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 2.2.12 $M_n \to M_\infty$ 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

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

Since $|\mathbf{1}_{\{A\}} M_m - \mathbf{1}_{\{A\}} M_\infty| \leq \mathbf{1}_{\{A\}} |M_m - M_\infty| \leq |M_m - M_\infty|$, 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

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

This implies (see Exercise 2.7(a)) that $\mathsf{E}(M_\infty \,|\, \mathcal{F}_n) \leq M_n$ 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 2.1.3) $\mathsf{E}(X \,|\, \mathcal{F}_n)$ 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 2.2.14 (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 $n \to \infty$*

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

a.s. and in L^1.

Proof. The process $M_n = \mathsf{E}(X \,|\, \mathcal{F}_n)$ is uniformly integrable (see Example 2.1.3), hence bounded in L^1 (explain!). By Theorem 2.2.13 $M_n \to M_\infty$ a.s. and in L^1, as $n \to \infty$ with M_∞ integrable and $\mathcal{F}_\infty$-measurable. It remains to show that $M_\infty = \mathsf{E}(X \,|\, \mathcal{F}_\infty)$ 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 2.2.13 for the first equality and the definition of M_n for the second. First assume that $X \geq 0$, then $M_n = \mathsf{E}(X \,|\, \mathcal{F}_n) \geq 0$ a.s. (see BN Lemma 7.2 (iv)), hence $M_\infty \geq 0$ 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 $M_\infty = \mathsf{E}(X \,|\, \mathcal{F}_\infty)$ 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 2.2.15 (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 2.2.12, 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 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 $M^- = \max\{-M, 0\}$. It is an non-decreasing, convex function of the submartingale $-M$, whence it is a submartingale (see Example 2.1.4). 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 $M_{-n} = \mathsf{E}(M_{-1} \mid \mathcal{F}_{-n})$ 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 2.2.14 as a special case. Indeed, if X is an integrable random variable and $\mathcal{F}_1 \supseteq \mathcal{F}_2 \supseteq \cdots \subseteq \cap_n \mathcal{F}_n = \mathcal{F}_\infty$ is a decreasing sequence of σ-algebras, then

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

a.s. and in L^1. This is generalised in the following corollary to Theorems 2.2.14 and 2.2.15. It will be useful in the sequel.

Corollary 2.2.16 *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 $\mathsf{E}(X_n \mid \mathcal{F}_n) \to \mathsf{E}(X \mid \mathcal{F}_\infty)$ 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 $\mathsf{E}(U_m \mid \mathcal{F}_n) \le \mathsf{E}(X_n \mid \mathcal{F}_n) \le \mathsf{E}(V_m \mid \mathcal{F}_n)$ 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

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

It follows that

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

Letting $\epsilon \downarrow 0$ yields that $\limsup \mathsf{E}(X_n \mid \mathcal{F}_n) = \liminf \mathsf{E}(X_n \mid \mathcal{F}_n)$ a.s. and so $\mathsf{E}(X_n \mid \mathcal{F}_n)$ 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

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

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

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

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 stopping theorems

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

Theorem 2.2.17 (Optional sampling theorem) *Let M be a uniformly integrable (super)martingale. Then the family of random variables $\{M_\tau \mid \tau \text{ is a finite stopping time}\}$ is uniformly integrable and for all stopping times $\sigma \leq \tau$ we have*

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

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

By Theorem 2.2.13, $M_\infty = \lim_{n\to\infty} M_n$ 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 $\tau \wedge n \leq n$, $\mathcal{F}_{\tau\wedge n} \subseteq \mathcal{F}_n$. 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 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 2.26)! For nonnegative supermartingales without additional integrability properties we only have an inequality.

Theorem 2.2.18 *Let M be a nonnegative supermartingale and let $\sigma \le \tau$ be stopping times. Then*

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

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

$$\mathsf{E}(M_{\tau\wedge n} \,|\, \mathcal{F}_\sigma) = \mathsf{E}(M_\infty^{\tau\wedge n} \,|\, \mathcal{F}_\sigma) \le M_\sigma^{\tau\wedge n} = M_{\sigma\wedge n} \quad \text{a.s.}$$

Since the limit exists, we have $M_\tau \mathbf{1}_{\{\tau=\infty\}} = M_\infty \mathbf{1}_{\{\tau=\infty\}}$. By conditional Fatou

$$\begin{aligned}
\mathsf{E}(M_\tau \,|\, \mathcal{F}_\sigma) &\le \mathsf{E}(\liminf M_{\tau\wedge n} \,|\, \mathcal{F}_\sigma) \\
&\le \liminf \mathsf{E}(M_{\tau\wedge n} \,|\, \mathcal{F}_\sigma) \\
&\le \liminf M_{\sigma\wedge n} = M_\sigma, \quad \text{a.s.}
\end{aligned}$$

This proves the result. QED

2.3 Continuous-time martingales

In this section we consider general martingales indexed by a subset T of $\mathbf{R}_+$. If the martingale $M = (M_t)_{t\ge 0}$ 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 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 2.2.10). 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 countable. 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 2.1.4). 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 2.2.11 $U_T[a,b]$ 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 $\Omega^* \in \mathcal{F}_\infty$ of probability 1, such that for all $\omega \in \Omega^*$ the limits*

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

exist and are finite for every $t \geq 0$.

The set of discontinuities of each path $M(\omega) = (M_t(\omega))_{t\in\mathrm{R}_+}$, $\omega \in \Omega^*$, is at most countable. This is a result from analysis. **This has to be checked still!**

Proof. 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 2.2.9

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

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 2.21. 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 $t \geq 0$

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

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 2.2.15, with $\sup_n \mathsf{E}M_{q_n} \leq \mathsf{E}M_t < \infty$. By that theorem, M_{t+} is integrable, and $M_{q_n} \xrightarrow{\mathsf{L}^1} M_{t+}$. As in the proof of Theorem 2.2.13, 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 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 2.2.15. Hence, $\mathsf{E}(M_{t+} \mid \mathcal{F}_{q'_n})$ converges to $\mathsf{E}(M_{t+} \mid \mathcal{F}_{t+})$ 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 §10.

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 $\{M_t - \widetilde{M}_t \neq 0\} \in \mathcal{F}_0$) 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 2.22).

'⇐' see Exercise 2.23. 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.*

The next example shows what can go wrong without right-continuity of the filtration.

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_t is a martingale, but it is not right-continuous, whereas X_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_t is not a cadlag modification of Y_t, and in particular Y_t 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_t cannot be a martingale w.r.t the filtration $\mathcal{F}_t$. By the same lemma, X_t is a right-continuous martingale, w.r.t to $\mathcal{F}_{t^+}$.

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 $|M_t(\omega) - M_\infty(\omega)| \leq |M_{q'}t(\omega) - M_\infty(\omega)| + |M_t(\omega) - M_{q'}(\omega)| < 2\epsilon$. 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 2.2.12, 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 2.2.13 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*

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

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 2.24. 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 $T_n \subseteq T$, such that $t \in T_n$ 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

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

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 stopping

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*

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

with equality if M is martingale.

Proof. By Lemma 1.6.14 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!)

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

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

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

Similarly

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

Hence, $(M_{\tau_n})_n$ is a 'backward' supermartingale in the sense of the Lévy-Doob downward theorem 2.2.15. 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

Corollary 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 1.6.14 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 2.2.15, 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 supermartingale and τ is a stopping time, then the stopped process is a supermartingale as well.*

Proof. Note that M^τ is right-continuous. By Lemmas 1.6.11 and 1.6.13 it is adapted.

Let $\sigma \leq \xi$ be bounded stopping times. By applying the previous corollary 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 corollary 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 2.2.18 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 a.s. M vanishes on $[\tau, \infty]$.*

Proof. Positive supermartingales are bounded in L^1. By Theorem 2.3.7, M converges a.s. to an integable limit M_∞ say. Note that 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-negative of M we get that $M_{\tau+q} = 0$ a.s. for all $q \in \mathsf{Q}_+$. But then also $\mathsf{P}\{\cap_q\{M_{\tau+q} = 0\}\} = 1$. By right continuity

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

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 1.14 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

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

Proof. For the first statement see Exercise 1.14 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 $\mathsf{E}(X_n \,|\, \mathcal{F}_{n+1}) = X_{n+1}$ a.s. Then, since $\sup \mathsf{E} X_n < \infty$, the Lévy-Doob downward theorem 2.2.15 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 $\mathsf{E}(X_n \,|\, \mathcal{F}_{n+1}) = X_{n+1}$ 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 1.11 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.

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

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 $0 = t_0^n \leq t_1^n \leq \cdots \leq t_{k_n}^n = t$. Suppose that the mesh $\|\pi_n\| = \max_k |t_k^n - t_{k-1}^n| \to 0$ as $n \to \infty$. Then

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

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 equality

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 $t \geq 0$ and $\lambda > 0$*

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

and

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

Proof. For $a > 0$ consider the exponential martingale M defined by $M_t = \exp\{aW_t - a^2t/2\}$ (see Example 2.1.5). Observe that

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

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

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

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

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

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 $\sup_{s\leq t} W_s \stackrel{\mathsf{d}}{=} |W_t|$ 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 $a > 0$*

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

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 $h(t) = \sqrt{2t \log\log 1/t}$ and 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 $M_s = \exp\{\alpha_n W_s - \alpha_n^2 s/2\}$:

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

for some constant $K > 0$ that does not depend on n. Applying the Borel-Cantelli lemma yields that

$$\sup_{s\le 1}(W_s - \alpha_n s/2) \le \beta_n,$$

for all sufficiently large n. In particular, for all sufficiently large n and $s \in [0, \theta^{n-1}]$

$$W_s \le \frac{\alpha_n s}{2} + \beta_n \le \frac{\alpha_n \theta^{n-1}}{2} + \beta_n = \Big(\frac{1+\delta}{2\theta} + \frac{1}{2}\Big) h(\theta^n).$$

Since h is increasing in a neighbourhood of 0, it follows that for all n large enough

$$\sup_{\theta^n \le s \le \theta^{n-1}} \frac{W_s}{h(s)} \le \Big(\frac{1+\delta}{2\theta} + \frac{1}{2}\Big).$$

Let $\theta \uparrow 1$ and $\delta \downarrow 0$ to find that

$$\limsup_{t \downarrow 0} \frac{W_t}{h(t)} \le 1.$$

QED

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

$$\sqrt{2\pi}\mathsf{P}\{A_n\} \ge \frac{a}{1+a^2} e^{-a^2/2}.$$

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 Borel-Cantelli Lemma

$$W_{\theta_n} \ge (1-\sqrt{\theta})h(\theta^n) + W_{\theta^{n+1}}$$

for infinitely many n. Since $-W$ is also a BM, the first part of the proof implies that

$$-W_{\theta^{n+1}} \le 2h(\theta^{n+1}),$$

for all n large enough. Note that

$$\log\log\theta^{-(n+1)} = (n+1)\log\log\theta^{-1} \le 2n\log\log\theta^{-1} = 2\log\log\theta^{-n},$$

we find that

$$h(\theta^{n+1}) \le \sqrt{2}\theta^{(n+1)/2} 2\log\log\theta^{-n} = 2\sqrt{\theta}h(\theta^n).$$

Combining this with the preceding inequality yields that

$$W_{\theta_n} \geq (1-\sqrt{\theta})h(\theta^n) - 2h(\theta^{n+1}) \geq h(\theta^n)(1-5\sqrt{\theta}),$$

for inifinitely many n. Hence

$$\limsup_{t\downarrow 0} \frac{W_t}{h(t)} \geq 1 - 5\sqrt{5}.$$

The proof is completed by letting θ tend to zero. QED

As a corollary we have the following result regarding the zero set of the BM that was considered in Exercise 1.26 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 2.1.5. 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 λ 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 $\omega_t \le \liminf_{s\uparrow t}\omega_s + 1$, 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-alagebra $\mathcal{F}$ is the associated σ-algebra that makes the projections on the t-coordinate measurable.

As we have seen in Example 2.1.6 , $\{N(t) - \lambda t\}_t$ 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

$$\frac{N(t)}{t} \stackrel{\text{a.s.}}{\to} \lambda$$

Proof. see Exercise 2.33. 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.

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

$$\int_0^t f(s)dN(s,\omega) = \sum_{n=0}^{N(t,\omega)} f(S_n(\omega)) = \sum_{n=0}^{\infty} f(S_n(\omega))\mathbf{1}_{\{S_n(\omega)\leq t\}},$$

where S_n are the successive discontinuities.

The PASTA property Let us now consider some $(E,\mathcal{E})$-valued stochastic process X on $(\Omega,\mathcal{F},\mathsf{P})$, 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: $U(t) = \mathbf{1}_{\{X(t)\in B\}}$.

Assumption A U has ladcag paths, i.o.w. all paths of U are left continuous and have right limits.

We further define

$$\begin{aligned}
\overline{U}(t) &= \frac{1}{t}\int_0^t U(s)ds \\
A(t) &= \int_0^t U(s)dN(s) \\
\bar{A}(t) &= \frac{A(t)}{N(t)}\mathbf{1}_{\{N(t)>0\}} \\
\mathcal{F}_t &= \sigma(U(s), s \leq t, N(s), s \leq t).
\end{aligned}$$

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 $\mathsf{E}A(t) = \lambda t \mathsf{E}\overline{U}(t) = \lambda \mathsf{E}\int_0^t U(s)ds$.

Proof. As in the proof of BN Lemma 3.3 the ladcag assumption implies that $U(s,\omega)$, $s \leq t$ has finitely many discontinuities. Together with the sample properties of N, this implies that we can approximate $A(t)$ by

$$A_n(t) = \sum_{k=0}^{n-1} U(\tfrac{kt}{n})[N(\tfrac{(k+1)t}{n}) - N(\tfrac{kt}{n})].$$

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

$$\mathsf{E}A(t) = \lim_{n\to\infty} \mathsf{E}A_n(t) = \lim_{n\to\infty} \frac{\lambda t}{n} \sum_{k=0}^{n-1} \mathsf{E}U(\tfrac{t}{n}) = \mathsf{E} \lim_{n\to\infty} \frac{\lambda t}{n} \sum_{k=0}^{n-1} U(\tfrac{t}{n}) = \lambda \mathsf{E} \int_0^t U(s)ds.$$

The interchange argument in the third step is also validated by dominated convergence. QED

Corollary 2.5.4 *Suppose that assumptions A and B hold.* $\mathsf{E}(A(t)-A(s) \,|\, \mathcal{F}_s) = \lambda \mathsf{E}(\int_s^t U(v)dv \,|\, \mathcal{F}_s)$ *a.s.*

Proof. The lemma implies that $\mathsf{E}(A(t) - A(s)) = \lambda \mathsf{E} \int_s^t U(v)dv$. Define

$$A_n(s,t) = \sum_{ns/t \leq k \leq n-1} U(\tfrac{kt}{n})[N(\tfrac{(k+1)t}{n}) - N(\tfrac{kt}{n})].$$

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 $\mathsf{E}(A_n(s,t) \,|\, \mathcal{F}_s) \to \mathsf{E}(A(t) - A(s) \,|\, \mathcal{F}_s)$ a.s. On the other hand

$$\mathsf{E}(A_n(s,t) \,|\, \mathcal{F}_s) = \mathsf{E}(\frac{\lambda t}{n} \sum_{ns/t \leq k \leq n-1} U(\tfrac{kt}{n}) \,|\, \mathcal{F}_s).$$

By another application of conditional dominated convergence, using boundedness of the function involved, the right-hand side converges a.s. to $\mathsf{E}(\lambda \int_s^t U(v)dv \,|\, \mathcal{F}_s)$. QED

Next define $R(t) = A(t) - \lambda t\overline{U}t)$. 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 BN Theorem 9.3

$$\frac{R(nh)}{n} \to 0,$$

on the set

$$A = \{\sum_{k=1}^{\infty} \frac{1}{k^2} \mathsf{E}((R((k+1)h) - R(kh))^2 \,|\, \mathcal{F}_{kh}) < \infty\}.$$

Note that

$$|R(t) - R(s)| \leq N(t) - N(s) + \lambda(t-s). \tag{2.5.1}$$

It follows that

$$\mathsf{E}(R(t) - R(s))^2 \leq \mathsf{E}(N(t) - N(s))^2 + 2\lambda^2(t-s)^2 = 3\lambda^2(t-s)^2 + \lambda(t-s). \tag{2.5.2}$$

Consider the random variables $Y_n = \sum_{k=1}^n \frac{1}{k^2} \mathsf{E}((R((k+1)h) - R(kh))^2 \,|\, \mathcal{F}_{kh})$, $n = 1, \ldots$. By (2.5.2),

$$\mathsf{E}|Y_n| = \mathsf{E}Y_n \leq \sum_{k=1}^n (3\lambda^2 h^2 + \lambda h) \frac{1}{k^2} < \infty,$$

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$N(t)/t \to \lambda$. 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 $\omega \in A$. Fix $t > n_\omega h$ and let n_t be such that $t \in [n_t h, (n_t + 1)h)$. By virtue of(2.5.1)

$$|R(t, \omega) - R(n_t h, \omega)| \leq N(t, \omega) - N(n_t h, \omega) + \lambda h.$$

By another application of Lemma 2.5.1

$$\frac{R(t, \omega)}{t} \leq \frac{n_t h}{t} \cdot \frac{R(n_t h, \omega) + N(t, \omega) - N(n_t h, \omega) + \lambda h}{n_t h} \to 0, \to \infty.$$

QED

Now we can finish the proof of the theorem. It follows from the relation

$$\frac{R(t)}{t} = \frac{A(t)}{N(t)} \frac{N(t)}{t} \mathbf{1}_{\{N(t)>0\}} - \lambda \overline{U}(t),$$

whilst noting that $(N(t)/t)\mathbf{1}_{\{N(t)>0\}} \to \lambda$ on Ω^*.

2.6 Exercises

Discrete-time martingales

Exercise 2.1 Prove the assertion in Example 2.1.3.

Exercise 2.2 Prove the assertion in Example 2.1.4.

Exercise 2.3 Show that the processes defined in Example 2.1.5 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 2.2.17. 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.

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 2.2.17 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 $\mathsf{E}\log(Z_n/Z_0) \le N\alpha$. 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,

$$T = \min\{n \mid n \ge 10, U_{n-9}U_{n-8}\cdots U_n = \mathbf{1231231231}\}.$$

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 2.20A 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 $\{-1, 0, 1\}$. 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 $\tau_y = \min\{n > 0 \,|\, X_n = y\}$ and let $\tau = \tau_a \wedge \tau_b$. We want to compute $\mathsf{P}_x\{\tau_b < \tau_a\}$ and $\mathsf{E}_x\{\tau\}$. 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 $f(x) = e^{\alpha x}$ 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

$$\mathsf{P}_x\{\tau_b < \tau_a\} = \frac{e^{\alpha x} - e^{\alpha a}}{e^{\alpha b} - e^{\alpha a}}.$$

If $\mathsf{P}\{X_i = 1\} = p = 1 - \mathsf{P}\{X_i = -1\}$ (that is: X_i takes only values ± 1), then

$$e^{\alpha} = \frac{1-p}{p} \quad \text{or } \alpha = \log(1-p) - \log p.$$

Show this.

iii) Show that $S_n - n\mathsf{E}X_1$ is martingale. Show that

$$\mathsf{E}_x\tau = \frac{(e^{\alpha x} - e^{\alpha b})(x-a) + (e^{\alpha x} - e^{\alpha a})(b-x)}{(e^{\alpha b} - e^{\alpha a})\mathsf{E}X_1}.$$

iv) Let now $\mathsf{E}X_1 = 0$. Show that $\tau < \infty$ P_x-a.s. (hint: use the Central Limit Theorem). Show for $x \in (a, b)$ that

$$\mathsf{P}_x\{\tau_a < \tau_b\} = \frac{b-x}{b-a}$$

and

$$\mathsf{E}_x\tau = \frac{(x-a)(b-x)}{\mathsf{E}X_1^2},$$

by constructing suitable martingales.

Continuous-time martingales

Exercise 2.21 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 2.22 Show that the process constructed in the proof of Theorem 2.3.5 is cadlag and a supermartingale, if M is a supermartingale.

Exercise 2.23 Prove the 'only if' part of Theorem 2.3.5.

Exercise 2.24 Prove Theorem 2.3.9 from LN. You may use Theorem 2.3.7.

Exercise 2.25 Show that for every $a \neq 0$, the exponential martingale of Example 2.1.5 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 2.26 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 2.25.

Exercise 2.27 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 abvoe the level x).

b) Let W be a standard BM. Using the exponential martingales of Example 2.1.5, 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 2.28 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 2.29 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 2.30 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 $M_t = \cosh(b|W_t|)\exp\{b^2t/2)$ is a martingale.

b) Find the Laplace transform of the stopping time $\tilde{\tau}_a$.

c) Calculate $\mathsf{E}\tilde{\tau}_a$.

Exercise 2.31 (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) Explain or prove that for $0 \leq t \leq u < 1$ 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) Set $M_n(t) = 1$ and put $\mathcal{G}_n(t) = \sigma(X_1, \ldots, X_n)$, for $t \geq 1$. 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 2.32 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 $x > 0$

$$\mathsf{P}\{H_x < \infty\} = \begin{cases} 1, & c \geq 0 \\ e^{-2|c|x}, & c < 0. \end{cases}$$

v) Explain why this result is reasonable.

Exercise 2.33 Let $N = \{N(t)\}_{t\geq 0}$ be a Poisson process (see Definition in Ch1.1). Show that $\{N(t) - \lambda t\}_{t\geq 0}$ is a martingale. Then prove Lemma 2.5.1. Hint: use the martingale LLN given in BN §9.

Chapter 3

Markov Processes

3.1 Basic definitions

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: $\mathcal{E} = \sigma\{\{x\} \,|\, x \in E\}$. Let $X = \{X_n\}_{n=0,1,\ldots}$ 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 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, 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 $n = 0, 1, \ldots, x_0, \ldots, x_n, x_{n+1} \in E$

$$\mathsf{P}_\nu\{X_{n+1} = x_{n+1} \,|\, X_0 = x_0, \ldots, X_n = x_n\} = \mathsf{P}_\nu\{X_{n+1} = x_{n+1} \,|\, X_n = x_n\} = P(x_n\,, x_{n+1}).$$

Recall that

$$\mathsf{P}_\nu\{X_{n+1} = x_{n+1} \,|\, \sigma(X_n)\} = \mathsf{E}_\nu\{\mathbf{1}_{\{x_{n+1}\}}(X_{n+1}) \,|\, \sigma(X_n)\}$$

is a function of X_n. In the case at hand this is $P(X_n, x_{n+1})$. Then $\mathsf{P}_\nu\{X_{n+1} = x_{n+1} \,|\, X_n = x_n\} = P(x_n, x_{n+1})$ is simply the evaluation of that function at the point $X_n = x_n$. These conditional probabilities can be computed by

$$\mathsf{P}_\nu\{X_{n+1} = x_{n+1} \,|\, X_n = x_n\} = \begin{cases} \dfrac{\mathsf{P}_\nu\{X_{n+1} = x_{n+1}, X_n = x_n\}}{\mathsf{P}\{X_n = x_n\}}, & \text{if } \mathsf{P}_\nu\{X_n = x_n\} > 0 \\ \text{anything you like} & \text{if } \mathsf{P}_\nu\{X_n = x_n\} = 0 \end{cases}$$

We can now rephrase the Markov property as follows: for all $n \in \mathbf{Z}_+$ and $y \in E$

$$\mathsf{P}_\nu\{X_{n+1} = y \,|\, \mathcal{F}_n^X\} = P(X_n, y), \quad \text{a.s.}$$

It is a straightforward computation that

$$\mathsf{P}_\nu\{X_{n+m} = y \,|\, \mathcal{F}_n^X\} = P^m(X_n, y),$$

is the (X_n, y)-th element of the m-th power of P. Indeed, for $m = 2$

$$\begin{aligned}
\mathsf{P}_\nu\{X_{n+2} = y \,|\, \mathcal{F}_n^X\} &= \mathsf{E}_\nu(\mathbf{1}_{\{y\}}(X_{n+2}) \,|\, \mathcal{F}_n^X) \\
&= \mathsf{E}_\nu(\mathsf{E}_\nu \mathbf{1}_{\{y\}}(X_{n+2}) \,|\, \mathcal{F}_{n+1}^X) \,|\, \mathcal{F}_n^X) \\
&= \mathsf{E}_\nu(P(X_{n+1}, y) \,|\, \mathcal{F}_n^X\} \qquad (3.1.1) \\
&= \mathsf{E}_\nu(\sum_{x \in E} \mathbf{1}_{\{x\}}(X_{n+1}), P(X_{n+1}, y) \,|\, \mathcal{F}_n^X\} \\
&= \sum_{x \in E} P(x, y) \mathsf{E}_\nu(\mathbf{1}_{\{x\}}(X_{n+1}) \,|\, \mathcal{F}_n^X\} \\
&= \sum_{x \in E} P(x, y) P(X_n, x) = P^2(X_n, y).
\end{aligned}$$

In step (3.1.1) to the next, we use discreteness of the state space as well linearity of conditions 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.

The one point sets now need not be measurable. The notion of stochastic matrix now 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 $x \mapsto P(x, B)$ 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 wth 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:

$$\mathsf{P}\{X_{n+1} \in B \,|\, \mathcal{F}_n^X\} = P(X_n, B), \quad B \in \mathcal{E}, n = 0, 1, 2 \ldots. \qquad (3.1.2)$$

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

$$\mathsf{P}\{X_{n+m} \in B \,|\, \mathcal{F}_n^X\} = P^m(X_n, B), \quad B \in \mathcal{E}, n = 0, 1, 2 \ldots,$$

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.(3.1.1))

$$\mathsf{P}\{X_{n+2} \in B \mid \mathcal{F}_n^X\} = \mathsf{E}(\mathsf{E}(\mathbf{1}_{\{B\}}(X_{n+2}) \mid \mathcal{F}_{n+1}^X) \mid \mathcal{F}_n^X\} = \mathsf{E}(P(X_{n+1}, B) \mid \mathcal{F}_n^X).$$

Can we apply the Markov property to the latter expression? We need to recast the Markov property (3.1.2) in terms of functions. Let us introduce some 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 $f : E \to \mathbf{R}$.

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

> Look up in BN section 3 Measurability what we mean by the 'standard machinery'.

yields that $Pf \in b\mathcal{E}$ for all $f \in b\mathcal{E}$. In other words P is a linear operator mapping $b\mathcal{E}$ to $b\mathcal{E}$.

The Markov property (3.1.2) 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.

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

$$P_{t+s}f = P_t(P_s f) = P_s(P_t f).$$

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.

Some further notation is enlightening. Let f, g, h be bounded (non-negative) measurable functions on E. As argued before $P_t f$ is bounded, measurable. Hence multiplying by g gives $gP_t f$, which is bounded, measurable. Here

$$gP_t f(x) = g(x) \cdot P_t f(x) = g(x) \int_y f(y) P_t(x, dy).$$

Then we can apply P_s to this function, yielding the bounded, measurable function $P_s g P_t f$, with

$$P_s g P_t f(x) = \int_y g(y) P_t f(y) P_s(x, dy) = \int_y g(y) \int_z f(z) P_t(y, dz) P_s(x, dy).$$

$hP_s g P_t f$ is again bounded, measurable and we can integrate over the probability distribution ν on $(E, \mathcal{E})$:

$$\begin{aligned} \nu h P_s g P_t f &= \int_x h(x) P_s g P_t f(x) \nu(dx) \\ &= \int_x h(x) \int_y g(y) \int_z f(z) P_t(y, dz) P_s(x, dy) \nu(dx). \end{aligned}$$

We can now give the definition of a Markov process.

Definition 3.1.3 Let $(E, \mathcal{E})$ be a measure 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 every bounded, measurable function $f : E \to \mathbf{R}$

$$\mathsf{E}_\nu(f(X_{t+s}) \mid \mathcal{F}_s) = P_t f(X_s) \quad \mathsf{P}_\nu - \text{a.s.} \tag{3.1.3}$$

To remind we get the following alternative notations that will be interchangedly used

$$\mathsf{E}_x \mathbf{1}_{\{A\}}(X_s) = \mathsf{P}_x\{X_s \in A\} = P_s(y, A).$$

Note that by the 'standard machinery', we may replace bounded f by non-negative f in the definition.

Definition 3.1.4 Let $(E,\mathcal{E})$ be a measure space and let $X : (\Omega,\mathcal{F}) \to (E^{\mathbb{R}_+},\mathcal{E}^{\mathbb{R}_+})$ be a map that is adapted to the filtration $(\mathcal{F}_t)_t$, $\mathcal{F}_t \subset \mathcal{F}$, $t \geq 0$. X is a **Markov process**, if for each distribution ν on $(E,\mathcal{E})$ there exists a probability distribution P_ν on $(\Omega,\mathcal{F})$

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 every bounded, measurable function $f : E \to \mathbf{R}$

$$\mathsf{E}_\nu(f(X_{t+s}) \,|\, \mathcal{F}_s) = P_t f(X_s) \quad \mathsf{P}_\nu - \text{a.s.} \tag{3.1.4}$$

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. (still have to give an example).

Lemma 3.1.5 *Let X be an $(E,\mathcal{E})$-valued stochastic process with transition function $(P_t)_{t\geq 0}$. Let ν be a distribution on $(E,\mathcal{E})$.*

Then X is Markov with initial distribution ν, with respect to its natural filtration $(\mathcal{F}_t^X)_t$ if and only if for all initial distributions ν, all $0 = t_0 < t_1 < \cdots < t_n$, and all bounded measurable functions $f_0,\ldots,f_n$ on E, $n \in \mathbf{Z}_+$,

$$\mathsf{E}_\nu \prod_{i=0}^n f_i(X_{t_i}) = \nu f_0 P_{t_1-t_0} f_1 \cdots P_{t_n-t_{n-1}} f_n. \tag{3.1.5}$$

In either case, (3.1.5) also holds for non-negative measurable functions $f_0,\ldots,f_n$.

Remark the proof of the Lemma shows that is it sufficient to check (3.1.5) for indicator functions.

Proof. Let X be a Markov process with initial distribution ν, with respect to its natural filtration. Then

$$\begin{aligned}
\mathsf{E}_\nu \prod_{i=0}^n f_i(X_{t_i}) &= \mathsf{E}_\nu \mathsf{E}_\nu \Big(\prod_{i=0}^n f_i(X_{t_i}) \,|\, \mathcal{F}^X_{t_{n-1}}\Big) \\
&= \mathsf{E}_\nu \prod_{i=0}^{n-1} f_i(X_{t_i}) \mathsf{E}_\nu(f(X_{t_n}) \,|\, \mathcal{F}^X_{t_{n-1}}) \\
&= \mathsf{E}_\nu \prod_{i=0}^{n-1} f_i(X_{t_i})\, P_{t_n-t_{n-1}} f_n(X_{t_{n-1}}).
\end{aligned}$$

Now, $P_{t_n-t_{n-1}} f_n$ is a bounded, measurable function, and so one has

$$\begin{aligned}
\mathsf{E}_\nu \prod_{i=0}^{n-1} f_i(X_{t_i})\, P_{t_n-t_{n-1}} f_n(X_{t_{n-1}}) &= \mathsf{E}_\nu \prod_{i=0}^{n-2} f_i(X_{t_i}) \mathsf{E}_\nu(f_{n-1}(X_{t_{n-1}})\, P_{t_n-t_{n-1}} f_n(X_{t_{n-1}}) \,|\, \mathcal{F}_{t_{n-2}}) \\
&= \mathsf{E}_\nu \prod_{i=0}^{n-2} f_i(X_{t_i})\, P_{t_{n-1}-t_{n-2}} f_{n-1}\, P_{t_n-t_{n-1}} f_n(X_{t_{n-2}}).
\end{aligned}$$

Interating 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 (3.1.5) holds for all $0 = t_0 < t_1 < \cdots < t_n$, all bounded measurable functions $f_0, \ldots, f_n$. 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 $A \in \mathcal{F}_s^X$

$$\mathsf{E}_\nu \mathbf{1}_{\{A\}} f(X_{t+s}) = \mathsf{E}_\nu \mathbf{1}_{\{A\}}\, P_t f(X_s). \tag{3.1.6}$$

Let $B \in \mathcal{E}$, put $n = 0$, $f_0 = \mathbf{1}_{\{B\}}$. (i) immediately follows.

We will show (ii). To derive (3.1.6), it is sufficient to check this for a π-system generating $\mathcal{F}_s$. As the π-system we take

$$\Big\{A = \{X_{t_0} \in A_0, X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n\} \,|\, t_0 = 0 < t_1 < \cdots < t_n \leq s, A_i \in \mathcal{F}_{t_i}, n = 1, \ldots\Big\}$$

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 $t_n < s$

$$\begin{aligned}&\mathsf{E}_\nu \prod_{i=0}^n \mathbf{1}_{\{A_i\}}(X_{t_i}) \mathbf{1}_{\{E\}}(X_s) f(X_{t+s}) \\ &\quad= \nu \mathbf{1}_{\{A_0\}}\, P_{t_1-t_0} \mathbf{1}_{\{A_1\}}\, P_{t_2-t-1} \cdots P_{t_n-t_{n-1}} \mathbf{1}_{\{A_n\}}\, P_{t+s-t_n} f \\ &\quad= \nu \mathbf{1}_{\{A_0\}}\, P_{t_1-t_0} \mathbf{1}_{\{A_1\}} \cdots P_{s-t_n}(\, P_t f) \\ &\quad= \mathsf{E}_\nu \prod_{i=0}^n \mathbf{1}_{\{A_i\}}(X_{t_i})(\, P_t f)(X_s),\end{aligned}$$

which we wanted to prove. The reasoning is similar if $t_n = s$.

This implies that (3.1.6) 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 3.1.6 (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 $\mathcal{S} = 2^S$. Put the probability measure P on this measurable space with $\mathsf{P}\{x\} = 1/9$. Define a sequence of i.i.d. random vectors $Z_k = (X_{3k}, X_{3k+1}, X_{3k+2})$, $k = 0, \ldots$ on $(\mathsf{S}, \mathcal{S}, \mathsf{P})$, $Z_k(s) = s$ for all $s \in \mathsf{S}$. 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. Let $\mathcal{F}_n^X$ be the natural filtration. Then $\mathsf{P}\{X_{n+1} = j \,|\, \sigma(X_n)\} = 1/3$ for each $j \in \{1,2,3\}$. However, $\{X_n\}_n$ is not a Markov chain, since $\mathsf{P}\{X_3 = 1 \,|\, \sigma(X_1, X_2)\} = f(X_1, X_2)$ with

$$f(X_1, X_2) = \begin{cases} 1/3, & (X_1, X_2) \in \{(2,3), (3,2)\} \\ 1, & (X_1, X_2) = (1,1) \\ 0, & \text{otherwise.} \end{cases}$$

Hence $f(1,1) \neq f(2,1)$, thus showing that the Markov property lacks.

Example 3.1.7 (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

$$\mathsf{E}_\nu(f(X_{t+s}) \,|\, \mathcal{F}_s^X) = \mathsf{E}_\nu\Big(f(Y_t + W_s + x) \,|\, \mathcal{F}_s^X) = g(X_s)$$

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

$$P_t f(z) = \int f(y) p(t,z,y) du,$$

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 kernel. Measurability of $P_t(x,B)$ in x for each Borel set B follows from continuity.

Example 3.1.7 ((B) Ornstein-Uhlenbeck process) Let W be a standard Brownian motion. Let $\alpha, \sigma^2 > 0 > 0$ 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 $X_t \stackrel{\mathcal{D}}{\to} \mathsf{N}(0, \sigma^2/2\alpha)$.

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 $r(s,t) = \sigma^2 \exp\{-\alpha|t-s|\}/2\alpha$.

Example 3.1.8 (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 3.1.7A). 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. Exercise3.7).

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 3.1.9 *Let X be a Markov process with state space $(E, \mathcal{E})$, initial distribution ν and transition function $(P_t)_t$. 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 its natural filtration, 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. Using the preceding lemma and the assumption, it is easily verified that Y has the Markov property (see Exercise 3.2). QED

Example 3.1.10 (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 function Q_t, $t \geq 0$, such that

$$P_t(f \circ \phi)(x) = (Q_t f) \circ \phi(x), x \in \mathbf{R} \tag{3.1.7}$$

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 $x \in \mathbf{R}$

$$\begin{aligned} P_t(f \circ \phi)(x) &= \int_{-\infty}^{\infty} p(t, x, y) f(y^2) dy \\ &= \int_0^{\infty} (p(t, x, y) + p(t, x, -y)) f(y^2) dy \\ &\overset{u=y^2 \Rightarrow y=\sqrt{u}, dy=du/2\sqrt{u}}{=} \int_0^{\infty} (p(t, x, \sqrt{u} + p(t, x, -\sqrt{u}) \frac{1}{2\sqrt{u}} f(u) du. \end{aligned} \tag{3.1.8}$$

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 (3.1.8) 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 (3.1.7) 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.

Recall that an E-valued process can be viewed as a random element of the space $E^{\mathbb{R}_+}$ of E-valued functions f on $\mathbf{R}_+$, or of a subspace $\Gamma \subset E^{\mathbb{R}_+}$ if X is known to have more structure. The σ-algebra $\Gamma \cap \mathcal{E}^{\mathbb{R}_+}$ is the smallest σ-algebra that makes all projections $f \to f(t)$ measurable.

Review BN section 2 on σ-cylinders, as well as Chapter 1.

As in Chapter 1, let $\Omega = \Gamma$ and $\mathcal{F} = \Gamma \cap \mathcal{E}^{\mathbb{R}_+}$. 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 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 $(E^{n+1}, \mathcal{E}^{n+1})$ 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 3.5). 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

$$\mathsf{P}_x\{X_t \in A\} = \int P_t(y, A)\delta_x(dy) = P_t(x, A)$$

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 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

$$\mathcal{G} = \{\{X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n\} \,|\, A_1 \ldots, A_n \in \mathcal{E}, 0 \leq t_1 < \cdots < t_n, n \in \mathbf{Z}_+\}$$

is a π-system for $\mathcal{F}^X_\infty = \mathcal{E}^{\mathrm{R}_+}$. So if we can show that $\mathcal{G} \subset \mathcal{S}$, then by BN Lemma 3.4 $\mathcal{F}^X_\infty \subset \mathcal{S}$. 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 3.6. 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* $\theta_t : E^{\mathrm{R}_+} \to E^{\mathrm{R}_+} \cup A$ by

$$(\theta_t \omega)_s = \omega_{t+s}, \quad s \geq 0, \quad \omega \in E^{\mathrm{R}_+}.$$

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 $(\Omega = E^{\mathbb{R}_+} \cap \Gamma, \mathcal{F} = \mathcal{E}^{\mathbb{R}_+} \cap \Gamma)$. 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 $\mathcal{F}_t^X = \mathcal{E}^{[0,t]} \cap \Gamma$ and θ_t is $\mathcal{F}$-measurable for every $t \geq 0$ (why?).

Theorem 3.2.4 (Generalised Markov property for canonical process) *Let Z be an $\mathcal{F}_\infty^X$-measurable random variable, non-negative or bounded. Then for every $t > 0$ and any initial distribution ν*

$$\mathsf{E}_\nu(Z \circ \theta_t \,|\, \mathcal{F}_t^X) = \mathsf{E}_{X_t} Z, \quad \mathsf{P}_\nu - \text{a.s.}$$

Before turning to the proof, note that the right-hand side of the above relation should be interpreted as $\mathsf{E}_x Z$ on $\{X_t = x\}$. By Lemma 3.2.3 this is a measurable function of X_t.

Proof. Fix an initial probability measure ν. We have to show that

$$\int_A Z \circ \theta_t d\mathsf{P}_\nu = \int_A \mathsf{E}_{X_t} Z d\mathsf{P}_\nu, \quad \forall A \in \mathcal{F}_t^X. \tag{3.2.1}$$

It is sufficient to show this for all sets A in a π-system generating $\mathcal{F}_t^X$. A convenient π-system is the collection $\mathcal{A}_t = \mathcal{A} \cap \mathcal{F}_t^X$ of cylinder sets contained in $\mathcal{F}_t^X$. Recall that $A \in \mathcal{A}_t$ whenever there exist $n \in \mathbf{Z}_+$, $0 = t_0 < t_1 < \cdots < t_n$, $A_0, \ldots, A_n \in \mathcal{E}$, $n \in \mathbf{Z}_+$, such that $A = \{X_{t_0} \in A_0, \ldots, X_{t_n} \in A_n\}$.

Now we will first show that (3.2.1) holds for all $Z = \mathbf{1}_{\{B\}}$, $B \in \mathcal{F}_\infty^X$ and $A \in \mathcal{A}_t$. Let

$$\mathcal{S} = \{B \in \mathcal{F}_\infty^X \mid \int_A \mathbf{1}_{\{B\}} \circ \theta_t d\mathsf{P}_\nu = \int_A \mathsf{E}_{X_t} \mathbf{1}_{\{B\}} d\mathsf{P}_\nu, \forall A \in \mathcal{A}_t\}.$$

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 $B_n, n = 1, \ldots, \in \mathcal{F}_\infty^X$ 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.

The collection $\mathcal{A}$ of all cylinder sets is a π-system generating $\mathcal{F}_\infty^X$. So, if we can show that $\mathcal{A} \subset \mathcal{S}$, then by BN Lemma 3.4 it follows that $\sigma(\mathcal{A}) = \mathcal{F}_\infty^X \subseteq \mathcal{S}$.

Take a cylinder set $B = \{X_{s_1} \in B_1, \ldots, X_{s_m} \in B_m\}$, where $0 \leq s_1 < \cdots < s_m$, $B_i \in \mathcal{E}$, $i = 1, \ldots, m$ and let $A \in \mathcal{A}_t$ with $A = \{X_{t_0} \in A_0, \ldots, X_{t_n} \in A_n\}$ for $t_0 = 0 < t_1 < \cdots < t_n \leq t$, $A_i \in \mathcal{E}$, $i = 0, \ldots, n$.

If $t_n < t$, using Lemma 3.1.5. it follows that

$$\begin{aligned}
&\int_A \mathbf{1}_{\{B\}} \circ \theta_t d\mathsf{P}_\nu \\
&= \int_A \mathbf{1}_{\{X_{t+s_1} \in B_1 \ldots, X_{t+s_m} \in B_m\}} d\mathsf{P}_\nu = \mathsf{E}_\nu \mathbf{1}_{\{A\}} \prod_{i=1}^m \mathbf{1}_{\{B_i\}}(X_{t+s_i}) \\
&= \mathsf{E}_\nu \prod_{j=0}^n \mathbf{1}_{\{A_j\}}(X_{t_j}) \prod_{i=1}^m \mathbf{1}_{\{B_i\}}(X_{t+s_i}) \\
&= \nu \mathbf{1}_{\{A_0\}} P_{t_1 - t_0} \cdots P_{t_n - t_{n-1}} \mathbf{1}_{\{A_n\}} P_{t+s_1-t_n} \mathbf{1}_{\{B_1\}} \cdots P_{s_m - s_{m-1}} \mathbf{1}_{\{B_m\}} \\
&= \nu \mathbf{1}_{\{A_0\}} P_{t_1 - t_0} \cdots P_{t_n - t_{n-1}} \mathbf{1}_{\{A_n\}} P_{t-t_n} P_{s_1} \mathbf{1}_{\{B_1\}} \cdots P_{s_m - s_{m-1}} \mathbf{1}_{\{B_m\}}
\end{aligned}$$

$$
\begin{aligned}
&= \nu \mathbf{1}_{\{A_0\}} P_{t_1-t_0} \cdots P_{t_n-t_{n-1}} \mathbf{1}_{\{A_n\}} P_{t-t_n} f \\
&= \mathsf{E}_\nu \prod_{j=0}^{n} \mathbf{1}_{\{A_j\}}(X_{t_j}) f(X_t) = \mathsf{E}_\nu \mathbf{1}_{\{A\}} f(X_t),
\end{aligned}
$$

with $f(x) = \delta_x P_{s_1} \mathbf{1}_{\{B_1\}} \cdots P_{s_m-s_{m-1}} \mathbf{1}_{\{B_m\}} = \mathsf{E}_x \mathbf{1}_{\{B\}}$. The argument for $t = t_n$ is analogous.

We have 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 3.2.5 (Markov jump process) Let E be a countable state space with σ-algebra $\mathcal{E}$ 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 $X_t = \mathbf{1}_{\{N_t<\infty\}} \sum_{n=0}^{\infty} \mathbf{1}_{\{N_t=n\}} Y_n$. By construction X_t has right-continuous paths.

3.3 Strong Markov property

3.3.1 Strong Markov property for right-continuous canonical Markov processes

Let X be a canonical Markov process with values in a Polish space E, equipped with the Borel-σ-algebra $\mathcal{E}$, w.r.t the natural filtration $(\mathcal{F}_t^X)_{t\geq 0}$. 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.

Definition 3.3.1 X is said to have the **strong Markov property** if for every $\mathcal{F}_\infty^X$-measurable random variable Z, with Z either bounded or non-negative, any adapted stopping time σ and any initial distribution ν

$$
\mathbf{1}_{\{\sigma<\infty\}} \mathsf{E}_\nu (Z \circ \theta_\sigma \mid \mathcal{F}_\sigma^X) = \mathbf{1}_{\{\sigma<\infty\}} \mathsf{E}_{X_\sigma} Z \quad \mathsf{P}_\nu \text{ a.s.}. \tag{3.3.1}
$$

We first prove a more general statement for stopping times taking only values from a countable set.

Lemma 3.3.2 *Let X be a canonical Markov process. Then (3.3.1) holds for any bounded or non-negative $\mathcal{F}_\infty^X$-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. First we prove that $\mathbf{1}_{\{\sigma<\infty\}}X_\sigma$ is $\mathcal{F}_\sigma^X$-measurable. To this end, check that $\{\sigma = t\} \in \mathcal{F}_t^X$ for any stopping time σ. Hence $A \in \mathcal{F}_\sigma^X$ implies that $A \cap \{\sigma = t\} \in \mathcal{F}_\sigma^X, \mathcal{F}_t^X$. Now

$$\mathbf{1}_{\{\sigma<\infty\}}\{X_\sigma \in B\} = \cup_{s\in S}\{X_s \in B\} \cap \{\sigma = s\}.$$

It is sufficient to check that $\{X_s \in B\} \cap \{\sigma = s\} \in \mathcal{F}_\sigma$. This is true iff

$$\{X_s \in B\} \cap \{\sigma = s\} \cap \{\sigma \le t\} \in \mathcal{F}_t^X$$

for any $t \ge 0$. This is easily checked. This implies that $\mathbf{1}_{\{\sigma<\infty\}}\mathsf{E}_{X_\sigma}Z$ is $\mathcal{F}_\sigma^X$-measurable as a composition of measurable maps.

The next step is to show that

$$\mathsf{E}_\nu \mathbf{1}_{\{A\}}\mathbf{1}_{\{\sigma<\infty\}} Z\circ\theta_\sigma = \mathsf{E}_\nu \mathbf{1}_{\{A\}}\mathbf{1}_{\{\sigma<\infty\}}\mathsf{E}_{X_\sigma}Z, \quad A \in \mathcal{F}_\sigma^X.$$

If $A \in \mathcal{F}_\sigma^X$ with $A \subset \{\sigma = s\}$ for some $s \in S$, then $A \in \mathcal{F}_s^X$. By the Markov property

$$\mathsf{E}_\nu \mathbf{1}_{\{A\}}\mathbf{1}_{\{\sigma<\infty\}} Z\circ\theta_\sigma = \mathsf{E}_\nu \mathbf{1}_{\{A\}} Z\circ\theta_s = \mathsf{E}_\nu \mathbf{1}_{\{A\}}\mathsf{E}_{X_s}Z = \mathsf{E}_\nu \mathbf{1}_{\{A\}}\mathbf{1}_{\{\sigma<\infty\}}\mathsf{E}_{X_\sigma}Z.$$

Let $A \in \mathcal{F}_\sigma^X$ be arbitrary. By the previous $A \cap \{\sigma = s\} \in \mathcal{F}_\sigma^X$. 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, w.r.t the natural filtration, has the strong Markov property.*

Theorem 3.3.4 *Let X be a canonical Markov process with right-continuous paths. Suppose that $x \mapsto \mathsf{E}_x f(X_s) = P_s f(x)$ is bounded continuous for each bounded continuous function f. Then the strong Markov property holds.*

Proof. Let σ be a $(\mathcal{F}_t^X)_t$-adapted stopping time. Let first $Z = f_1(X_{t_1})f_2(X_{t_2})\cdots f_n(X_{t_n})$, with $n \in \mathbf{Z}_+$, $t_1 < \cdots < t_n$, $f_1,\ldots,f_n$ bounded, continuous functions. Consider

$$\sigma_m = \sum_{k=1}^\infty \frac{k}{2^m}\cdot \mathbf{1}_{\{\frac{k-1}{2^m}<\sigma\le\frac{k}{2^m}\}} + \infty\cdot\mathbf{1}_{\{\sigma=\infty\}}.$$

Then σ_m takes countably many different values and $\sigma_m \downarrow \sigma$. By virtue of Lemma 3.3.2 for all $A \in \mathcal{F}_{\sigma_m}^X$

$$\mathsf{E}_\nu \mathbf{1}_{\{A\}}\mathbf{1}_{\{\sigma_m<\infty\}} Z\circ\theta_{\sigma_m} = \mathsf{E}_\nu \mathbf{1}_{\{A\}}\mathbf{1}_{\{\sigma_m<\infty\}}\mathsf{E}_{X_{\sigma_n}}Z.$$

Next, use that if $A \in \mathcal{F}_\sigma^X$, then $A \in \mathcal{F}_{\sigma_m}^X$. Moreover, $\mathbf{1}_{\{A\}}\mathbf{1}_{\{\sigma_m<\infty\}}Z\circ\theta_{\sigma_m} \to \mathbf{1}_{\{A\}}\mathbf{1}_{\{\sigma<\infty\}}Z\circ\theta_\sigma$ and $\mathbf{1}_{\{A\}}\mathbf{1}_{\{\sigma_m<\infty\}}Z \to \mathbf{1}_{\{A\}}\mathbf{1}_{\{\sigma<\infty\}}\mathsf{E}_{X_\sigma}Z$, $m \to \infty$. Apply dominated convergence.

Let next $Z = \prod_{i=1}^n \mathbf{1}_{\{A_i\}}(X_{t_i})$, with $n \in \mathbf{Z}_+$, $t_1 < \cdots < t_n$, $A_1, \ldots, A_n \in \mathcal{E}$. Let f_i^m be given by

$$f_i^m(x) = 1 - m \cdot (m^{-1} \wedge d(x, A_i)),$$

where d is a metric on E, consistent with the topology. Then f_i^m are continuous, bounded functions and by the previous

$$\mathsf{E}\mathbf{1}_{\{A\}}\mathbf{1}_{\{\sigma<\infty\}} \prod_i f_i^m(X_{t_i}) \circ \theta_\sigma = \mathsf{E}\mathbf{1}_{\{A\}}\mathbf{1}_{\{\sigma<\infty\}}\mathsf{E}_{X_\sigma} \prod_i f_i^m(X_{t_i}), \quad A \in \mathcal{F}_\sigma^X.$$

The random variable on the left-handside converges pointwise to $\mathbf{1}_{\{A\}}\mathbf{1}_{\{\sigma<\infty\}} Z \circ \theta_\sigma$, the one on right-handside converges pointwise to $\mathbf{1}_{\{A\}}\mathbf{1}_{\{\sigma<\infty\}}\mathsf{E}_{X_\sigma} Z$. Use monotone convergence.

Finally, we apply the d-system recipe to show that the strong Markov property holds for $Z = \mathbf{1}_{\{A\}}$ with $A \in \mathcal{F}_\infty^X$. Then use the standard machinery. QED

Corollary 3.3.5 *Assume that X is a right-continuous canonical process with state space $E \subset \mathbf{Z}_+^d$, $d < \infty$, and $\mathcal{E} = 2^E$. Then X has the strong Markov property.*

The corollary implies that the canonical Poisson process has the strong Markov property, as well as the canonical right-continuous Markov jump process.

Corollary 3.3.6 *Canonical* BM *has the strong Markov property.*

We discuss some general applications of the strong Markov property.

Corollary 3.3.7 *Assume the conditions of Theorem 3.3.4 and let Y be a bounded $\mathcal{F}_\tau^X$-measurable random variable. Then*

$$\mathsf{E}_\nu Y(Z \circ \theta_\tau) = \mathsf{E}_\nu(Y \mathsf{E}_{X_\tau} Z).$$

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 $A \in \mathcal{F}_0^X$ then $\mathsf{P}_x(A) = 0$ or 1 for all $x \in E$.*

ii) *If τ is an $(\mathcal{F}_t^X)_t$-stopping time, then $\mathsf{P}_x\{\tau = 0\} = 0$ or 1, for all $x \in E$.*

Proof. For (i) use Corollary 3.3.7 with $Y = Z = \mathbf{1}_{\{A\}}$ and $\tau = 0$ (see Exercise 3.27). QED

Example 3.3.9 (Strong Markov property fails) Consider Example 3.1.7 (A). We slightly adapt the definition of the process X:

$$X_t = X_0 + \mathbf{1}_{\{X_0 \neq 0\}} W_t,$$

with $X_0 \stackrel{\mathsf{d}}{=} \delta_x$, independent of W. One can show that X is a Markov process. Let $\tau = \inf\{t \geq 0 \,|\, X_t = 0\}$. Consider for instance $Z = \mathbf{1}_{\{[1,\infty)\}}(X_1)$ for $y > 0$, Then $\mathbf{1}_{\{\tau<\infty\}}\mathsf{E}_{X_\tau} Z = 0$. However, for $x \neq 0$ one has $\mathbf{1}_{\{\tau<\infty\}}\mathsf{E}_x(Z \circ \theta_\tau \,|\, \mathcal{F}_\tau) = P_1(0, [1, \infty))\,(> 0)$ on $\tau < \infty$. (cf. Exercise 3.28).

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. The *Lévy processes* are a class of processes with this property of which canonical BM and the canonical Poisson process are well-known examples.

Lemma 3.3.10 *Let X be a right-continuous canonical Markov process. Suppose that X has stationary, independent increments and τ is a finite stopping time. Then the process $X(\tau) = (X_{\tau+t} - X_\tau)_{t\geq 0}$ is independent of $\mathcal{F}^X_\tau$ and for each initial distribution ν, the distribution of $X(\tau)$ under P_ν is the same as the distribution of X under P_x, for any $x \in E$.*

Proof. Put $Y_t = X_{\tau+t} - X_\tau$, $t \geq 0$. For $t_1 < \cdots < t_n$ and bounded, measurable functions $f_1, \ldots, f_n$ we have

$$\begin{aligned}\mathsf{E}_\nu\Big(\prod_k f_k(Y_{t_k}) \,|\, \mathcal{F}^X_\tau\Big) &= \mathsf{E}_\nu\Big(\prod_k f_k(X_{\tau+t_k} - X_\tau) \,|\, \mathcal{F}^X_\tau\Big) \\ &= \mathsf{E}_{X_\tau} \prod_k f_k(X_{t_k} - X_0),\end{aligned}$$

by the strong Markov property. As a consequence, the proof is complete once we have shown that for arbitrary $x \in E$

$$\mathsf{E}_x \prod_{k=1}^n f_k(X_{t_k} - X_0) = P_{t_1} f_1 \cdots P_{t_n - t_{n-1}} f_n(0),$$

with $0 \in E$ a selected state (cf. Characterisation Lemma 3.1.5). We prove this by induction on n. Suppose first that $n = 1$. By stationarity of the increments, the distribution of $X_{t_1} - X_0$ under P_x is independent of x. In particular, we can take $x = 0$, obtaining

$$\mathsf{E}_x f_1(X_{t_1} - X_0) = \mathsf{E} f_1(X_{t_1}) = P_{t_1} f_1(0).$$

Now suppose that the statement is true for for $n-1$ and all bounded, measurable functions $f_1, \cdots, f_{n-1}$. We have

$$\begin{aligned}\mathsf{E}_x \prod_{k=1}^n f_k(X_{t_k} - X_0) &= \mathsf{E}_x \mathsf{E}_x\Big(\prod_{k=1}^n f_k(X_{t_k} - X_0) \,|\, \mathcal{F}^X_{t_{n-1}}\Big) \\ &= \mathsf{E}_x \prod_{k=1}^{n-1} f_k(X_{t_k} - X_0) \mathsf{E}_x(f_n(X_{t_n} - X_0) \,|\, \mathcal{F}^X_{t_{n-1}}).\end{aligned}$$

By independence of the increments

$$\begin{aligned}\mathsf{E}_x(f_n(X_{t_n} - X_0) \,|\, \mathcal{F}^X_{t_{n-1}}) &= \mathsf{E}_x(f_n(X_{t_n} - X_{t_{n-1}} + X_{t_{n-1}} - X_0) \,|\, \mathcal{F}^X_{t_{n-1}}) \\ &= g_x(X_{t_{n-1}} - X_0),\end{aligned}$$

where

$$g_x(y) = \mathsf{E}_x f_n(X_{t_n} - X_{t_{n-1}} + y).$$

The P_x-distribution of $X_{t_n} - X_{t_{n-1}}$ is the same as the distribution of $X_{t_n - t_{n-1}} - X_0$, and is independent of x. Hence, we may put $x = y$, so that $X_0 = y$, P_y-a.s., and

$$g_x(y) = \mathsf{E}_y f_n(X_{t_n - t_{n-1}} - X_0 + y) = \mathsf{E}_y f_n(X_{t_n - t_{n-1}}) = P^{\delta}_{t_n - t_{n-1}} f_n(y).$$

We finally obtain

$$\begin{aligned}\mathsf{E}_x \prod_{k=1}^{n} f_k(X_{t_k} - X_0) &= \mathsf{E}_x \Big(\prod_{k=1}^{n-1} f_k(X_{t_k} - X_0) P_{t_n - t_{n-1}} f_n(X_{t_{n-1}} - X_0) \Big) \\ &= \mathsf{E}_x \prod_{k=1}^{n-2} (f_{n-1} P_{t_n - t_{n-1}} f_n(X_{t_{n-1}} - X_0)).\end{aligned}$$

By the induction hypothese, this equals $P_{t_1} f_1 \cdots P_{t_n - t_{n-1}} f_n(0)$, thus completing the proof. QED

The following lemma is often useful in connection with the strong Markov property.

Lemma 3.3.11 *If σ and τ are finite $(\mathcal{F}_t)_t$-stopping times, then $\sigma + \tau \circ \theta_\sigma$ is also a finite $(\mathcal{F}_t)_t$-stopping time.*

Proof. Since $(\mathcal{F}_t)_t$ is right-continuous, it suffices to prove that $\{\sigma + \tau \circ \theta_\sigma < t\} \in \mathcal{F}_t$ for every $t > 0$ (cf. Lemma 1.6.6). Observe that

$$\{\sigma + \tau \circ \theta_\sigma < t\} = \cup_{q \geq 0} \{\tau \circ \theta_\sigma < q\} \cap \{\sigma \leq t - q\}.$$

The indicator of the event $\{\tau \circ \theta_\sigma < q\}$ can be written as $\mathbf{1}_{\{\tau < q\}} \circ \theta_\sigma$. By Exercise 3.23, it follows that $\{\tau \circ \theta_\sigma < q\} \in \mathcal{F}_{\sigma + q}$. By definition of the latter

$$\{\tau \circ \theta_\sigma < q\} \cap \{\sigma \leq t - q\} = \{\tau \circ \theta_\sigma < q\} \cap \{\sigma + q \leq t\} \in \mathcal{F}_t.$$

This completes the proof.

3.3.2 Applications to Brownian Motion

In this subsection W is the canonical BM on $(\Omega = \mathcal{C}[0,\infty), \mathcal{F} = \mathcal{C}(0,\infty] \cap \mathcal{B}^{\mathrm{R}_+})$ with associated Markov process X. Since BM has stationary, independent increments, Corollary 3.3.10 implies that for every $(\mathcal{F}^X_t)_t$-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.

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 $(\mathcal{F}^X_t)_t$-stopping time, w.r.t P_0. Denote $\tau'_0 = \inf\{t > 0 \,|\, t \geq \tau_1, X_t = 0\}$.

Lemma 3.3.12 **i)** *$\mu(A, \tau)$ is a measurable function on $(\Omega, \mathcal{F})$.*

ii) $\mu(A) := \mathsf{E}_0 \mu(A, \tau_1) = 2 \int_{-\infty}^{0} \mathbf{1}_{\{A\}}(x) d\lambda(x) + 2 \int_0^1 (1 - x) \mathbf{1}_{\{A\}}(x) d\lambda(x)$.

ii) $\mu'(A) = \mathsf{E}_0 \mu(A, \tau'_0) = 2\lambda(A)$.

Proof. See exercise 3.9. For the proof of (ii), note that

$$\begin{aligned}
\mathsf{E}_0\mu(A,\tau_1) &= \mathsf{E}_0\int_0^\infty \mathbf{1}_{\{A\}}(X_t)\mathbf{1}_{\{t\le\tau_1\}}dtd\mathsf{P}_0\\
&= \int_0^\infty\int_\Omega \mathbf{1}_{\{A\}}(X_t)\mathbf{1}_{\{t\le\tau\}}d\mathsf{P}_0dt\\
&= \int_0^\infty \mathsf{P}_0\{X_t\in A, t\le\tau_1\}dt\\
&= \int_0^\infty\int_A w(t,x)d\lambda(x)dt\\
&= \int_A\int_0^\infty w(t,x)dtd\lambda(x),
\end{aligned}$$

where

$$w(t,x)=\frac{1}{\sqrt{2\pi t}}\Big(e^{-x^2/2t}-e^{-(x-2)^2/2t}\Big),\quad x\le 1.$$

This follows from

$$\mathsf{P}_0\{X_t\in A,\tau_1\le t\}=\mathsf{P}_2\{X_t\in A\}.$$

Why is this true? (ii) can then be shown by writing

$$w(t,x)=-\int_{x-2}^x\frac{u}{t^{3/2}\sqrt{2\pi}}e^{-u^2/2t},$$

applying Fubini, and do a substitution $s=t^{-1/2}$. Distinguish the cases that $x\le 0$ and $0<x\le 1$. QED

let $f,g:\mathbf{R}\to\mathbf{R}$ be Lebesgue measurable, integrable functions with $\int_{\mathrm{R}} g(x)dx\neq 0$.

Theorem 3.3.13

$$\lim_{T\to\infty}\frac{\int_0^T f(W_t)dt}{\int_0^T g(W_s)ds}=\frac{\int_R f(x)dx}{\int_R g(x)dx},\quad \text{a.s.}$$

Proof. Put $\tau_0^1=\tau_0$ and $\tau_1^1=\tau_1$. Inductively define for $n\ge 2$: $\tau_0^n=\inf\{t\ge\tau_1^{n-1}\,|\,X_t=0\}$, and $\tau_1^n=\inf\{t\ge\tau_0^n\,|\,X_t=1\}$. By virtue of the standard machinery, one has

$$\mathsf{E}_0\int_0^{\tau_0^1}f(X_t)dt=2\int_{\mathrm{R}}f(x)dx.$$

Now, for any $T\ge 0$ define

$$K(T)=\max\{n\,|\,\tau_0^n\le T\}.$$

Then $\lim_{T\to\infty}\int_0^T f(X_t)dt/K(T)=2\int_{\mathrm{R}}f(x)dx$, P_0-a.s. The result then follows. QED

The second example that we give, is the so-called reflection principle (compare with Ch.1, Exercise 1.11). Recall that we denote the hitting time of $x\in\mathbf{R}$ by τ_x. This is a finite stopping time with respect to the natural filtration of the BM(see Example 1.6.9).

Theorem 3.3.14 (Reflection principle) *Let $x \in \mathbf{R}$ be given. Define the process W' by*

$$W'_t = \begin{cases} W_t, & t \le \tau_x \\ 2x - W_t, & t > \tau_x. \end{cases}$$

Then W' is a standard BM.

Proof. Define the processes Y and Z by $Y = W^{\tau_x}$ and $Z_t = W_{\tau_x + t} - W_{\tau_x} = W_{\tau_x+t} - x$, $t \ge 0$. By Corollary 3.3.10 the processes Y and Z are independent and Z is a standard BM. By symmetry of BM, it follows that $-Z$ is also a BM that is independent of Y, and so the two pairs (Y, Z) and $(Y, -Z)$ have the same distribution (i.e. the fdd's are equal)

Now, observe that for $t \ge 0$

$$W_t = Y_t + Z_{t-\tau_x}\mathbf{1}_{\{t>\tau_x\}}, \quad W'_t = Y_t - Z_{t-\tau_x}\mathbf{1}_{\{t>\tau_x\}}.$$

In other words, we have $W = \phi(Y, Z)$ and $W' = \phi(Y, -Z)$, where $\phi : C[0,\infty) \times \{\omega \in C[0,\infty) \,|\, \omega_0 = 0\} \to C[0,\infty)$ is given by

$$\phi(y,z)(t) = y(t) + z(t - \psi(y))\mathbf{1}_{\{t>\psi(y)\}},$$

where $\psi : C[0,\infty) \to [0,\infty]$ is defined by $\psi(y) = \inf\{t > 0 \,|\, y(t) = x\}$. Consider the induced σ-algebra on $C[0,\infty)$ and $\{\omega \in C[0,\infty) \,|\, \omega_0 = 0\}$ inherited from the σ-algebra $\mathcal{B}^{[0,\infty)}$ on $\mathbf{R}^{[0,\infty)}$. It is easily verified that ψ is a Borel-measurable map, and that ϕ is measurable as the composition of measurable maps (cf. Exercise 3.29). Since $(Y, Z) \stackrel{\mathsf{d}}{=} (Y, -Z)$, it follows that $W = \phi(Y, Z) \stackrel{\mathsf{d}}{=} \phi(Y, -Z) = W'$. 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

$$S_t = \sup_{s \le t} W_s.$$

Corollary 3.3.15 *Let W be a standard* BM *and S its running maximum. Then*

$$\mathsf{P}\{W_t \le x, S_t \ge y\} = \mathsf{P}\{W_t \le x - 2y\}, \quad x \le y.$$

The pair (W_t, S_t) has joint density

$$(x,y) \mapsto \frac{(2y-x)e^{-(2y-x)^2/2t}}{\sqrt{\pi t^3/2}}\mathbf{1}_{\{x \le y\}},$$

with respect to the Lebesgue measure.

Proof. Let W' be the process obtained by reflecting W at the hitting time τ_y. Observe that $S_t \ge y$ if and only if $t \ge \tau_y$. Hence, the probability of interest equals $\mathsf{P}\{W_t \le x, t \ge \tau_y\}$. On the event $\{t \ge \tau_y\}$ we have $W_t = 2y - W'_t$, and so we have to calculate $\mathsf{P}\{W'_t \ge 2y - x, t \ge \tau_y\}$. Since $x \le y$, we have $2y - x \ge y$. hence $\{W'_t \ge 2y - x\} \subseteq \{W'_t \ge y\} \subseteq \{t \ge \tau_y\}$. It follows that $\mathsf{P}\{W'_t \ge 2y - x, t \ge \tau_y\} = \mathsf{P}\{W'_t \ge 2y - x\}$. By the reflection principle and symmetry of BM this proves the first statement. The second follows from Exercise 3.30. 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 3.31). This shows in particular that $S_t \stackrel{\mathsf{d}}{=} |W_t|$ for every $t \geq 0$. This allows to 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 2.29 of Chapter 2.

Corollary 3.3.16 *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 3.32. 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 3.3.17 *The zero set $Z = \{t \geq 0 \,|\, W_t = 0\}$ of standard* BM *is a.s. closed, unbounded, contains no isolated points and has Lebesgue measure 0.*

Proof. In view of Exercise 1.27, 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 3.3.11 the random time σ_q is a stopping time. The strong Markov property implies that $W_{\sigma_q+t} - W_{\sigma_q}$ is a standard BM. 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 Feller-Dynkin processes

The first main goal is to study conditions under which the canonical Markov process on $(E_+^{\mathbb{R}}, \mathcal{E}_+^{\mathbb{R}})$ can be modified to have right-continuous paths without losing the Markov property.

For this regularisation procedure we will invoke regularisation of supermartingales. This requires constructing a suitable supermartingale associated with the Markov process. To this end we will first study a number of properties of a special class of transition functions that are important by ithemselves.

3.4.1 Feller-Dynkin transition functions and resolvents

In section 3.1 we saw that a homogeneous transition function $(P_t)_{t\geq 0}$ on a measurable space $(E, \mathcal{E})$ can be viewed as a semigroup of operators on the space of bounded, measurable functions on E.

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. In general one may assume that E is a locally compact Hausdorff space with countable base.

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

$$\|f\|_\infty = \sup_{x\in E} |f(x)|.$$

In these notes, we can formally describe $C_0(E)$ by

$$C_0(E) = \left\{ f : E \to \mathbf{R} \;\middle|\; \begin{array}{l} f \text{ continuous and} \\ \text{for each } \epsilon > 0 \text{ there exists a compact set } K = K(\epsilon, f), \\ \text{such that } |f(x)| \leq \epsilon, \text{ for } x \notin K \end{array} \right\}$$

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 $C_0(E)$.

Definition 3.4.1 The transition function $(P_t)_{t\geq 0}$ is called a *Feller-Dynkin transition function* if

i) $P_t C_0(E) \subseteq C_0(E)$;

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

$$\|P_t f\|_\infty = \sup_{x\in E} \left| \int_E f(y) P_t(x, dy) \right| \leq \|f\|_\infty.$$

So, for all $t \geq 0$ we have $\|P_t\|_\infty \leq 1$, 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 3.4.1. 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

$$R_\lambda f(x) = \int_0^\infty e^{-\lambda t} P_t f(x) dt.$$

If we view $t \mapsto P_t$ as an operator-valued map, then R_λ is simply its Laplace transform calculated at the 'frequency' λ. The next theorem collects the most important properties of the operators R_λ. In particular, it states that for all $\lambda > 0$, R_λ is in fact an operator that maps C_0 to itself. It is called the *resolvent of order* λ.

Theorem 3.4.2 *Let R_λ, $\lambda > 0$, be the resolvents of a Feller-Dynkin transition function. Then $R_\lambda C_0(E) \subseteq C_0(E)$ for every $\lambda > 0$ and the resolvent equation*

$$R_\mu - R_\lambda + (\mu - \lambda)R_\mu R_\lambda = 0$$

holds for all $\lambda, \mu > 0$. Moreover the image $R_\lambda C_0(E)$ does not depend on λ and is dense in $C_0(E)$.

Proof. Denote the transition function by $(P_t)_t$. Fix $f \in C_0(E)$. Then $P_t f \in C_0(E)$ for every $t \geq 0$. Hence, if $x_n \to x$ in E, then the function $t \mapsto P_t f(x_n)$ converges pointwise to the function $t \mapsto P_t f(x)$. Since $|P_t f(y)| \leq \|f\|_\infty$ for all $y \in E$ and $t \geq 0$, we can use dominated convergence to obtain that

$$R_\lambda f(x_n) = \int_0^\infty e^{-\lambda t} P_t f(x_n) dt \to \int_0^\infty e^{-\lambda t} Ptf(x) dt = R_\lambda f(x).$$

Consequently $R_\lambda f$ is a continuous function.

By the same reasoning, we find that $R_\lambda f(x_n) \to 0$ as $\|x_n\| \to \infty$. This implies that $R_\lambda f \in C_0(E)$, for all $f \in C_0(E)$. In other words $R_\lambda C_0(E) \subseteq C_0(E)$.

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 Fubini's theorem. A change of variables, the semigroup property of the transition function and another application of Fubini 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.

The resolvent equation implies that $R_\lambda = R_\mu(\mathbf{I} + (\mu - \lambda) R_\lambda)$. This shows that $R_\lambda C_0(E) \subseteq R_\mu C_0(E)$. Reversing the roles of μ and λ shows that $R_\lambda C_0(E) = R_\mu C_0(E)$ is independent of λ.

To prove that $R_\lambda C_0(E)$ is dense in $C_0(E)$ (i.e. $\overline{R_\lambda C_0(E)} = C_0(E)$), consider an arbitrary, bounded linear functional A on $C_0(E)$ that vanishes on $R_\lambda(C_0(E)$. By the Riesz representation theorem (BN §9 Theorem ??) there exist finite Borel measures ν and ν' on E such that

$$A(f) = \int_E f d\nu - \int_E f d\nu' = \int_E f d(\nu - \nu'),$$

for every $f \in C_0(E)$. By part (ii) of Definition 3.4.1 and dominated convergence, for every $x \in E$

$$\lambda R_\lambda f(x) = \int_0^\infty \lambda e^{-\lambda t} P_t f(x) dt = \int_0^\infty e^{-s} P_{s/\lambda} f(x) ds \to f(x), \quad \lambda \to \infty. \tag{3.4.1}$$

Note that $\|\lambda R_\lambda f\|_\infty \le \|f\|_\infty$. Then dominated convergence implies

$$0 = A(\lambda R_\lambda f) = \int_E \lambda R_\lambda f(x)(\nu - \nu')dx \to \int_E f(x) d(\nu - \nu')(dx) = A(f), \quad \lambda \to \infty.$$

We conclude that the functional A vanishes on the entire space $C_0(E)$. By the Corollary to the Hahn-Banach theorem (BN §11), this shows that $R_\lambda C_0(E)$ is dense in $C_0(E)$. QED

Observe that for every $f \in C_0(E)$ the resolvent R_λ satisfies

$$\|R_\lambda f\|_\infty \le \int_0^\infty e^{-\lambda t} \|P_t f\|_\infty dt \le \|f\|_\infty \int_0^\infty e^{-\lambda t} dt = \frac{\|f\|_\infty}{\lambda}.$$

Hence, for the linear transformation $R_\lambda : C_0(E) \to C_0(E)$ has $\|R_\lambda\|_\infty \le 1/\lambda$. Let us also note that in the proof of Theorem 3.4.2 we saw that

$$P_t R_\lambda f(x) = \int_0^\infty e^{-\lambda u} P_{t+u} f(x) du. \tag{3.4.2}$$

By the semigroup property of the transition function, the right-hand side of the above equation equals $R_\lambda P_t f(x)$. In other words the operators R_λ and P_t commute.

Corollary 3.4.3 *For a Feller-Dynkin transition function $(P_t)_t$ and its resolvents R_λ, it holds for every $f \in C_0(E)$ that $\|P_t f - f\|_\infty \to 0$, as $t \downarrow 0$, and $\|\lambda R_\lambda f - f\|_\infty \to 0$ as $\lambda \to \infty$. As a consequence $\|P_{t+h} f - P_t f\|_\infty \to 0$, $h \to 0$, so that $P_t f$ is continuous in t, for $t > 0$ and $f \in C_0(E)$.*

Proof. Relation (3.4.2) shows that for $f \in C_0(E)$

$$P_t R_\lambda f(x) = e^{\lambda t} \int_t^\infty e^{-\lambda s} P_s f(x) ds.$$

It follows that

$$P_t R_\lambda f(x) - R_\lambda f(x) = (e^{\lambda t} - 1) \int_t^\infty e^{-\lambda s} P_s f(x) ds - \int_0^t e^{-\lambda s} P_s f(x) ds.$$

And so

$$\|P_t R_\lambda f - R_\lambda f\|_\infty \le (e^{\lambda t} - 1)\|R_\lambda |f|\|_\infty + t\|f\|_\infty.$$

Since the right-hand side tends to 0 as $t \downarrow 0$, this shows the statement of the corollary 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 $\lim_{t\downarrow 0} \|P_t f - f\|_\infty = 0$.

The second statement follows from the first by dominated convergence (see (3.4.1)).

For the third, note for $h > 0$ that $\|P_{t-h} f - P_t f\|_\infty = \|P_{t-h}(f - P_h f)\|_\infty \leq \|f - P_h f\|_\infty \to 0$ as $h \downarrow 0$. A similar reasoning shows that $\|P_{t+h} f - P_t f\|_\infty \to 0$, $h \downarrow 0$. QED

Example 3.4.4 The BM-process from Example 3.1.7 (A) is a Feller-Dynkin process. 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 3.12)

Example 3.4.5 Let X be the Markov jump process defined in Example 3.2.5. It is a Feller-Dynkin process if $PC_0(E) \subseteq C_0(E)$ (cf. Exercise 3.10). The resolvent is given by (cf. Exercise 3.13).

$$R_\mu f = \frac{1}{\lambda + \mu} \sum_{n\geq 0} \Big(\frac{\lambda}{\lambda + \mu}\Big)^n P^n f = ((\lambda + \mu)\mathbf{I} - \lambda P)^{-1} f \quad f \in C_0(E).$$

This is by a direct computation, using Fubini, dominated convergence and the fact that (proved using induction)

$$\int_0^\infty t^n e^{-\xi t} dt = \frac{n!}{\xi^{n+1}}.$$

3.4.2 Generators

In the discrete-time case, the resolvent can be expressed in terms of the transition kernel P by

$$R_\lambda = \sum_{n=0}^\infty e^{-\lambda n} P^n = (\mathbf{I} - e^{-\lambda} P)^{-1}.$$

In the continuous-time case, we have a similar relation between resolvent and *generator*, which is a linear operator determining a Feller-Dynkin transition function.

Definition 3.4.6 Let $(P_t)_{t\geq 0}$ be a Feller-Dynkin transition function. The infinitesimal generator A of $(P_t)_t$ is the linear operator $\mathsf{A} : \mathcal{D}(\mathsf{A}) \subseteq C_0(E) \to C_0(E)$ defined as follows. Write $f \in \mathcal{D}(\mathsf{A})$ if there exists $g \in C_0(E)$, such that

$$\Big\| \frac{P_h f - f}{h} - g \Big\|_\infty \to 0, \quad h \downarrow 0,$$

and then we define $g = \mathsf{A} f$.

From the definition we immediately see that for $f \in \mathcal{D}(\mathsf{A})$

$$\mathsf{E}_\nu(f(X_{t+h}) - f(X_t) \,|\, \mathcal{F}_t^X) = h\mathsf{A}f(X_t) + o(h), \quad \mathsf{P}_\nu - \text{a.s.},$$

as $h \downarrow 0$. In this sense the generator describes the motion in an infinitesimal time-interval.

Example 3.4.7 Consider the Markov jump process from Example 3.2.5. The generator is given by $\mathsf{A} = \lambda(P - \mathbf{I})$. Check that

$$R_\mu f = (\mu\mathbf{I} - \mathsf{A})^{-1} f, \quad f \in C_0(E). \tag{3.4.3}$$

The following lemma gives some basic properties of the generator. Then we will derive that (3.7.1) is a generic property of the generator.

Lemma 3.4.8 *Suppose that $f \in \mathcal{D}(\mathsf{A})$.*

i) *$P_t\mathcal{D}(\mathsf{A}) \subseteq \mathcal{D}(\mathsf{A})$, for $t \geq 0$.*

ii) *The function $t \mapsto P_t f$ is differentiable in $C_0(E)$ 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,

$$\lim_{h\downarrow 0} \Big\| \frac{P_{t+h} f - P_t f}{h} - P_t \mathsf{A} f \Big\|_\infty = \lim_{h\downarrow 0} \Big\| \frac{P_{t+h} f - P_t f}{h} - \mathsf{A} P_t f \Big\|_\infty = 0.$$

iii) *For every $t \geq 0$*

$$P_t f - f = \int_0^t P_s \mathsf{A} f ds = \int_0^t \mathsf{A} P_s f ds.$$

Proof. (i) follows from the second relation in (ii). For (ii) note that

$$\Big\| \frac{P_{t+h} f - P_t f}{h} - P_t \mathsf{A} f \Big\|_\infty = \Big\| P_t \Big(\frac{P_h f - f}{h} - \mathsf{A} f \Big) \Big\|_\infty \leq \Big\| \frac{P_h f - f}{h} - \mathsf{A} f \Big\|_\infty.$$

Taking the limit $h \downarrow 0$ yields that

$$\lim_{h\downarrow 0} \Big\| \frac{P_{t+h} f - P_t f}{h} - P_t \mathsf{A} f \Big\|_\infty = 0. \tag{3.4.4}$$

Since $\mathsf{A}f \in C_0(E)$, $g = P_t\mathsf{A}f \in C_0(E)$. Rewriting (3.4.4) gives

$$\lim_{h\downarrow 0} \Big\| \frac{P_h(P_tf) - P_tf}{h} - g \Big\|_\infty = 0.$$

Hence $g = \mathsf{A}P_tf$. Consequently $P_t\mathsf{A}f = g = \mathsf{A}P_tf = (d^+/dt)P_tf$ (d^+/dt stands for the right-derivative). To see that the left derivative exists and equals the right-derivative, observe for $h > 0$ that

$$\begin{aligned}\Big\| \frac{P_tf - P_{t-h}f}{h} - P_t\mathsf{A}f \Big\|_\infty &\le \Big\| \frac{P_tf - P_{t-h}f}{h} - P_{t-h}\mathsf{A}f \Big\|_\infty + \Big\| P_{t-h}\mathsf{A}f - P_t\mathsf{A}f \Big\|_\infty \\ &\le \Big\| \frac{P_hf - f}{h} - \mathsf{A}f \Big\|_\infty + \Big\| P_{t-h}\mathsf{A}f - P_t\mathsf{A}f \Big\|_\infty \to 0, \quad h \downarrow 0,\end{aligned}$$

where we have used the strong continuity (Corollary 3.4.3) and the fact that $\mathsf{A}f \in C_0(E)$. For (iii) note that $(d/dt)P_tf = P_t\mathsf{A}f$ is a continuous function of t by Corollary 3.4.3. It is therefore integrable, and so

$$P_tf - f = \int_0^t \frac{d}{ds}P_sf\,ds = \int_0^t \mathsf{A}P_sf\,ds = \int_0^t P_s\mathsf{A}f\,ds.$$

QED

The next theorem gives a full description of the generator in terms of the resolvents R_λ. In the proof we need the following lemma.

Lemma 3.4.9 *For $h > 0$ define the linear operators*

$$A_hf = \frac{1}{h}(P_hf - f), \quad B_hf = \frac{1}{h}\int_0^h P_tf\,dt.$$

Then $B_hf \in \mathcal{D}(\mathsf{A})$ for all $h > 0$ and $f \in C_0(E)$, and $\mathsf{A}B_h = A_h$.

Proof. First note that $A_hB_s = A_sB_h$. This follows from

$$\begin{aligned}A_h(B_sf) &= \frac{1}{h}(P_hB_sf - B_sf) \\ &= \frac{1}{hs}\int_0^s (P_{t+h}f - P_tf)dt = \frac{1}{hs}\int_0^s P_t\int_0^h P_u\mathsf{A}f\,du\,dt \\ &= \frac{1}{hs}\int_0^h P_u\int_0^s P_t\mathsf{A}f\,dt\,du = \frac{1}{hs}\int_0^h (P_{s+u}f - P_uf)du \\ &= \frac{1}{s}(P_sB_hf - B_hf) = A_sB_hf.\end{aligned}$$

In the third equality we have used Lemma 3.4.8 (iii), in the other equalities we have interchange arguments relying on Fubini's theorem.

Note that for $f \in C_0(E)$

$$\|B_hf - f\|_\infty \le \frac{1}{h}\int_0^h \|P_tf - f\|_\infty dt \to 0, \quad h \downarrow 0,$$

by strong continuity. Now

$$\lim_{h\downarrow 0} \|A_h B_s f - A_s f\|_\infty = \lim_{h\downarrow 0} \|A_s(B_h f - f)\|_\infty \le \lim_{h\downarrow 0} \|A_s\|_\infty \|B_h f - f\|_\infty = 0.$$

This shows that $B_s f \in \mathcal{D}(\mathsf{A})$ with $\mathsf{A} B_s f = A_s f$, which is what we wanted to prove. QED

Theorem 3.4.10 *$\mathcal{D}(\mathsf{A}) = R_\lambda C_0(E)$ (for any $\lambda > 0$), whence $\mathcal{D}(\mathsf{A})$ is dense in $C_0(E)$. For every $\lambda > 0$ the transformation $\lambda \mathbf{I} - A : \mathcal{D}(\mathsf{A}) \to C_0(E)$ is invertible with inverse R_λ.*

Proof. Let $f \in \mathcal{D}(\mathsf{A})$. Then

$$\begin{aligned} R_\lambda[(\lambda \mathbf{I} - \mathsf{A})f] &= \int e^{-\lambda t} P_t(\lambda f - \mathsf{A} f)dt \\ &= \lambda \int e^{-\lambda t} P_t f dt - \int e^{-\lambda t} \frac{d}{dt} P_t f dt. \end{aligned}$$

Integration by parts of the last integral shows that $f = R_\lambda[(\lambda \mathbf{I} - \mathsf{A})f]$. Since $(\lambda \mathbf{I} - \mathsf{A})f \in C_0(E)$, we conclude that $\mathcal{D}(\mathsf{A}) \subseteq R_\lambda C_0(E)$ and $R_\lambda(\lambda \mathbf{I} - \mathsf{A}) = \mathbf{I}$ on $\mathcal{D}(\mathsf{A})$.

To prove the converse, recall that P_h and R_λ commute. For $f \in C_0(E)$ we want to show that $R_\lambda f \in \mathcal{D}(\mathsf{A})$ and that $(\lambda \mathbf{I} - \mathsf{A}) R_\lambda f = f$. Fix $f \in C_0(E)$ and let $h > 0$. Then

$$A_h R_\lambda f = R_\lambda A_h f = R_\lambda \mathsf{A} B_h f = \int_0^\infty e^{-\lambda t} P_t A B_h f dt.$$

By virtue of Lemma 3.4.8 and integration by parts, the right-handside equals

$$\int_0^\infty e^{-\lambda t} \frac{d}{dt} P_t B_h f dt = \lambda R_\lambda B_h f - B_h f.$$

Recall that $\|B_h f - f\|_\infty \to 0$ as $h \downarrow 0$. Hence

$$\begin{aligned} \lim_{h\downarrow 0} \|A_h R_\lambda f - (\lambda R_\lambda f - f)\|_\infty &= \lim_{h\downarrow 0} \|\lambda R_\lambda (B_h f - f) - (B_h f - f)\|_\infty \\ &\le \lim_{h\downarrow 0} (\|\lambda R_\lambda\|_\infty + 1) \|B_h f - f\|_\infty = 0. \end{aligned}$$

In other words, $\mathsf{A} R_\lambda f = \lambda R_\lambda f - f \in C_0(E)$. It follows that $R_\lambda f \in \mathcal{D}(\mathsf{A})$ and that $(\lambda \mathbf{I} - \mathsf{A}) R_\lambda f = f$ on $C_0(E)$. QED

The theorem states for $\lambda > 0$ that $R_\lambda = (\lambda \mathbf{I} - \mathsf{A})^{-1}$, and so the generator determines the resolvents. By uniqueness of Laplace transforms, the resolvents determine the semigroup $(P_t)_t$ (this is not a triviality, but a seep theorem called the Hille-Yosida theorem). This shows that the generator determines the semigroup. In other words, two Feller-Dynkin processes with the same generator have the same semigroup.

Corollary 3.4.11 *The generator determines the semigroup.*

The preceding theorem also shows that for all $\lambda > 0$, the generator is given by $\mathsf{A} = \lambda \mathbf{I} - R_\lambda^{-1}$. This gives us a method for explicit computation. The domain of the generator is not always easily determined. The following theorem is helpful, but we need to introduce a new concept first.

The definition of generator implies that if $f \in \mathcal{D}(\mathsf{A})$ and there exists $x \in E$, such that $f(y) \le f(x)$ for all $y \in E$, then $\mathsf{A} f(x) \le 0$. An operator with this property is said to satisfy the *maximum principle*.

Theorem 3.4.12 *Suppose that the generator A extends to $A' : D' \to C_0(E)$, where $\mathcal{D}(\mathsf{A}) \subseteq D' \subseteq C_0(E)$, D' is a linear space and A' satisfies the maximum principle. Then $D' = \mathcal{D}(\mathsf{A})$.*

Proof. In the first part of the proof we show that for a linear space $D \subseteq C_0(E)$ any operator $B : D \to C_0(E)$ is dissipative, in other words

$$\lambda \|f\|_\infty \le \|(\lambda \mathbf{I} - B)f\|_\infty, \quad f \in D, \lambda > 0.$$

To prove this, let $x \in E$ be such that $|f(x)| = \|f\|_\infty$, and define $g(y) = f(y) \operatorname{sgn} f(x)$. Then $g \in D$ since D is linear, and $g(y) \le g(x)$ for all $y \in E$. By the maximum principle $Bg(x) \le 0$. It follows that

$$\lambda \|f\|_\infty = \lambda g(x) \le \lambda g(x) - Bg(x) \le \|(\lambda \mathbf{I} - B)g\|_\infty = \|(\lambda \mathbf{I} - B)f\|_\infty,$$

as claimed.

For the proof of the theorem, take $f \in D'$ and define $g = (\mathbf{I} - A')f$. We claim that $f \in \mathcal{D}(\mathsf{A})$. By the first part of the proof A' is dissipative, hence

$$\|f - R_1 g\|_\infty \le \|(\mathbf{I} - A')(f - R_1 g)\|_\infty.$$

By Theorem 3.4.10 we have $(\mathbf{I} - A)R_1 = \mathbf{I}$ on $C_0(E)$. Since $R_1 g \in \mathcal{D}(\mathsf{A})$, one has $(\mathbf{I} - A')R_1 g = (\mathbf{I} - A)R_1 g$. Hence

$$(\mathbf{I} - A')(f - R_1 g) = g - (\mathbf{I} - A')R_1 g = 0.$$

It follows tat $f = R_1 g$, whence $f \in \mathcal{D}(\mathsf{A})$. QED

Example 3.4.13 For the BM process we have that $\mathcal{D}(\mathsf{A}) = C_0^2(\mathbf{R})$, the space of $C_0(\mathbf{R})$ functions that have continuous first and second derivatives. For $f \in \mathcal{D}(\mathsf{A})$ we have $\mathsf{A}f = f''/2$.

The procedure to prove this, 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 $\mathcal{D}(\mathsf{A}) \subseteq C_0^2(\mathbf{R})$. Hence A' defined by $\mathsf{A}'h = \frac{1}{2}h''$, $h \in \mathsf{C}_0^2(\mathbf{R})$ is an extension of A on $\mathcal{D}(\mathsf{A}') = \mathsf{C}_0^2(\mathbf{R})$.

We claim that A' satisfies the maximum principle. Let $h \in \mathcal{D}(\mathsf{A}') = \mathsf{C}_0^2(\mathbf{R})$, and choose $x \in \mathbf{R}$ with $h(x) \ge h(y)$, $\forall y \in \mathbf{R}$. Then $h'(x) = 0$ and so $h''(x) \le 0$ by a second order Taylor expansion. By virtue of Theorem 3.4.12, $C_0^2(\mathbf{R}) = \mathcal{D}(\mathsf{A})$.

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,

$$h'(x) = \int_{\mathrm{R}} f(y) \sqrt{2\lambda}\, r_\lambda(x,y) \operatorname{sgn}(y-x)\,dy.$$

The integrand is not continuous in $y = x$! We have to show that h' is differentiable. For $\delta > 0$ we have

$$\begin{aligned} h'(x+\delta) - h'(x) &= \int_{y<x} f(y)\sqrt{2\lambda}\Big(\frac{r_\lambda(x,y) - r_\lambda(x+\delta,y)}{\delta}\Big)dy \\ &\quad + \int_{y>x+\delta} f(y)\sqrt{2\lambda}\Big(\frac{r_\lambda(x+\delta,y) - r_\lambda(x,y)}{\delta}\Big)dy \\ &\quad - \int_{x\le y\le x+\delta} f(y)\sqrt{2\lambda}\Big(\frac{r_\lambda(x+\delta,y) + r_\lambda(x,y)}{\delta}\Big)dy \end{aligned}$$

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,

$$\begin{aligned} \int_x^{x+\delta}\cdots &= \int_{u=0}^{\delta}\Big(\frac{\exp^{\sqrt{2\lambda}(u-\delta)} + \exp^{-\sqrt{2\lambda}u}}{\delta}\Big)f(x+u)du \\ &= \int_{u=0}^{1}\mathbf{1}_{\{u\le\delta\}}\Big(\frac{\exp^{\sqrt{2\lambda}(u-\delta)} + \exp^{-\sqrt{2\lambda}u}}{\delta}\Big)f(x+u)du \\ &\to 2f(x), \quad \delta\downarrow 0, \end{aligned}$$

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.

Example 3.4.14 (Ornstein-Uhlenbeck process) Consider the Orstein-Uhlenbeck process in Example 3.1.7 (B). The generator is given by

$$\mathsf{A}f(x) = \tfrac{1}{2}\sigma^2 f''(x) - \alpha x f'(x), \quad x \in \mathbf{R},$$

for $f \in C_0^2(\mathbf{R})$ (cf. Exercise 3.18).

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.

The next lemma gives condition under which a function ϕ of a Feller-Dynkin process is Feller-Dynkin, and it provides a relation between the corresponding generators.

Lemma 3.4.15 *Let X be a Feller-Dynkin process with state space $(E,\mathcal{E})$, initial distribution ν and transition function $(P_t)_t$. 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 3.4.16 In Example 3.1.10 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 3.15.

3.4.3 Applications to the generator

Generators provide an important link between Feller processes and martingales.

Theorem 3.4.17 *For every $f \in \mathcal{D}(\mathsf{A})$ and initial 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.

Proof. Since f and $\mathsf{A}f$ are in $C_0(E_\delta)$, M_t^f is integrable for every $t \geq 0$. For $s \leq t$

$$\mathsf{E}_\nu(M_t^f \mid \mathcal{F}_s) = M_s^f + \mathsf{E}_\nu\big(f(X_t) - f(X_s) - \int_s^t \mathsf{A}f(X_u)du \mid \mathcal{F}_s\big).$$

By the Markov property, the conditional expectation on the right-hand side equals

$$\mathsf{E}_{X_s}\Big(f(X_{t-s}) - f(X_0) - \int_0^{t-s} \mathsf{A}f(X_u)du\Big).$$

But for every $x \in E$

$$\mathsf{E}_x\Big(f(X_{t-s}) - f(X_0) - \int_0^{t-s} \mathsf{A}f(X_u)du\Big) = P_{t-s}f(x) - f(x) - \int_0^{t-s} P_u\mathsf{A}f(x)du = 0,$$

by Fubini's theorem and Lemma 3.4.8 (iii). QED

This gives rise to the following extremely convenient formula.

Corollary 3.4.18 (Dynkin's formula) *For every $f \in \mathcal{D}(\mathsf{A})$ and every 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 3.4.17 and the optional stopping 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 $\mathsf{A}f \in C_0(E_\delta)$, we have $\|\mathsf{A}f\|_\infty < \infty$ and so

$$\Big| \int_{\tau\wedge n} \mathsf{A}f(X_s)ds \Big| \leq \|\mathsf{A}f\|_\infty \tau.$$

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. Call a point $x \in E_\delta$ *absorbing* if for all $t \geq 0$ it holds that $P_t^\delta(x, \{x\}) = 1$. This means that if the process starts at an absorbing point x, it never leaves x (cf. Exercise 3.16).

Example 3.4.19 Let X be a canonical cadlag Feller-Dynkin process on the countable state space E_δ, equipped with the σ-algebra $\mathcal{E}_\delta$, where $\mathcal{E}$ is generated by the one-point sets of E.

Suppose that $\mathbf{1}_{\{x\}} \in \mathcal{D}(\mathsf{A})$ for all $x \in E$. The following extension of Exercise 3.11 holds: there exists an $E_\delta \times E_\delta$ matrix Q, with $Q(x,x) \leq 0 \leq Q(x,y)$ for all $x, y \in E_\delta$, $y \neq x$, such that $\mathsf{A}\mathbf{1}_{\{x\}}(y) = Q(x,y)$. In particular, $\sum_{y\neq x} Q_{xy} = 0$ for all $x, y \in E_\delta$ and $Q(\delta,\delta) = 0$.

Define $\sigma_x = \inf\{t > 0 \,|\, X_t \neq x\}$. In Exercises 3.16 and 3.24 you are asked to show that $\mathsf{P}_x\{\sigma_x > t\} = e^{-at}$, $t \geq 0$, for some $a \in [0,\infty]$. Further, if $a \in (0,\infty)$, then x is called a *holding point*, and the holding time (in state x) is a (negative) exponentially distributed random variable with parameter a. Moreover, $\mathsf{P}_x\{X_{\sigma_x} = x\} = 0$, so that a holding point x can only be left by a jump.

With these results, one can give an interpretation to the elements of Q: if x is a holding point then $a = |Q(x,x)|$, so that $\mathsf{E}_x\sigma_x = |Q(x,x)|^{-1}$. Further, $\mathsf{P}_x\{X_{\sigma_x} = y\} = Q(x,y) \cdot |Q(x,x)|^{-1}$. If x is absorbing, then $\mathsf{E}_x\sigma_x = \infty$.

Example 3.4.20 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 want 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.

Consider the (canonical cadlag) 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. 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 2.28).

Characteristic operator We will now give a probabilistic intepretation of the generator. For $r > 0$, define the stopping

$$\eta_r = \inf\{t \geq 0 \mid \|X_t - X_0\| > r\}. \tag{3.4.5}$$

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 3.4.21 *If $x \in E$ is not absorbing, then $\mathsf{E}_x \eta_r < \infty$ for all $r > 0$ sufficiently small.*

Proof. Let $B_x(\epsilon) = \{y \mid \|y - x\| \leq \epsilon\}$ be the closed ball of radius ϵ around the point x. If x is not absorbing, then $P_t^\delta(x, B_x(\epsilon)) < p < 1$ for some $t, \epsilon > 0$.

By the Feller-Dynkin property of the semi-group P_t^δ we have that $P_t^\delta(y, \cdot) \to P_t^\delta(x, \cdot)$ as $y \to x$. Hence, by the Portmanteau theorem, the fact that $B_x(\epsilon)$ is closed implies that

$$\limsup_{y \to x} P_t^\delta(y, B_x(\epsilon)) \leq P_t^\delta(x, B_x(\epsilon)).$$

It follows that for all y sufficiently close to x, say $y \in B_x(r)$ for some $r \in (0, \epsilon)$, we have $P_t^\delta(y, B_x(r)) \leq P_t^\delta(x, B_x(r)) < p$. Using the Markov property it is easy to show (cf. Exercise 3.33) that $\mathsf{P}_x(\eta_r \geq nt\} \leq p^n$, $n = 0, 1, \ldots$. Hence,

$$\mathsf{E}_x \eta_r = \int_0^\infty \mathsf{P}_x\{\eta_r \geq s\} ds \leq t \sum_{n=0}^\infty \mathsf{P}_x(\eta_r \geq nt\} \leq \frac{t}{1-p} < \infty.$$

This completes the proof. QED

We can now prove the following alternative description of the generator.

Theorem 3.4.22 *For $f \in \mathcal{D}(\mathsf{A})$ we have $\mathsf{A}f(x) = 0$ if x is absorbing, and otherwise*

$$\mathsf{A}f(x) = \lim_{r \downarrow 0} \frac{\mathsf{E}_x f(X_{\eta_r}) - f(x)}{\mathsf{E}_x \eta_r}. \tag{3.4.6}$$

Proof. If x is absorbing, we have $P_t^\delta f(x) = f(x)$ for all $t \geq 0$ and so $\mathsf{A}f(x) = 0$. For non-absorbing $x \in E_\delta$ 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

$$\begin{aligned} \Big| \frac{\mathsf{E}_x f(X_{\eta_r}) - f(x)}{\mathsf{E}_x \eta_r} - \mathsf{A}f(x) \Big| &\leq \frac{\mathsf{E}_x \int_0^{\eta_r} |\mathsf{A}f(X_s) - \mathsf{A}f(x)| ds}{\mathsf{E}_x \eta_r} \\ &\leq \sup_{\|y-x\| \leq r} |\mathsf{A}f(y) - \mathsf{A}f(x)|. \end{aligned}$$

This completes the proof, since $\mathsf{A}f \in C_0(E_\delta)$. QED

The operator defined by the right-handside of (3.4.6) is called the *characteristic operator* of the Markov process X. Its domain is simply the collection of all functions $f \in C_0(E_\delta)$ for which the limit in (3.4.6) exists. The theorem states that for Feller-Dynkin processes the characteristic operator extends the infinitesimal generator. It is easy to see that the characteristic operator satisfies the maximum principle. By Theorem 3.4.12, the characteristic operator and the generator coincide.

3.5 Regularisation

3.5.1 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 $(\Omega, \mathcal{F}) = (E_\delta^{\mathbf{R}_+}, \mathcal{E}_\delta^{\mathbf{R}_+})$, 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.

We need the following lemma, which will allow us to use the regularisation results for supermartingales from the preceding chapter.

Lemma 3.5.1 *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

$$e^{-\lambda t} P_{t-s} R_\lambda f(x) \le e^{-\lambda} R_\lambda f(x), \quad x \in E. \tag{3.5.1}$$

This is a straightforward calculation (cf. Exercise 3.19). QED

Theorem 3.5.2 *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 3.20). By the second statement of Corollary 3.4.3, the class

$$\mathcal{H}' = \{nR_n h \,|\, h \in \mathcal{H}, n \in \mathbf{Z}_+\}$$

has the same property. The proof of Theorem 2.3.2 can be adapted to show that the set

$$\Omega_{h'} = \{\omega \,|\, \lim_{q\downarrow t} h'(X_r)(\omega), \lim_{q\uparrow t} h'(X_r)(\omega) \text{ exist as finite limits for all } t > 0\} \qquad (3.5.2)$$

is $\mathcal{F}_\infty^X$-measurable. By virtue of Lemma 3.5.1 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 3.21, 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

$$\begin{aligned}
\mathsf{E}_\nu f(X_t) g(Y_t) &= \lim_{q\downarrow t} \mathsf{E}_\nu f(X_t) g(X_q) \\
&= \lim_{q\downarrow t} \mathsf{E}_\nu \mathsf{E}_\nu (f(X_t) g(X_q) \,|\, \mathcal{F}_t^X) \\
&= \lim_{q\downarrow t} \mathsf{E}_\nu f(X_t) P_{q-t} g(X_t).
\end{aligned}$$

By Corollary 3.4.3, $P_{q-t} g(X_t) \to g(X_t)$, $q \downarrow t$, 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 3.22 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 3.21, 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 $*$).

3.5.2 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 3.5.3 *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

$$\mathbf{1}_{\{B\}} = \mathsf{E}_\nu(\mathbf{1}_{\{B\}} \mid \mathcal{F}_{t+}^\nu) \overset{\mathsf{P}_\nu-\text{a.s.}}{=} \mathsf{E}_\nu(\mathbf{1}_{\{B'\}} \mid \mathcal{F}_{t+}^\nu).$$

It therefore suffices to show (explain!) that

$$\mathsf{E}_\nu(\mathbf{1}_{\{B'\}} \mid \mathcal{F}_{t+}^\nu) = \mathsf{E}_\nu(\mathbf{1}_{\{B'\}} \mid \mathcal{F}_t^\nu), \mathsf{P}_\nu - \text{a.s.}$$

To this end, define

$$\mathcal{S} = \{A \in \mathcal{F}_\infty^X \mid \mathsf{E}_\nu(\mathbf{1}_{\{A\}} \mid \mathcal{F}_{t+}^\nu) = \mathsf{E}_\nu(\mathbf{1}_{\{A\}} \mid \mathcal{F}_t^\nu), \mathsf{P}_\nu - \text{a.s.}\}.$$

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 are the finite cylinders. Let A be a finite cylinder, 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

$$\mathbf{1}_{\{A\}} = \prod_{k=1}^n \mathbf{1}_{\{A_k\}}(X_{t_k}). \tag{3.5.3}$$

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

$$\mathsf{E}(Z \mid \mathcal{F}_t^\nu) = \mathsf{E}(Z \mid \mathcal{F}_{t_+}^\nu), \quad \mathsf{P}_\nu - \text{a.s.} \tag{3.5.4}$$

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 (Corollary 3.4.3), $\|g^h - g^0\|_\infty \to 0$, 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 2.2.15

$$\mathsf{E}_\nu(Z \mid \mathcal{F}_{t+h}^\nu) \to \mathsf{E}_\nu(Z \mid \mathcal{F}_{t+}^\nu), \quad \mathsf{P}_\nu - \text{a.s..}$$

This implies (3.5.4).

The only thing left to prove is that we can replace Z by $\mathbf{1}_{\{A\}}$ from (3.5.3). 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 $Z = \prod_{i=1}^n f_i^m(X_{t_i})$ (3.5.4) holds. Use monotone convergence, to obtain that (3.5.4) holds for $Z = \mathbf{1}_{\{A\}}$ given in (3.5.3). Hence, the d-system $\mathcal{S}$ contains all finite cylinder sets, 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 3.5.4 *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 3.5.5 *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 3.5.6 *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 3.5.7 (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 3.5.4 and 3.5.5 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

$$\int_A Z \circ \theta_t d\mathsf{P}_\nu = \int_A \mathsf{E}_{X_t} Z d\mathsf{P}_\nu. \tag{3.5.5}$$

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 (3.5.5) 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 (3.5.5) 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 agan 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 3.5.7 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 3.5.8 (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.*

$$\mathsf{E}_\nu(Z \circ \theta_\tau \,|\, \mathcal{F}_\tau) = \mathsf{E}_{X_\tau}(Z). \tag{3.5.6}$$

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 1.6.13 and 1.6.14. 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 (3.5.6) holds for Z an $\mathcal{F}^X_\infty$-measurable random variable. A similar reasoning as in the proof of Theorem 3.5.3 shows (check yourself) that we can restrict to showing (3.5.6) 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 1.17. 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 2.2.16 the left-handside converges P_ν-a.s. to

$$\mathsf{E}_\nu(\prod_i f_i X_{t_i}) \circ \theta_\tau \,|\, \mathcal{F}_\tau),$$

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 (3.5.6) 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 3.5.7.

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

$$\begin{aligned}
\mathsf{E}_\nu(Z'' \circ \theta_\tau - Z' \circ \theta_\tau) &= \mathsf{E}_\nu \mathsf{E}_\nu(Z'' \circ \theta_\tau - Z' \circ \theta_\tau \,|\, \mathcal{F}_\tau) \\
&= \mathsf{E}_\nu \mathsf{E}_{X_\tau}(Z'' - Z') \\
&= \int \mathsf{E}_x(Z'' - Z')\mu(dx) = \mathsf{E}_\mu(Z'' - Z') = 0.
\end{aligned}$$

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.

$$\mathsf{E}(Z' \circ \theta_\tau \,|\, \mathcal{F}_\tau) \leq \mathsf{E}(Z \circ \theta_\tau \,|\, \mathcal{F}_\tau) \leq \mathsf{E}(Z'' \circ \theta_\tau \,|\, \mathcal{F}_\tau).$$

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

3.6 Regenerative processes and killed Feller processes

3.6.1 Regenerative processes

3.6.2 Killed Feller processes

3.7 Exercises

Exercise 3.1 Consider the Ornstein Uhlenbeck process in example 3.1.7(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 (see BN section 5 version April 20 or later). If $X_0 \stackrel{\mathsf{d}}{=} \mathsf{N}(0, \sigma^2/2\alpha)$, show that $X_t \stackrel{\mathsf{d}}{=} \mathsf{N}(0, \sigma^2/2\alpha)$ (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 Complete the proof of Lemma 3.1.9.

Exercise 3.3 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 3.4 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 3.5 Prove that the probability measures $\mu_{t_1,\ldots,t_n}$ defined in the proof of Corollary 3.2.2 form a consistent system.

Exercise 3.6 Work out the details of the proof of Lemma 3.2.3.

Exercise 3.7 Show for the Poisson process X with initial distribution $\nu = \delta_x$ in Example 3.1.8, that X is a Markov process w.r.t. the natural filtration, with the transition function specified in the example.

Exercise 3.8 Show Corollary 3.3.6 that canonical Brownian motion has the strong Markov property.

Exercise 3.9 Prove Lemma 3.3.12 and Theorem 3.3.13.

Exercise 3.10 Show that the Markov jump process from Example 3.2.5 is a Feller-Dynkin process iff $PC_0(E) \subseteq C_0(E)$. Give an example of a Markov jump process that is not a Feller-Dynkin process.

Exercise 3.11 Let X be a Feller-Dynkin Markov process on a countable state space E, without accumulation points. Suppose that $\mathbf{1}_{\{x\}} \in \mathcal{D}(\mathsf{A})$ for any $x \in E$. Show that the generator A can be represented by a matrix Q, with $\sum_y Q(x,y) \leq 0$, $-\infty < Q(x,x) \leq 0$, $Q(x,y) \geq 0$, $y \neq x$, $x \in E$.

Exercise 3.12 Prove the claims made in Example 3.4.4. Hint: to derive the explicit expression for the resolvent kernel it is needed to calculate integrals of the form

$$\int_0^\infty \frac{e^{-a^2t-b^2/t}}{\sqrt{t}}dt.$$

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 3.13 Prove the validity of the expression for the resolvent of the Feller-Dynkin Markov jump process given in Example 3.4.5.

Exercise 3.14 Show that the generator of the Poisson process is given by

$$\mathsf{A}(x,y) = \begin{cases} \lambda, & y = x+1 \\ -\lambda, & y = x, \end{cases}$$

for $x \in \mathbf{Z}_+$.

Exercise 3.15 Prove Lemma 3.3.15. Use this lemma to show the validity of the expression for the generator of W_t^2, with W_t a standard BM, given in Example 3.4.16.

Exercise 3.16 Let X be a canonical, right-continuous Markov process with state space $(E, \mathcal{E})$ and for $x \in E$. Consider the stopping time $\sigma_x = \inf\{t > 0 \mid X_t \neq x\}$.

i) 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 3.2.5) 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 3.17 Branching model in continuous time Let $E = \mathbf{Z}_+ = \{0, 1, 2, \ldots\}$. 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, where $\mathcal{E}$ is the collection of all subsets of E. Assume that it is a Markov jump process.

i) Argue that the generator Q is given by

$$Q(i,j) = \begin{cases} \lambda i & j = i+1 \\ -(\lambda+\mu)i, & j = i \\ \mu i, & j = i-1, i > 0. \end{cases}$$

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 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.?

Exercise 3.18 Show that the generator of the Ornstein-Uhlenbeck process (cf. Example 3.1.7 (B) and 3.4.14) is given by

$$\mathsf{A}f(x) = \tfrac{1}{2}\sigma^2 f''(x) - \alpha x f'(x), \quad x \in \mathbf{R}, f \in C_0^2(\mathbf{R}).$$

You may use that expression for the generator of Brownian motion derived in Example 3.4.13. 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 3.19 Prove (3.5.1) in the proof of Lemma 3.5.1.

Exercise 3.20 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 3.21 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 3.22 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 3.23 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 3.24 Consider the situation of Exercise 3.16. Suppose in addition that X has the strong Markov property. 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 $\{X_{\sigma_x} = x\} \subseteq \{\sigma_x \circ \theta_{\sigma_x} = 0\}$.

ii) Using the strong Markov property, show that

$$\mathsf{P}_x\{X_{\sigma_x} = x\} = \mathsf{P}_x\{X_{\sigma_x} = x\}\mathsf{P}_x\{\sigma_x = 0\}.$$

iii) Conclude that $\mathsf{P}_x\{X_{\sigma_x} = x\} = 0$, 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 3.25 Show that a Markov process with stationary and independent increments is a Feller-Dynkin process.

Exercise 3.26 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 3.22.

Exercise 3.27 Prove Corollary 3.3.7 and Lemma 3.3.8.

Exercise 3.28 Show for Example 3.3.9 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 3.29 Show that the maps ϕ and ψ in the proof of Theorem 3.3.14 are Borel measurable.

Exercise 3.30 Derive the expression for the joint density of BM and its running maximum given in Corollary 3.3.15.

Exercise 3.31 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 3.32 Prove Corollary 3.3.16

Exercise 3.33 In the proof of Lemma 3.4.21, show that $\mathsf{P}\{\eta_r \geq nt\} \leq p^n$ for $n = 0, 1, \ldots$.

Exercise 3.34 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 $f \in C_0^2(\mathbf{R})$, where α is an arbitrary but fixed continuous, bounded function on $\mathbf{R}$. Suppose that there exists a $C_0^2(\mathbf{R})$ function $f \not\equiv 0$, such that

$$\mathsf{A}f(x) = 0, \quad x \in \mathbf{R}.$$

Then the martingale M_t^f has a simpler structure, namely $M_t^f = f(X_t) - f(X_0)$.

i) Show that in this case 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 $\tau = \inf\{t > 0 \mid X_t \in (-\infty, a] \cup [b, \infty)\}$. 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 $p_x = \frac{f(x)-f(a)}{f(b)-f(a)}$, $x \in (a, b)$.

iii) Let X be a real-valued canonical, cadlag Feller-Dynkin process, such that $X_t = X_0 + bt + \sigma W_t$, where X_0 and $(W_t)_t$ are independent, and $(W_t)_t$ a standard BM. Show that the generator A has domain $\mathcal{D}(\mathsf{A}) = C_0^2(\mathbf{R})$ and is given by

$$\mathsf{A}f = bf' + \tfrac{1}{2}\sigma^2 f''$$

for $f \in C_0^2(\mathbf{R})$ (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 $bf' + \sigma^2 f''/2 = 0$. 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 3.1.5 (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 3.18 on the generator of the Ornstein-Uhlenbeck process. See also hint of (iii).

Exercise 3.35 **i)** Use the characteristic operator to compute the generator of the Brownian motion process (cf. Example 3.1.5 (A)). You may use the results of Exercises 2.28 and 2.29.

ii) Determine generator and domain for the process in 3.34 (iii) by means of the characteristic operator (without using the generator for BM).

Exercise 3.36 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 3.37 (Continuation of Exercise 3.36) 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 $(W_t)_t$. (Think about how to construct this). Notice that X is strong Markov.

We would like to show that the generator is defined by

$$\mathsf{A} f(x) = \tfrac{1}{2} \Delta f(x), \tag{3.7.1}$$

where $\Delta f(x) = \sum_{i=1}^d \frac{\partial^2}{\partial x_i^2} f(x)$ is the Laplacian of f, with domain $\mathcal{D}(\mathsf{A}) \supset C_0^2(\mathbf{R}^d)$. We again want to use the characteristic operator. To this end, define for $r > 0$

$$\tau_r = \inf\{t \geq 0 \,|\, \|X_t - X_0\| \geq r\}.$$

i) Argue that τ_r is a finite $(\mathcal{F}_t^X)_t$-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 $\{y \,|\, \|y - x\| = r\}$.

ii) Show the validity of (3.7.1) for $f \in C_0^2(\mathbf{R}^d)$ (use the characteristic operator). Argue that this implies $\mathcal{D}(\mathsf{A}) \supset C_0^2(\mathbf{R}^d)$.

iii) For $0 < a < \|x\| < b$, show that

$$\mathsf{P}_x\{\tau_a < \tau_b\} = \begin{cases} \frac{\log b - \log \|x\|}{\log b - \log a}, & d = 2 \\ \frac{\|x\|^{2-d} - b^{2-d}}{a^{2-d} - b^{2-d}}, & d \geq 3. \end{cases}$$

Hint: a similar procedure as in Exercise 3.34.

iv) Compute $\mathsf{P}_x\{\tau_a < \infty\}$ for x with $a < \|x\|$.

Exercise 3.38 (cf. Example 3.5.13) Let X be a canonical Feller-Dynkin process on the countable state space E_δ, $E \subset \mathbf{Z}^d$, equipped with the σ-algebra $\mathcal{E}_\delta$, where $\mathcal{E}$ is generated by the one-point sets of E. We assume that $E \subset \mathbf{Z}^d$, for some $d < \infty$. You may assume the results in Exercises 3.16 and 3.24.

i) Argue that no $x \in E$ can be a regular point.

In other words $\mathsf{P}_x\{\sigma_x > t\} = e^{-at}$, for some $a \in [0, \infty)$, where $\sigma_x = \inf\{t > 0 \,|\, X_t \neq x\}$. Hence $\mathsf{E}_x \sigma_x = a^{-1} \in (0, \infty]$.

ii) Use the characteristic operator to show that $\mathbf{1}_{\{x\}} \in \mathcal{D}(\mathsf{A})$ for all $x \in E_\delta$.

iii) Show that there exists an $E_\delta \times E_\delta$ matrix Q, with $Q(x,x) \le 0 \le Q(x,y)$ for all $x, y \in E_\delta$, $y \neq x$, such that $\mathsf{A}\mathbf{1}_{\{x\}}(y) = Q(x,y)$. In particular, show that $\sum_y Q(x,y) = 0$ for all $x \in E_\delta$ and $Q(\delta,\delta) = 0$.

Exercise 3.39 (Continuation of Exercise 3.38) i) Show for each $x \in E_\delta$ that $\mathsf{E}_x \sigma_x = 1/|Q(x,x)|$ and that $\mathsf{P}_x\{X_{\sigma_x} = y\} = Q(x,y)/|Q(x,x)|$, if $Q(x,x) \neq 0$.

ii) Compute the generator of the Poisson process and domain. Determine the distribution of X_{σ_x} conditional on $X_0 = x$. In other words, compute $\mathsf{P}_x\{X_{\sigma_x} = y\}$, $y \in \mathbf{Z}_+$.

iii) Let $E = \mathbf{Z}$. Let be given the Q-matrix with non-zero elements $Q(x,x) = -Q(x,x+1) = -2^x$. Can one construct a canonical right-continuous Feller-Dynkin process X, such that $\mathsf{A}f(x) = Qf(x)$, $f \in C_0(\mathbf{Z}_+)$? If yes, what is $\lim_{t\to\infty} X_t$? Hint: use that $\sum_{x=0}^\infty 2^{-x} < \infty$.