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 $m = 2$

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

In step (3.1.1) to the next, we use discreteness of the state space as well linearity of conditional expectations. In fact we have proved a more general version of the Markov property to hold. To formulate it, we need some more notation. But first we will move on to Markov chains on a general measurable space.

The one point sets need not be measurable in general space. The notion of a stochastic matrix generalises to the notion of a transition kernel.

Definition 3.1.1 Let $(E, \mathcal{E})$ be a measurable space. A *transition kernel* on E is a map $P : E \times \mathcal{E} \to [0, 1]$ such that

i) for every $x \in E$, the map $B \mapsto P(x, B)$ is a probability measure on $(E, \mathcal{E})$,

ii) for every $B \in \mathcal{E}$, the map $x \mapsto P(x, B)$ measurable.

Let X be an $(E, \mathcal{E})$ valued stochastic process defined on some underlying probability space $\Omega, \mathcal{F}, \mathsf{P}_\nu)$. Then X is a Markov chain with initial distribution ν if (i) $\mathsf{P}_\nu\{X_0 \in B\} = \nu(B)$ for all $B \in \mathcal{E}$; (ii) if there exists a transition kernel P such that the Markov property holds:

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

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

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

In other words, the linear operators $(P_t)_{t\geq 0}$ form a *semigroup* of operators on the space of non-negative (or bounded) functions on E. In the sequel we will not distinguish between this semigroup and the corresponding (homogeneous) transition function on $(E,\mathcal{E})$, since there is a one-to-one relation between the two concepts.

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

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

Then we can apply P_s to this function, yielding the bounded, measurable function P_sgP_tf, with

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

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

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

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 $f: E\to\mathbf{R}$

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

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 (cf. Example 3.1.6).

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 $0 = t_0 < t_1 < \cdots < t_n$, 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

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

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 $\mathcal{F}_s^X$. As the π-system we take

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

Let $f_i = \mathbf{1}_{\{A_i\}}$, then $\prod_{i=0}^n f_i(X_{t_i}) = \mathbf{1}_{\{X_{t_0} \in A_0, \ldots, X_{t_n} \in A_n\}}$ and so, assuming that $t_n < s$

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

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 $j \in \{1,2,3\}$ and n, meaning that the motion is determined by the 3×3 stochastic matrix with all elements equal to $1/3$.

However, $\{X_n\}_n$ is not a Markov chain, since $\mathsf{P}\{X_2 = 1 \,|\, \sigma(X_0, X_1)\} = f(X_0, X_1)$ with

$$f(X_0, X_1) = \begin{cases} 1, & (X_0, X_1) \in \{(1,1), (2,3), (3,2)\}\\ 0, & \text{otherwise.}\end{cases}$$

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

Example 3.1.7 (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 $Y_t \overset{\mathsf{d}}{=} \mathsf{N}(0,t)$ 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 $X_t \xrightarrow{\mathcal{D}} \mathsf{N}(0, \sigma^2/2\alpha)$.

If $X_0 \overset{\mathsf{d}}{=} \mathsf{N}(0, \sigma^2/2\alpha)$ 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 $x \in \mathbf{R}$

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

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 $\mathcal{F}_\infty^X$ 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

$$\mathcal{S} = \{\Gamma \in \mathcal{F}_\infty^X \,|\, x \to \mathsf{E}_x \mathbf{1}_{\{\Gamma\}} \text{ is measurable and } \mathsf{E}_\nu \mathbf{1}_{\{\Gamma\}} = \int \mathsf{E}_x \mathbf{1}_{\{\Gamma\}} \nu(dx)\}.$$

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 $\mathcal{F}_\infty^X = \mathcal{E}^{\mathrm{R}_+}$. So if we can show that $\mathcal{G} \subset \mathcal{S}$, then by BN Lemma 3.4 $\mathcal{F}_\infty^X \subset \mathcal{S}$. But this follows from Lemma 3.1.5.

It follows that the statement of the lemma is true for $Z = \mathbf{1}_{\{\Gamma\}}$, $\Gamma \in \mathcal{F}_\infty^X$. Apply the standard machinery to obtain the validity of the lemma for $\mathcal{F}_\infty^X$-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* $\theta_t : E^{\mathrm{R}_+} \to E^{\mathrm{R}_+}$ 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 we introduced new notation: $\mathsf{E}_{X_t} Z$ is a random variable with value $\mathsf{E}_x Z$ on the event $\{X_t = x\}$. By Lemma 3.2.3 this is a measurable function of X_t.

Proof. Fix an initial probability measure ν. We 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

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

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

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

$$\begin{aligned}
&= \nu \mathbf{1}_{\{A_0\}} P_{t_1-t_0} \cdots P_{t_n-t_{n-1}} \mathbf{1}_{\{A_n\}} P_{t+s_1-t_n} \mathbf{1}_{\{B_1\}} \cdots P_{s_m-s_{m-1}} \mathbf{1}_{\{B_m\}} \\
&= \nu \mathbf{1}_{\{A_0\}} P_{t_1-t_0} \cdots P_{t_n-t_{n-1}} \mathbf{1}_{\{A_n\}} P_{t-t_n} P_{s_1} \mathbf{1}_{\{B_1\}} \cdots P_{s_m-s_{m-1}} \mathbf{1}_{\{B_m\}} \\
&= \nu \mathbf{1}_{\{A_0\}} P_{t_1-t_0} \cdots P_{t_n-t_{n-1}} \mathbf{1}_{\{A_n\}} P_{t-t_n} f \\
&= \mathsf{E}_\nu \prod_{j=0}^{n} \mathbf{1}_{\{A_j\}}(X_{t_j}) f(X_t) = \mathsf{E}_\nu \mathbf{1}_{\{A\}} f(X_t),
\end{aligned}$$

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.*

Without the required continuity properties, the strong Markov property may fail, as illustrated in Example 3.3.15. 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 3.12). QED

The strong Markov property has interesting consequences for right-continuous canonical Markov processes, X with so-called stationary and independent increments. This means that $X_t - X_s$ is independent of $\mathcal{F}_s$ for $s \leq t$, and for each initial distribution ν, the P_ν-distribution of $X_t - X_s$ only depends on the difference $t - s$, and is independent of ν. In other words: the P_ν-distribution of $X_t - X_s$ and the P_μ distribution $X_{t-s} - X_0$ are equal for all initial distributions ν and μ. 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 3.3.9 *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 $\mathcal{F}_\tau^X$ 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

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

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

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

By independence of the increments

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

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 3.3.10 *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 3.34, 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 3.3.9 implies that for every $(\mathcal{F}_t^X)_t$-stopping time τ, the proces $(X_{\tau+t}-X_\tau)_t$ is a BM. This can be used to prove an interesting ratio limit result (originally derived by Cyrus Derman 1954).

To this end, let $A\in\mathcal{B}$ be a bounded set. Define $\mu(A,\tau)=\lambda\{t\leq\tau : X_t\in A\}$, where λ is the Lebesgue measure (on $(\mathbf{R},\mathcal{B})$ and τ a finite $(\mathcal{F}_t^X)_t$-stopping time, w.r.t P_0. Denote $\tau_0'=\inf\{t>0\,|\,t\geq\tau_1,X_t=0\}$.

Lemma 3.3.11 **i)** *$\mu(A,\tau)$ is a measurable function on $(\Omega,\mathcal{F})$.*

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

iii) *$\mu'(A)=\mathsf{E}_0\mu(A,\tau_0')=2\lambda(A)$.*

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

$$\begin{aligned}\mathsf{E}_0\mu(A,\tau_1) &= \mathsf{E}_0\int_0^\infty\mathbf{1}_{\{A\}}(X_t)\mathbf{1}_{\{[t,\infty)\}}(\tau_1)dt=\int_\Omega\int_0^\infty\mathbf{1}_{\{A\}}(X_t)\mathbf{1}_{\{[t,\infty)\}}(\tau_1)dt\,d\mathsf{P}_0\\ &= \int_0^\infty\int_\Omega\mathbf{1}_{\{A\}}(X_t)\mathbf{1}_{\{[t,\infty)\}}(\tau_1)d\mathsf{P}_0\,dt\end{aligned}$$

$$
\begin{aligned}
&= \int_0^\infty \mathsf{P}_0\{X_t \in A, t \le \tau_1\} dt \\
&= \int_0^\infty \int_A w(t,x) d\lambda(x) dt \\
&= \int_A \int_0^\infty w(t,x) dt d\lambda(x),
\end{aligned}
$$

where

$$
w(t,x) = \begin{cases} \frac{1}{\sqrt{2\pi t}} \Big(e^{-x^2/2t} - e^{-(x-2)^2/2t} \Big), & x \le 1 \\ 0, & x > 1. \end{cases}
$$

This follows from

$$
\mathsf{P}_0\{X_t \in A \cap [-\infty, 1], \tau_1 \le t\} = \mathsf{P}_2\{X_t \in A \cap [-\infty, 1]\}.
$$

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

$$
w(t,x) = -\mathbf{1}_{\{x \le 1\}} \int_{x-2}^{x} \frac{u}{t^{3/2}\sqrt{2\pi}} e^{-u^2/2t} du,
$$

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 $\int_{\mathrm{R}} g(x) dx \ne 0$.

Theorem 3.3.12

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

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

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

Now, for any $T \ge 0$ define

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

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 3.3.13 (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 $Z_t = W_{\tau_x+t} - W_{\tau_x} = W_{\tau_x+t} - x$, $t \geq 0$. By Corollary 3.3.9 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 $t \geq 0$

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

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 3.14). Since $(Y, Z) \overset{\mathsf{d}}{=} (Y, -Z)$, it follows that $W = \phi(Y, Z) \overset{\mathsf{d}}{=} \phi(Y, -Z) = W'$. 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

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

Corollary 3.3.14 *Let W be a standard* BM *and S its running maximum. Then*

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

The pair (W_t, S_t) has joint density

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

with respect to the Lebesgue measure.

Proof. Let W' be the process obtained by reflecting W at the hitting time τ_y. Observe that $S_t \geq y$ if and only if $t \geq \tau_y$. Hence, the probability of interest equals $\mathsf{P}\{W_t \leq x, t \geq \tau_y\}$. On the event $\{t \geq \tau_y\}$ we have $W_t = 2y - W'_t$, and so we have to calculate $\mathsf{P}\{W'_t \geq 2y - x, t \geq \tau_y\}$. Since $x \leq y$, we have $2y - x \geq y$. hence $\{W'_t \geq 2y - x\} \subseteq \{W'_t \geq y\} \subseteq \{t \geq \tau_y\}$. It follows that $\mathsf{P}\{W'_t \geq 2y - x, t \geq \tau_y\} = \mathsf{P}\{W'_t \geq 2y - x\}$. By the reflection principle and symmetry of BM this proves the first statement. The second follows from Exercise 3.15. 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 3.16). This shows in particular that $S_t \overset{\mathsf{d}}{=} |W_t|$ for every $t \geq 0$. This allows to construct an example of a Markov process that lacks the strong Markov property.

Example 3.3.15 (Strong Markov property fails (Yushkevich)) 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 $X_0 \overset{\mathsf{d}}{=} \nu$, independent of W. For initial distribution $\nu = \delta_x$, $x \neq 0$, we therefore do not change the underlying distribution.

X is a Markov process with transition function

$$P_t(x,B) = \begin{cases} \displaystyle\int_B \frac{1}{\sqrt{2\pi t}} e^{-(y-x)^2/2t}\,dy, & x \neq 0 \\ \delta_x\{B\}, & x = 0. \end{cases}$$

Suppose that X has the strong Markov property. Let $\tau = \inf\{t \geq 0 \,|\, X_t = 0\}$. It is a stopping time for X. Let $\sigma = (1-\tau) \vee 0$. Put $Z = \mathbf{1}_{\{\mathbb{R}\setminus\{0\}\}}(X_\sigma)$. Then

$$\mathbf{1}_{\{\tau<\infty\}} \mathsf{E}_{X_\tau} Z = 0.$$

Take initial distribution $\nu = \delta_x$ for some $x > 0$. By the above $\mathsf{E}_x \mathbf{1}_{\{\tau<\infty\}} \mathsf{E}_{X_\tau} Z = 0$. On the other hand,

$$\mathbf{1}_{\{\tau<\infty\}}(Z \circ \theta_\tau) = \mathbf{1}_{\{\mathbb{R}\setminus\{0\}\}}(X_1)\mathbf{1}_{\{\tau\leq 1\}}.$$

Hence

$$\mathsf{E}_x \mathbf{1}_{\{\tau<\infty\}} \mathsf{E}_x(Z\circ\theta_\tau \,|\, \mathcal{F}_\tau) = \mathsf{E}_x \mathbf{1}_{\{\mathbb{R}\setminus\{0\}\}}(X_1)\mathbf{1}_{\{\tau\leq 1\}} = \mathsf{P}_x\{\tau \leq 1\} = \mathsf{P}\{|W_1| \geq x\} > 0.$$

(cf. Exercise 3.13). A contradiction.

We also may derive an explicit expression for the density of the hitting time τ_x. It is easily seen from this expression that $\mathsf{E}\tau_x = \infty$, as was proved by martingale methods in Exercise 2.29 of Chapter 2.

Corollary 3.3.16 *The first time τ_x that the standard* BM *hits the level $x > 0$ has density*

$$t \mapsto \frac{x e^{-x^2/2t}}{\sqrt{2\pi t^3}} \mathbf{1}_{\{t\geq 0\}},$$

with respect to the Lebesgue measure.

Proof. See Exercise 3.17. 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 $\mathcal{Z} = \{t \geq 0 \,|\, W_t = 0\}$ 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 3.3.10 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 Generator of a Markov process

In the case of a discrete Markov chain, the transition kernel completely determines the finite dimensional distribution of the process and hence the distribution of the process. The question arises whether the cases for continuous time processes is analogous: is the distribution of the process determined by one operator? In general the answer is no, but under certain conditions it will be yes.

The question is mainly analytic, and so this section will be analytically oriented. The starting point is a transition function $\{ P_t\}_{t\geq 0}$ on the measurable space $(E, \mathcal{E})$. Let $\mathcal{S}$ be a Banach space of real-valued measurable functions on E, and let $\|\cdot\|$ denote the corresponding norm. By virtue of the Chapman-Kolmogorov equations, $\{ P_t\}_t$ is a so-called *semigroup*.

Definition 3.4.1 $\{ P_t\}_t$ is a strongly continuous semigroup on $\mathcal{S}$, if (i) $P_t : \mathcal{S} \to \mathcal{S}$ is a bounded linear operator for each $t \geq 0$ and (ii) $\lim_{t\downarrow 0} \| P_t f - f\| = 0$ for each $f \in \mathcal{S}$.

A main notion that we will use is *closedness*.

Let $B : \mathcal{D} \to \mathcal{S}$ be a linear operator defined on $\mathcal{D} \subset \mathcal{S}$. The set $\mathcal{G}(B) = \{(f, Bf) \mid f \in \mathcal{D}\} \subset \mathcal{S}\times\mathcal{S}$ is called the graph of B. Note that $\mathcal{S}\times\mathcal{S}$ is a Banach space with norm $\|(f, g)\| = \|f\|+\|g\|$. Then we call B is *closed* iff $\mathcal{G}(B) = \overline{\mathcal{G}(B)}$.

It is easily checked that P_t is a closed linear operator for each $t \geq 0$, if $\{ P_t\}_{t\geq 0}$ is a SCSG($\mathcal{S}$).

Example 3.4.2 Consider Example 3.1.7 (A, B) Brownian motion and the Ornstein-Uhlenbeck process. Let

$$\mathcal{S} = C_0(\mathbf{R}) = \{f : \mathbf{R} \to \mathbf{R} \mid f \text{ continuous with } \lim_{x\to\pm\infty} f(x) = 0\},$$

and let $\|f\| = \sup_{x\in\mathrm{R}} |f(x)|$. One can show that the associated transition functions are strongly continuous semigroups on $C_0(\mathbf{R})$ (cf. Section 3.5).

Example 3.4.3 Consider the Markov jump process in Example 3.2.5. Suppose that there exists a function $F : E \to \mathbf{R}_+$ an a constant $c > 0$ such that $PF(x) \leq cF(x)$, for all $x \in E$. Let

$$\mathcal{S} = \{f : E \to \mathbf{R} \mid \|f\| := \sup_x \frac{|f(x)|}{F(x)} < \infty\}.$$

Then the associated transition function is a strongly continuous semigroup on $\mathcal{S}$, with

$$\| P_t\| \leq \sum_{n=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^n}{n!} c^n = e^{(c-1)\lambda t}.$$

This norm is a *weighted supremum norm* and the Banach space $\mathcal{S}$ is used in much of modern Markov chain theory with applications in queueing and control. The choice $F \equiv 1$ often applies.

The norm of an SCSG($\mathcal{S}$) cannot grow quicker than exponential. This follows from the following lemma.

Lemma 3.4.4 *Let* $\{ P_t\}_{t\geq 0}$ *be a SCSG(*$\mathcal{S}$*). There are constants* $M \geq 1$, $\alpha \geq 0$, *such that* $\| P_t\| \leq Me^{\alpha t}$.

Proof. By the Banach-Steinhaus theorem (cf. BN ...) or the Principle of Uniform Boundedness, it is sufficient to show that, given any $t_0 > 0$, $\sup_{t\leq t_0} \| P_t f\| < \infty$, for all $f \in \mathcal{S}$.

Suppose not, then there exists $f \in \mathcal{S}$ with $\sup_{t\leq t_0} \| P_t f\| = \infty$. By the semigroup property, this implies that for any sequence $\{t_n\}_n$, $t_n \downarrow 0$, $n \to \infty$, that $\sup_{t\leq t_n} \| P_t f\| = \infty$. Hence, there exists a sequence $\{t_n\}_n$, $t_n \downarrow 0$, $n \to \infty$, with $\| P_{t_n} f\| \to \infty$, $n \to \infty$. This contradicts strong continuity. Hence there exists a constant $M \geq 1$, such that $\| P_t\| \leq M$ for $t \leq t_0$.

Finally, put $\alpha = (\log M)/t_0$. Let $t \in [0, \infty)$. Then with $k = \lfloor t/t_0 \rfloor$, we get

$$\| P_t\| = \| P_{kt_0} P_{t-kt_0}\| \leq \| P_{t_0}\|^k \| P_{t-kt_0}\| \leq e^{\alpha t} \cdot M.$$

QED

Corollary 3.4.5 $t \mapsto P_t f$ *is continuous (i.o.w.* $\lim_{s\to t} \| P_t f - P_s f\| = 0$*) for all* $f \in \mathcal{S}$, *and* $t \geq 0$*).*

Proof. We will only prove right-continuity. Let $h > 0$. Then

$$\| P_{t+h} f - P_t f\| \leq \| P_t\| \| P_h f - f\| \to 0, \quad h \downarrow 0.$$

QED

Let next

$$\mathcal{D} = \{f \in \mathcal{S} \,|\, \exists g \in \mathcal{S} \text{ such that } \lim_{t\downarrow 0} \| \frac{P_t f - f}{t} - g\| = 0\}.$$

A priori it is not clear whether $\mathcal{D}$ is even non-empty! For each $f \in \mathcal{D}$ we write

$$Af = g = \lim_{t\downarrow 0} \frac{P_t f - f}{t}.$$

$\mathsf{A} : \mathcal{D} \to \mathcal{S}$ is a (generally unbounded) linear operator, with domain $\mathcal{D}(\mathsf{A}) = \mathcal{D}$, A is called the *generator*.

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

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

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

Example 3.4.6 Brownian motion has $\mathcal{D}(A) = \{f \in C_0(\mathbf{R}) \,|\, f', f'' \text{ exist and are continuous}\}$. Further $\mathsf{A}f = f''/2$ for $f \in \mathcal{D}(\mathsf{A})$. The proof is given in §3.5.

Example 3.4.7 (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 3.19).

Recall that we introduced Brownian motion as a model for the position of a particle. The problem however is that Brownian motion paths are nowhere differentiable, whereas the derivative of the position of a particle is its velocity, hence it should be differentiable. It appears that the Ornstein-Uhlenbeck process is a model for the velocity of a particle, and then its position at time t is given by

$$S_t = \int_0^t X_u du.$$

It can be shown that $\alpha S_{nt}/\sqrt{n} \to W_t$ in distribution, as $n \to \infty$. Hence, for large time scales, Brownian motion may be accepted as a model for particle motion.

Example 3.4.8 Markov jump process in Example 3.2.5 satisfying the conditions in Example 3.4.3. Here we have $\mathcal{D}(\mathsf{A}) = \mathcal{S}$ and $\mathsf{A}f = \lambda(\, P - \mathbf{I})f$ for $f \in \mathcal{S}$.

We next derive the important Kolmogorov forward and backward equations. This requires integrating $\mathcal{S}$-valued functions of t.

Denote by $C_{\mathcal{S}}(a,b) = \{u : [a,b] \to \mathcal{S} \,|\, u \text{ is continuous}\}$. A function $u : [a,b] \to \mathcal{S}$ is said to be (Rieman) integrable over $[a,b]$ if $\lim_{h\to 0} \sum_{k=1}^n u(s_k)(t_k - t_{k-1})$ exists, where $a = t_0 \le s_1 \le t_1 \le \cdots \le t_{n-1} \le s_n \le t_n = b$ and $h = \max_k(t_k - t_{k-1})$, and the limit is independent of the particular sequence $t_0, s_1, \ldots, t_k$. It is then denoted by $\int_a^b u(t)dt$. If a and/or $b = \infty$, the integral is defined as an improper integral. The following result holds.

Integration Lemma

a) If $u \in C_{\mathcal{S}}(a,b)$ and $\int_a^b \|u(t)\| dt < \infty$, then u is integrable over $[a,b]$ and

$$\| \int_a^b u(t)dt \| \le \int_a^b \|u(t)\| dt.$$

If a, b finite then every function in $C_{\mathcal{S}}(a,b)$ is integrable over $[a,b]$.

b) Let B be a closed linear operator on $\mathcal{S}$. Suppose that $u \in C\mathcal{S}(a,b)$, $u(t) \in \mathcal{D}(B)$ for all $t \in [a,b]$, and both u, Bu are integrable over $[a,b]$. Then $\int_a^b u(t)dt \in \mathcal{D}(B)$ and

$$B \int_a^b u(t)dt = \int_a^b Bu(t)dt.$$

c) If $u \in C_{\mathcal{S}}[a,b]$ and u continuously differentiable on $[a,b]$ then

$$\int_a^b \frac{d}{dt} u(t)dt = u(b) - u(a).$$

Proof. See Exercise 3.20. QED

The consequence is that we can interchange of integral and closed linear operators.

The following theorem holds.

Theorem 3.4.9 *Let* $\{ P_t\}_t$ *be an SCSG*$(\mathcal{S})$.

i) *Let* $f\in\mathcal{S}$, $t\geq 0$. *Then* $\int_0^t P_s f ds\in\mathcal{D}(\mathsf{A})$ *and*

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

ii) *Let* $f\in\mathcal{D}(\mathsf{A})$ *and* $t\geq 0$. *Then* $P_t f\in\mathcal{D}(\mathsf{A})$. *The function* $t\mapsto P_t f$ *is differentiable in* $\mathcal{S}$ *and the Kolmogorov backward and forward equations hold:*

$$\frac{d}{dt}P_t f = \mathsf{A}P_t f = P_t\mathsf{A}f.$$

More precisely,

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

iii) *Let* $f\in\mathcal{D}(\mathsf{A})$, $t\geq 0$. *Then*

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

Proof. For the proof of (i) note that

$$\begin{aligned}\frac{1}{h}(P_h - \mathbf{I})\int_0^t P_s f ds &= \frac{1}{h}\int_0^t\Big(P_{s+h}f - P_s f\Big)ds\\ &= \frac{1}{h}\int_t^{t+h} P_s f ds - \frac{1}{h}\int_0^t P_s f ds.\end{aligned}$$

For the second term we get

$$\left\|\frac{1}{h}\int_0^t P_s f ds - f\right\|\leq\frac{1}{h}\int_0^h\| P_s f - f\| ds\to 0,\quad h\downarrow 0.$$

Similarly, for the first term

$$\left\|\frac{1}{h}\int_t^{t+h} P_s f ds - P_t f\right\|\leq\frac{\| P_t\|}{h}\int_0^h\| P_s f - f\| ds\to 0,\quad h\downarrow 0.$$

The result follows. For (ii) note that

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

Taking the limit $h \downarrow 0$ yields that

$$\lim_{h\downarrow 0} \left\| \frac{P_{t+h}f - P_t f}{h} - P_t \mathsf{A} f \right\| = 0. \tag{3.4.1}$$

Since $\mathsf{A}f \in \mathcal{S}$, $g = P_t\mathsf{A}f \in \mathcal{S}$. Rewriting (3.4.1) gives

$$\lim_{h\downarrow 0} \left\| \frac{P_h(P_t f) - P_t f}{h} - g \right\| = 0.$$

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

$$\begin{aligned}\left\| \frac{P_t f - P_{t-h} f}{h} - P_t\mathsf{A}f \right\| &\leq \left\| \frac{P_t f - P_{t-h} f}{h} - P_{t-h}\mathsf{A}f \right\| + \left\| P_{t-h}\mathsf{A}f - P_t\mathsf{A}f \right\| \\ &\leq \left\| \frac{P_h f - f}{h} - \mathsf{A}f \right\| + \left\| P_{t-h}\mathsf{A}f - P_t\mathsf{A}f \right\| \to 0, \quad h \downarrow 0,\end{aligned}$$

where we have used strong continuity and the fact that $\mathsf{A}f \in \mathcal{S}$. For (iii) note that $(d/dt)P_t f = P_t\mathsf{A}f$ is a continuous function of t by Corollary 3.4.5. It is therefore integrable, and so

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

QED

The previous theorem (i) shows that $\mathcal{D}(\mathsf{A})$ is non-empty. In fact it is dense in $\mathcal{S}$ and A is a so-called closed operator.

Corollary 3.4.10 *Let $\{ P_t\}_t$ be an SCSG($\mathcal{S}$). Then $\overline{\mathcal{D}(\mathsf{A})} = \mathcal{S}$ and A is a closed operator.*

Proof. Theorem 3.4.9 (i) and the fact that $\| \int_0^t P_s f ds/t - f\| \to 0$, $t \downarrow 0$ immediately imply that $\overline{\mathcal{D}(\mathsf{A})} = \mathcal{S}$.

Let $\{f_n\}_n \subset \mathcal{D}(\mathsf{A})$ be any sequence with the property that there exist $f, g \in \mathcal{S}$ such that $f_n \to f$ and $\mathsf{A}f_n \to g$ as $n \to \infty$. We need to show that $g = \mathsf{A}f$.

To this end, note that $P_t f_n - f_n = \int_0^t P_s(\mathsf{A}f_n)ds$, for all $t > 0$, by virtue of Theorem 3.4.5 (iii). Since $\|(P_t f_n - f_n) - (P f - f)\| \to 0$ and $\| \int_0^t P_s\mathsf{A}f_n ds - \int_0^t P_s g ds\| \to 0$ as $n \to \infty$, necessarily $P_t f - f = \int_0^t P_s g ds$, for all $t > 0$. Hence

$$\left\| \frac{P_t f - f}{t} - \frac{\int_0^t P_s g ds}{t} \right\| = 0, \quad \forall t > 0.$$

It follows that

$$\lim_{t\downarrow 0} \left\| \frac{P_t f - f}{t} - g \right\| \leq \lim_{t\downarrow 0} \left\| \frac{P_t f - f}{t} - \frac{\int_0^t P_s g ds}{t} \right\| + \lim_{t\downarrow 0} \left\| \frac{\int_0^t P_s g ds}{t} - g \right\| = 0,$$

so that $g = \mathsf{A}f$. QED

By virtue of Corollary 3.4.5, the map $t \to P_t f$ is continuous for each $f \in \mathcal{S}$. Recall that $\| P_t\| \le Me^{\alpha t}$ for some constants $M \ge 1$ and $\alpha \ge 0$. By the Integration Lemma, for all $\lambda > \alpha$ we may define

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

R_λ is simply the Laplace transform of the semigroup calculated at the 'frequency' λ. The next lemma collects preliminary properties of the operators R_λ. In particular, it states that for all $\lambda > 0$, R_λ is in fact an operator that maps $\mathcal{S}$ into itself. It is called the *resolvent of order* λ.

Lemma 3.4.11 *Let* $\{ P_t\}_t$ *be a SCSG(**S**).*

i) $\|R_\lambda\| \le M/(\lambda - \alpha)$.

ii) *The resolvent equation*

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

holds for all $\lambda, \mu > \alpha$.

Proof. The first part is straightforward. To prove the resolvent equation, note that

$$e^{-\mu t} - e^{-\lambda t} = (\lambda - \mu) e^{-\lambda t} \int_0^t e^{(\lambda-\mu)s} ds.$$

Hence,

$$\begin{aligned} R_\mu f(x) - R_\lambda f(x) &= \int_0^\infty (e^{-\mu t} - e^{-\lambda t}) P_t f(x) dt \\ &= (\lambda - \mu) \int_0^\infty e^{-\lambda t} \Big(\int_0^t e^{(\lambda-\mu)s} P_t f(x) ds \Big) dt \\ &= (\lambda - \mu) \int_0^\infty e^{-\mu s} \Big(\int_s^\infty e^{-\lambda(t-s)} P_t f(x) dt \Big) ds, \end{aligned}$$

by the integration Lemma. A change of variables, the semigroup property of the transition function and another application of Integration Lemma show that the inner integral equals

$$\begin{aligned} \int_0^\infty e^{-\lambda u} P_{s+u} f(x) du &= \int_0^\infty e^{-\lambda u} P_s P_u f(x) du \\ &= \int_0^\infty e^{-\lambda u} \Big(\int_E P_u f(y) P_s(x, dy) \Big) du \\ &= \int_E \Big(\int_0^\infty e^{-\lambda u} P_u f(y) du \Big) P_s(x, dy) \\ &= P_s R_\lambda f(x). \end{aligned}$$

Inserting this in the preceding equation yields the resolvent equation. QED

The following important connection between resolvent and generator is easily derived. By **I** we mean the identity operator.

Theorem 3.4.12 *Let* $\{ P_t\}_t$ *be a SCSG(**S**) with* $\| P_t\| \le M \cdot e^{\alpha t}$. *For all* $\lambda > \alpha$ *the following hold.*

i) *$R_\lambda \mathcal{S} = \mathcal{D}(\mathsf{A})$.*

ii) *$\lambda \mathbf{I} - \mathsf{A} : \mathcal{D}(\mathsf{A}) \to \mathcal{S}$ is a 1-1 linear operator with $(\lambda \mathbf{I} - \mathsf{A})\mathcal{D}(\mathsf{A}) = \mathcal{S}$.*

iii) *$(\lambda \mathbf{I} - \mathsf{A})^{-1} : \mathcal{S} \to \mathcal{D}(\mathsf{A})$ exists as a bounded linear operator. In particular $(\lambda \mathbf{I} - \mathsf{A})^{-1} = R_\lambda$.*

iv) *$R_\lambda(\lambda \mathbf{I} - \mathsf{A})f = f$ for all $f \in \mathcal{D}(\mathsf{A})$.*

v) *$(\lambda \mathbf{I} - \mathsf{A})R_\lambda g = g$ for all $g \in \mathcal{S}$.*

Proof. The proof consists of 2 main steps: *Step 1* Proof of (v); and *Step 2* Proof of (iv). As a consequence, (iv) implies (i) and the first part of (ii); (v) implies the second part of (ii). (iii) then follows by combining (iv,v).

Proof of Step 1. Let $g \in \mathcal{S}$. By the Integration Lemma we may write

$$P_h R_\lambda g = \int_0^\infty e^{-\lambda t}\, P_{t+h} g dt.$$

Hence

$$\begin{aligned} \frac{P_h R_\lambda g - R_\lambda g}{h} &= \frac{1}{h}\Big[\int_0^\infty e^{-\lambda t}\, P_{t+h} g dt - \int_0^\infty e^{-\lambda t}\, P_t g dt\Big] \\ &= \frac{e^{\lambda h} - 1}{h} R_\lambda g - \frac{1}{h} e^{\lambda h} \int_0^h e^{-\lambda t}\, P_t g dt. \end{aligned}$$

The right-hand side converges to $\lambda R_\lambda g - g$. It follows that

$$\Big\| \frac{P_h R_\lambda g - R_\lambda g}{h} - (\lambda R_\lambda g - g) \Big\| \to 0, \quad h \downarrow 0.$$

By definition, $R_\lambda g \in \mathcal{D}(\mathsf{A})$ and

$$\mathsf{A} R_\lambda g = \lambda R_\lambda g - g. \tag{3.4.2}$$

The result follows by rewriting.

Proof of Step 2. Let $f \in \mathcal{D}(\mathsf{A})$, then by definition $\mathsf{A} f \in \mathcal{S}$. We have

$$R_\lambda[\mathsf{A} f] = \int_0^\infty e^{-\lambda t}\, P_t[\mathsf{A} f] dt = \int_0^\infty e^{-\lambda t} \mathsf{A}[\, P_t f] dt = \mathsf{A} \int_0^\infty e^{-\lambda t}\, P_t f dt = \mathsf{A} R_\lambda f. \tag{3.4.3}$$

The last equality follows from the Integration Lemma by using that A is closed. The second follows from Theorem 3.4.9 (ii). The rest follows by inserting (3.4.3) into (3.4.2) and rewriting. QED

Due to the importance of resolvents, we will explicitly compute these for two examples.

Example 3.4.13 Consider the BM-process from Example 3.1.7 (A). Its resolvents are given by

$$R_\lambda f(x) = \int_{\mathrm{R}} f(y) r_\lambda(x, y) dy,$$

where $r_\lambda(x, y) = \exp\{-\sqrt{2\lambda}|x - y|\}/\sqrt{2\lambda}$ (see Exercise 3.21).

Example 3.4.14 Let X be the Markov jump process from Example 3.4.3. The resolvent is given by (cf. Exercise 3.22).

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

for $\mu > (c-1)\lambda$. This is by a direct computation, and the fact that (proved using induction)

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

We have proved one part of the Hille-Yosida theorem specifying the precise relation between generator and semigroup. It will however be convenient to restrict to so-called contraction SCSG's: the transition function $\{ P_t \}_T$ is a *strongly continuous contracting semigroup on the Banach space* $\mathcal{S}$ (SCCSG($\mathcal{S}$)) if it is a SCSG($\mathcal{S}$) with $\| P_t \| \leq 1$ for all $t \geq 0$.

This is no restriction. Suppose that $\| P_t \| \leq M \cdot e^{\alpha t}$ for constants $M \geq 1$ and $\alpha \geq 0$. Then $P_t e^{-\alpha t}$ is a SCSG($\mathcal{S}$) with $\| P_t e^{-\alpha t} \| \leq M$. Define a new norm $\| \cdot \|^*$ by

$$\|f\|^* = \sup_{t \geq 0} \| P_t e^{-\alpha t} f \|,$$

then $\|f\| \leq \|f\|^* \leq M\|f\|$. Hence $\| \cdot \|$ and $\| \cdot \|^*$ are equivalent norms and $\mathcal{S}$ is a Banach space with respect to $\| \cdot \|^*$. It easily follows that $\{ P_t e^{-\alpha t} \}_t$ is a SCCSG($\mathcal{S}$) with respect to the new norm.

Theorem 3.4.15 (Hille-Yosida Theorem) *Suppose that there exists a linear operator* $\mathsf{A} : \mathcal{D}(\mathsf{A}) \to \mathcal{S}$, *where* $\mathcal{D}(\mathsf{A})$ *is a linear spubspace of* $\mathcal{S}$. *Then* A *is the generator of a SCCSG($\mathcal{S}$) semigroup if and only if the following three properties hold:*

i) $\overline{\mathcal{D}(\mathsf{A})} = \mathcal{S}$;

ii) A *is* $(0,1)$*-dissipative. In other words:*

$$\|\lambda f - \mathsf{A} f\| \geq \lambda \|f\|, \quad \forall f \in \mathcal{D}(\mathsf{A}), \lambda > 0;$$

iii) $(\lambda \mathbf{I} - \mathsf{A})\mathcal{D}(\mathsf{A}) = \mathcal{S}$.

Proof of "⇒". The only thing left to prove is (ii) (why?). By Theorem 3.4.12 (i) there exists $g \in \mathcal{S}$, such that $f = R_\lambda g$. By the same theorem (v) $(\lambda \mathbf{I} - \mathsf{A}) f = g$. We have

$$\|f\| = \|R_\lambda g\| \leq \frac{1}{\lambda} \|g\| = \frac{1}{\lambda} \|\lambda f - \mathsf{A} f\|.$$

QED

For proving '⇐', we need to derive a number of lemmas. We will formulate these for A being (α, M)-dissipative and return to the $(0,1)$-dissipative case at the moment we really need it.

Lemma 3.4.16 *Let* $\mathsf{A} : \mathcal{D}(\mathsf{A}) \to \mathcal{S}$, *be an* (α, M) *dissipative linear operator, with* $\mathcal{D}(\mathsf{A}) \subset \mathcal{S}$ *a linear subspace, and* $M \geq 1$, $\alpha \geq 0$. *Then* A *is a closed operator if and only if* $(\lambda \mathbf{I} - \mathsf{A})\mathcal{D}(\mathsf{A})$ *is a closed set (in* $\mathcal{S}$*), for some* $\lambda > \alpha$. *Under either condition* $(\lambda \mathbf{I} - \mathsf{A})\mathcal{D}(\mathsf{A})$ *is a closed set (in* $\mathcal{S}$*) is closed for all* $\lambda > \alpha$.

Proof. Assume that A is closed. Let $\lambda > \alpha$. Let $\{f_n\}_n \subset \mathcal{D}(\mathsf{A})$ be a sequence such that $(\lambda\mathbf{I} - \mathsf{A})f_n \to h$, $n \to \infty$, for some $h \in \mathcal{S}$. We have prove that $h \in (\lambda\mathbf{I} - \mathsf{A})\mathcal{D}(\mathsf{A})$.

By (α, M)-dissipativeness of A we have

$$\|(\lambda\mathbf{I} - \mathsf{A})(f_n - f_m)\| \geq \frac{\lambda - \alpha}{M}\|f_n - f_m)\|.$$

Hence $\{f_n\}_n$ is a Cauchy-sequence in $\mathcal{S}$, and so it has a limit $f \in \mathcal{S}$. It holds that

$$\mathsf{A}f_n = -(\lambda\mathbf{I} - \mathsf{A})f_n + \lambda f_n \to -h + \lambda f, \quad n \to \infty.$$

A is closed, hence $f \in \mathcal{D}(\mathsf{A})$, $\mathsf{A}f = -h + \lambda f$. Therefore $h = (\lambda\mathbf{I} - \mathsf{A})f$. The conclusion is that $(\lambda\mathbf{I} - \mathsf{A})\mathcal{D}(\mathsf{A})$ is closed.

Next we assume that $(\lambda\mathbf{I} - \mathsf{A})\mathcal{D}(\mathsf{A})$ is closed in $\mathcal{S}$ for some $\lambda > \alpha$. Let $\{f_n\}_n \subset \mathcal{D}(\mathsf{A})$, with $f_n \to f$, $\mathsf{A}f_n \to g$ for some $f, g \in \mathcal{S}$. Then

$$(\lambda\mathbf{I} - \mathsf{A})f_n \to \lambda f - g, \quad n \to \infty,$$

and so $\lambda f - g \in (\lambda\mathbf{I} - \mathsf{A})\mathcal{D}(\mathsf{A})$. Hence, there exists $h \in \mathcal{D}(\mathsf{A})$, such that $\lambda h - \mathsf{A}h = \lambda f - g$. A is (α, M)-dissipative, so that

$$\|\lambda\mathbf{I} - \mathsf{A})(f_n - h)\| = \|(\lambda(f_n - h) - \mathsf{A}(f_n - h)\| \geq \frac{\lambda - \alpha}{M}\|f_n - h\|.$$

The left-hand side converges to 0, as $n \to \infty$. Hence $f_n \to h$ and so $f = h \in \mathcal{D}(\mathsf{A})$ and $g = \mathsf{A}f$. This shows that A is closed. QED

Lemma 3.4.17 *Let* $\mathsf{A} : \mathcal{D}(\mathsf{A}) \to \mathcal{S}$*, be an* (α, M) *dissipative, closed linear operator, with* $\mathcal{D}(\mathsf{A}) \subset \mathcal{S}$ *a linear subspace, and* $M \geq 1$, $\alpha \geq 0$. *Let*

$$\lambda(\mathsf{A}) = \left\{ \lambda > \alpha \;\middle|\; \begin{array}{l} (\lambda\mathbf{I} - \mathsf{A}) \text{ is 1-1}, \\ (\lambda\mathbf{I} - \mathsf{A})\mathcal{D}(\mathsf{A}) = \mathcal{S}, \\ (\lambda\mathbf{I} - \mathsf{A})^{-1} \text{exists as a bounded linear operator on } \mathcal{S} \end{array} \right\}$$

Then $\lambda(\mathsf{A}) \neq \emptyset \Rightarrow \lambda(\mathsf{A}) = (\alpha, \infty)$.

Proof. It is sufficient to show that $\lambda(\mathsf{A})$ is both open and closed in (α, ∞). First we will show that $\lambda(\mathsf{A})$ is open in (α, ∞). To this end, let $\lambda \in \lambda(\mathsf{A})$. Let

$$B = \{\mu \in (\alpha, \infty) \,|\, |\lambda - \mu| < \|(\lambda\mathbf{I} - \mathsf{A})^{-1}\|^{-1}\}.$$

B is open in (α, ∞). We will show that $B \subset \lambda(\mathsf{A})$.

Let $\mu \in B$. Then

$$C = \sum_{n=0}^{\infty} (\lambda - \mu)^n \Big((\lambda\mathbf{I} - \mathsf{A})^{-1}\Big)^{n+1} : \mathcal{S} \to \mathcal{S}$$

is a bounded linear operator. We claim that $C = (\mu\mathbf{I} - \mathsf{A})^{-1}$. Use $\mu\mathbf{I} - \mathsf{A} = \lambda\mathbf{I} - \mathsf{A} + (\mu - \lambda)\mathbf{I}$ to show that $C(\mu\mathbf{I} - \mathsf{A}) = (\mu\mathbf{I} - \mathsf{A})C = \mathbf{I}$. Then it easily follows that $\mu \in \lambda(\mathsf{A})$.

We next show that $\lambda(\mathsf{A})$ is closed in (α,∞). To this end, let $\{\lambda_n\}_n \subset \lambda(\mathsf{A})$, with $\lambda_n \to \lambda$, $n \to \infty$, for some $\lambda \in \mathbf{R}$, $\lambda > \alpha$. We have to show that $\lambda \in \lambda(\mathsf{A})$.

The proof consists of two steps: Step 1 $(\lambda\mathbf{I}-\mathsf{A})\mathcal{D}(\mathsf{A}) = \mathcal{S}$; Step 2 $(\lambda\mathbf{I}-\mathsf{A})$ is 1-1. Steps 1 and 2 then imply that for each $g \in \mathcal{S}$ there exists precisely one element $f \in \mathcal{D}(\mathsf{A})$ with $g = (\lambda\mathbf{I}-\mathsf{A})f$. This means that the inverse $(\lambda\mathbf{I}-\mathsf{A})^{-1}$ exists. By (α, M)-dissipativeness

$$\|(\lambda\mathbf{I}-\mathsf{A})^{-1}g\| = \|f\| \le \frac{M}{\lambda-\alpha}\|\lambda f - \mathsf{A}f\| = \frac{M}{\lambda-\alpha}\|g\|.$$

Since $g \in \mathcal{S}$ was arbitrary, $\|(\lambda\mathbf{I}-\mathsf{A})^{-1}\| \le M/(\lambda-\alpha) < \infty$. This shows that $\lambda \in \lambda(\mathsf{A})$.

Proof of Step 1. Let $g \in \mathcal{S}$. Put $g_n = (\lambda\mathbf{I}-\mathsf{A})(\lambda_n\mathbf{I}-\mathsf{A})^{-1}g$, for all n. Clearly $g = (\lambda\mathbf{I}-\mathsf{A})(\lambda_n\mathbf{I}-\mathsf{A})^{-1}g$ and so $g_n - g = (\lambda-\lambda_n)(\lambda_n\mathbf{I}-\mathsf{A})^{-1}g$. A is (α, M)-dissipative, and so

$$\|g_n - g\| \le |\lambda_n-\lambda|\|(\lambda_n\mathbf{I}-\mathsf{A})^{-1}g\| \le |\lambda_n-\lambda|\frac{M}{\lambda-\alpha}\|g\| \to 0, \quad n\to\infty.$$

It follows that $\overline{(\lambda\mathbf{I}-\mathsf{A})\mathcal{D}(\mathsf{A})} = \mathcal{S}$. Since $(\lambda\mathbf{I}-\mathsf{A})\mathcal{D}(\mathsf{A})$ is closed by Lemma 3.4.16, $\mathcal{S} = (\lambda\mathbf{I}-\mathsf{A})\mathcal{D}(\mathsf{A})$.

Proof of Step 2. This follows immediately from (α, M)-dissipativeness. QED

Lemma 3.4.18 *Let $\mathsf{A} : \mathcal{D}(\mathsf{A}) \to \mathcal{S}$, be an (α, M) dissipative, closed linear operator, with $\mathcal{D}(\mathsf{A}) \subset \mathcal{S}$ a linear subspace, and $M \ge 1$, $\alpha \ge 0$. Suppose that $\overline{\mathcal{D}(\mathsf{A})} = \mathcal{S}$ and $\lambda(\mathsf{A}) = (\alpha,\infty)$. Define the Yosida-approximation $\mathsf{A}_\lambda = \lambda\mathsf{A}(\lambda\mathbf{I}-\mathsf{A})^{-1}$, $\lambda > \alpha$. Then A_λ, $\lambda > \alpha$, have the following properties.*

- **a)** *A_λ is a bounded linear operator $\mathcal{S} \to \mathcal{S}$ and $e^{t\mathsf{A}_\lambda} = \sum_{n=0}^\infty t^n\mathsf{A}_\lambda^n/n!$ is a SCSG($\mathcal{S}$) with generator A_λ.*
- **b)** $\mathsf{A}_\lambda\mathsf{A}_\mu = \mathsf{A}_\mu\mathsf{A}_\lambda$.
- **c)** $\|\mathsf{A}f - \mathsf{A}_\lambda f\| \to 0$, *$\lambda \to \infty$, for all $f \in \mathcal{D}(\mathsf{A})$.*

Proof. For all $\lambda > \alpha$ write $U_\lambda = (\lambda\mathbf{I}-\mathsf{A})^{-1}$. In the proof of Lemma 3.4.18 we have seen that $\|U_\lambda\| \le M/(\lambda-\alpha)$. Further $U_\lambda U_\mu = U_\mu U\lambda$. This follows by a straightforward computation. Hence

$$\begin{aligned}(\lambda\mathbf{I}-\mathsf{A})U_\lambda &= \mathbf{I}, \quad \text{on } \mathcal{S} && (3.4.4)\\ U_\lambda(\lambda\mathbf{I}-\mathsf{A}) &= \mathbf{I}, \quad \text{on } \mathcal{D}(\mathsf{A}). && (3.4.5)\end{aligned}$$

We first prove (a). Using the above, we may rewrite $\mathsf{A}_\lambda = \lambda\mathsf{A}U_\lambda$ by

$$\begin{aligned}\mathsf{A}_\lambda &= \lambda^2U_\lambda - \lambda\mathbf{I}, \quad \text{on } \mathcal{S} && (3.4.6)\\ &= \lambda U_\lambda\mathsf{A}, \quad \text{on } \mathcal{D}(\mathsf{A}). && (3.4.7)\end{aligned}$$

(3.4.6) implies that A_λ is bounded, with

$$\|e^{t\mathsf{A}_\lambda}\| \le e^{-t\lambda}e^{t\lambda^2\|U_\lambda\|} \le e^{t\lambda^2M/(\lambda-\alpha)-t\lambda}. \tag{3.4.8}$$

Further, for all $f \in \mathcal{S}$

$$\|e^{t\mathsf{A}_\lambda} f - f\| \le \sum_{n=1}^{\infty} t^n \|\mathsf{A}_\lambda\|^n \|f\|/n! \to 0, \quad t \downarrow 0.$$

Hence $\{e^{t\mathsf{A}_\lambda}\}_t$ is a SCSG($\mathcal{S}$). In a similar way, one can prove that A_λ is the generator.

The proof of (b) follows by using the expression (3.4.6) for A_λ on $\mathcal{S}$ and the fact that U_λ and U_μ commute. We will finally prove (c). First we show that $\|\lambda U_\lambda f - f\| \to 0$, $\lambda \to \infty$, for all $f \in \mathcal{S}$.

For $f \in \mathcal{D}(\mathsf{A})$, we use (3.4.5) to obtain

$$\|\lambda U_\lambda f - f\| = \|U_\lambda \mathsf{A} f\| \le \frac{M}{\lambda - \alpha}\|\mathsf{A} f\| \to 0, \quad \lambda \to \infty.$$

Let $f \in \mathcal{S}$ and let $\{f_n\}_n \subset \mathcal{D}(\mathsf{A})$ converge to f. Then for all n

$$\limsup_{\lambda\to\infty} \|\lambda U_\lambda f - f\| \le \limsup_{\lambda\to\infty} \Big[\|\lambda U_\lambda f_n - f_n\| + \frac{M\lambda}{\lambda-\alpha}\|f_n - f\| + \|f_n - f\|\Big]$$

Let $\lambda \to \infty$. The first term on the right-hand side converges to 0, the second converges to $M\|f_n - f\|$ and the third equals $\|f_n - f\|$. Since the left-hand side is independent of n, we can take the limit $n \to \infty$ and obtain that the left-hand side must equal 0. (c) follows by combining with (3.4.7). QED

Lemma 3.4.19 *Suppose that B, C are bounded linear operators on $\mathcal{S}$ with $\|e^{tB}\|, \|e^{tC}\| \le 1$, that commute: $BC = CB$. Then*

$$\|e^{tB} f - e^{tC} f\| \le t\|Bf - Cf\|$$

for every $f \in \mathcal{S}$ and $t \ge 0$.

Proof. Use the identity

$$\begin{aligned} e^{tB}f - e^{tC}f &= \int_0^t \frac{d}{ds}[e^{sB}e^{(t-s)C} f ds = \int_0^t e^{sB}(B-C)e^{(t-s)C} ds \\ &= \int_0^t e^{sB}e^{(t-s)C}(B-C) f ds \le t\|Bf - Cf\|. \end{aligned}$$

QED

Continuation of the proof of the Hille-Yosida Theorem: '$\Leftarrow$' The idea is the define a strongly continuous contraction semigroup $\{P_t\}_t$ and then show that it has generator A.

Conditions (ii), (iii) and Lemma 3.4.16 imply that A is closed. By inspection of the proof of closedness of $\lambda(\mathsf{A})$ of Lemma 3.4.17, that $\lambda(\mathsf{A}) \neq \emptyset$. Lemma 3.4.17 implies $\lambda(\mathsf{A}) = (0, \infty)$.

Using the notation in Lemma 3.4.18, define for each $\lambda > 0$ the SCCSG($\mathcal{S}$) $\{e^{t\mathsf{A}_\lambda}\}_t$. Let us check that it is indeed a SCCSG($\mathcal{S}$).

By virtue of (3.4.8), $\|e^{t\mathsf{A}_\lambda}\| \le 1$. By virtue of Lemmas 3.4.18(b) and 3.4.19

$$\|e^{t\mathsf{A}_\lambda} f - e^{t\mathsf{A}_\mu} f\| \le t\|\mathsf{A}_\lambda f - \mathsf{A}_\mu f\|,$$

for all $t \geq 0$, and $f \in \mathcal{S}$. By virtue of Lemma 3.4.18 (c), $\lim_{\lambda\to\infty} e^{t\mathsf{A}_\lambda} f$ exists for all $t \geq 0$, uniformly in $t \in [0,T]$, for each $T > 0$, for all $f \in \mathcal{D}(\mathsf{A})$. Define $P_t f = \lim_{\lambda\to\infty} e^{t\mathsf{A}_\lambda} f$, for all $f \in \mathcal{D}(\mathsf{A})$. By uniform convergence, $t \to e^{t\mathsf{A}_\lambda} f$ is continuous in t for each $f \in \mathcal{D}(\mathsf{A})$.

The fact that $\overline{\mathcal{D}(\mathsf{A})} = \mathcal{S}$ allows to define $P_t f$ on all of $\mathcal{S}$, for all $t \geq 0$.

Next It holds that

$$P_{t+s}f - P_t P_s f \;=\; P_{t+s}f - e^{(t+s)\mathsf{A}_\lambda} f + e^{t\mathsf{A}_\lambda}(e^{s\mathsf{A}_\lambda} f - P_s) f + (e^{s\mathsf{A}_\lambda} - P_s)\, P_t f.$$

This allows conclude that the Chapman-Kolmogorov equations apply. We may similarly prove that strong continuity holds and that P_t is a bounded linear operator with norm at most 1, $t \geq 0$. We may then conclude that $\{P_t\}_t$ is a SCCSG($\mathcal{S}$).

Finally we will show that this SCSG has generator A. By Theorem 3.4.9 (iii)

$$e^{t\mathsf{A}_\lambda} f - f = \int_0^t e^{s\mathsf{A}_\lambda} \mathsf{A}_\lambda f ds, \tag{3.4.9}$$

for all $f \in \mathcal{S}, t \geq 0, \lambda > \alpha$. For all $f \in \mathcal{D}(\mathsf{A})$ and $t \geq 0$

$$e^{s\mathsf{A}_\lambda} \mathsf{A}_\lambda f - P_s f = e^{s\mathsf{A}_\lambda}(\mathsf{A}_\lambda f - \mathsf{A} f) + (e^{s\mathsf{A}_\lambda} - P_s) f.$$

By virtue of Lemma 3.4.17 (iii) this implies that $\|e^{s\mathsf{A}_\lambda}\mathsf{A}_\lambda f - P_s \mathsf{A} f\| \to 0, \quad \lambda \to \infty$, uniformly in $s \in [0,t]$. Combining with (3.4.9) yields $P_t f - f = \int_0^t P_s \mathsf{A} f ds$, for all $f \in \mathcal{D}(\mathsf{A})$ and $t \geq 0$.

Suppose that $\{P_t\}_t$ has generator B, with domain $\mathcal{D}(\mathsf{B})$. The above implies that $\mathcal{D}(\mathsf{B}) \supset \mathcal{D}(\mathsf{A})$ and $\mathsf{B} = \mathsf{A}$ on $\mathcal{D}(\mathsf{A})$. Hence B extends A.

By Theorem 3.4.12 $\lambda\mathbf{I} - \mathsf{B}$ is 1-1 for λ large enough. Since $\mathcal{S} = (\lambda\mathbf{I} - \mathsf{A})\mathcal{D}(\mathsf{A}) = (\lambda\mathbf{I} - \mathsf{B})\mathcal{D}(\mathsf{A})$, $\mathcal{D}(\mathsf{B}) \setminus \mathcal{D}(\mathsf{A}) = \emptyset$, otherwise we would get a contradiction with the fact that $\lambda\mathbf{I} - \mathsf{B}$ is 1-1. QED

It is generally hard to determine the domain of a generator. The following lemma may be of use.

Lemma 3.4.20 *Let $\{P_t\}_t$ be a SCSG($\mathcal{S}$) with $\|P_t\| \leq Me^{\alpha t}$. Suppose that the linear operator B is an extension of the generator A. In other words, $B : \mathcal{D} \to \mathcal{S}$ is a linear operator with $\mathcal{D} \supset \mathcal{D}(\mathsf{A})$, and $Bf = \mathsf{A}f$ for $f \in \mathcal{D}(\mathsf{A})$. If $\lambda\mathbf{I} - B$ is 1-1 for some value $\lambda > \alpha$, then B equals A, that is, $\mathcal{D} = \mathcal{D}(\mathsf{A})$. A sufficient condition for $\lambda\mathbf{I} - B$ to be 1-1 is that there is no non-trivial $f \in \mathcal{D}$ with $Bf = \lambda f$.*

Proof. Suppose that $f \in \mathcal{D}$. Put $g = \lambda f - Bf$. Then $h = R_\lambda g \in \mathcal{D}(\mathsf{A})$ and so

$$\lambda f - Bf = g = \lambda h - \mathsf{A}h = \lambda h - Bh, \tag{3.4.10}$$

since $\mathsf{A} = B$ on $\mathcal{D}(\mathsf{A})$. Hence $f = h \in \mathcal{D}(\mathsf{A})$, if $\lambda\mathbf{I} - B$ is 1-1.

Suppose that we only know that there does not exist any function $F \in \mathcal{S}$ with $BF = \lambda F$. Then substracting the right side in (3.4.10) from the left, we obtain that $f = h$, and thus $\lambda\mathbf{I} - B$ is 1-1. QED

We may finally ask ourselves whether the transition function is *uniquely* determined by the generator. The answer is again in the case of a SCSG($\mathcal{S}$). In general this need not be true.

In the case of special Banach spaces, more interesting properties prevail.

3.5 Feller-Dynkin transition functions

From now on consider semigroups with additional properties. For simplicity, the state space E is assumed to be a closed or open subset of $\mathbf{R}^d$ with $\mathcal{E}$ its Borel-σ-algebra, or of $\mathbf{Z}^d$ with the discrete topology and with the $\mathcal{E}$ the σ-algebra generated by the one-point sets. 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

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

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

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

Note that $C_0(E)$ is a subset of the space of $b\mathcal{E}$ of bounded, measurable functions on E, so we can consider the restriction of the transition operators $(P_t)_t$ to $C_0(E)$. It is also a Banach space.

We will now introduce a seemingly weaker condition on the semigroup than strong continuity. This notion is not uni-valent in the literature: sometimes it is called the Feller property.

Definition 3.5.1 The transition function $(P_t)_{t\ge 0}$ is called a *Feller-Dynkin transition function* if

i) $P_tC_0(E) \subseteq C_0(E)$;

ii) $P_tf(x) \to f(x)$, $t \downarrow 0$, 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

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

So, for all $t \ge 0$ we have $\|P_t\|_\infty \le 1$, where $\|P_t\|_\infty$ is the norm of P_t as a linear operator on the normed linear space $C_0(E)$, endowed with the supremum norm (see Appendix B LN, or BN section 11).

If $f \in C_0(E)$, then $P_tf \in C_0(E)$ by part (i) of Definition 3.5.1. By the semigroup property and part (ii) it follows that

$$P_{t+h}f(x) = P_h(P_tf)(x) \to P_tf(x), \quad h \downarrow 0.$$

In other words, the map $t \mapsto P_tf(x)$ is right-continuous for all $f \in C_0(E)$ and $x \in E$.

Right-continuous functions $f : \mathbf{R} \to \mathbf{R}$ are $\mathcal{B}(\mathbf{R})/\mathcal{B}(\mathbf{R})$-measurable. See BN§3.

In particular, this map is measurable, and so for all $\lambda > 0$ we may define the resolvent R_λ. Also in this weaker case, it is a bounded linear operator with norm $\|R_\lambda\| \leq 1$. We will now show that in the case of the sapce $C_0(E)$, pointwise continuity implies strong continuity.

Theorem 3.5.2 *Suppose that* $\{ P_t\}_t$ *is a Feller-Dynkin transition function. Then* $\overline{R_\lambda C_0(E)} = C_0(E)$*, and* $\{ P_t\}_t$ *is a SCCSG(*$C_0(E)$*).*

Proof. In order that a Feller-Dynkin transition function be a SCCSG($C_0(E)$), we only need to show strong continuity. We will first show that this is easily checked, provided that $R_\lambda C_0(E)$ is dense in $C_0(E)$.

Since $P_t R_\lambda f(x) = e^{\lambda t} \int +t^\infty e^{-\lambda s}\, P_s f(x) ds$ by the Integration Lemma,

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

Therefore

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

Since the right-hand side tends to 0 as $t \downarrow 0$, this shows desired norm continuity for functions in the dense subset $R_\lambda C_0(E)$ of $C_0(E)$. Now let $f \in C_0(E)$ be arbitrary. Then for every $g \in R_\lambda C_0(E)$ it holds that

$$\begin{aligned} \|P_t f - f\|_\infty &\leq \|P_t f - P_t g\|_\infty + \|P_t g - g\|_\infty + \|g - f\|_\infty \\ &\leq \|P_t g - g\|_\infty + 2\|g - f\|_\infty. \end{aligned}$$

Taking the $\limsup_{t\downarrow 0}$ and using the first part of the proof, we get

$$\limsup_{t\downarrow 0} \|P_t f - f\|_\infty \leq 2\|g - f\|_\infty,$$

for every $g \in R_\lambda C_0(E)$. The right-hand side can be made arbitrarily small, since $R_\lambda C_0(E)$ is dense in $C_0(E)$. Hence $\lim_{t\downarrow 0} \|P_t f - f\|_\infty = 0$.

Next we will show that $R_\lambda C_0(E)$ is dense in $C_0(E)$. Suppose that this is not true and that $\overline{R_\lambda C_0(E)} \subsetneq C_0(E)$. By the Hahn-Banach theorem (BN §9, Corollary 11.2) there exists a non-trivial bounded linear functional B on $C_0(E)$ that vanishes on $R_\lambda(C_0(E))$. By the Riesz representation theorem (BN §9 Theorem 11.3) there exist finite Borel measures ν and ν' on E such that

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

for every $f \in C_0(E)$. By part (ii) of Definition 3.5.1 and dominated convergence, for every $x \in E$

$$\lambda R_\lambda f(x) = \int_0^\infty \lambda e^{-\lambda t} P_t f(x) dt = \int_0^\infty e^{-s} P_{s/\lambda} f(x) ds \to f(x), \quad \lambda \to \infty. \tag{3.5.1}$$

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

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

We conclude that the functional B vanishes on the entire space $C_0(E)$ and so B is trivial. A contradiction. QED

Example 3.5.3 Brownian motion and the Ornstein-Uhlenbeck process from Example 3.1.7 (A, B) are Feller-Dynkin processes. Hence their semigroups are SCCSG($C_0(\mathbf{R})$) (cf. Example 3.4.2).

We will show that the BM transition function has the Feller-Dynkin property. First we check continuity of $x \mapsto P_t f(x)$ for each $t \geq 0$ and $f \in C_0(\mathbf{R})$. To this end let $x \in \mathbf{R}$ and let $\{x_n\}_n \subset \mathbf{R}$ be a converging sequence with limit x and let $f \in C_0(\mathbf{R})$. Then

$$P_t f(x_n) = \int_{\mathrm{R}} f(x_n - u) \frac{1}{\sqrt{2\pi t}} e^{-u^2/2t} du.$$

Define the functions f_n by $f_n(u) = f(x_n - u)$, $n = 1, \ldots$. We have $f_n \in C_0(\mathbf{R})$, and $\sup_n \|f_n\| = \|f\| < \infty$. Note that $f_n(u) \to f(x-u)$, $n \to \infty$, by continuity of f. By dominated convergence it follows that

$$\lim_{n\to\infty} P_t f(x_n) = \lim_{n\to\infty} \int_{\mathrm{R}} f_n(u) \frac{1}{\sqrt{2\pi t}} e^{-u^2/2t} du = \int_{\mathrm{R}} f(x-u) \frac{1}{\sqrt{2\pi t}} e^{-u^2/2t} du = P_t f(x).$$

The proof that $P_t f(x) \to 0$ as $|x| \to \infty$, is proved similarly.

We will prove pointwise continuity of the function $t \mapsto P_t f(x)$ as $t \downarrow 0$, $x \in \mathbf{R}$. Note that this amounts to proving that $\mathsf{E}_x f(X_t) \to \mathsf{E}_x f(X_0)$ for each $f \in C_0(\mathbf{R})$. First, by sample path continuity $X_t \to X_0$, $t \downarrow 0$, P_x-a.s. Hence by BN Lemma 5.4 $X_t \xrightarrow{\mathcal{D}} X_0$, $t \downarrow 0$. The BN Portmanteau theorem 5.3 implies desired convergence.

Example 3.5.4 Consider the Markov jump process in Example 3.2.5. It is a Feller-Dynkin process if $PC_0(E) \subseteq C_0(E)$ (cf. Exercise 3.23).

3.5.1 Computation of the generator

Let us first look compute the generator of Brownian motion in a straightforward manner.

Example 3.5.5 (cf. Example 3.4.6). 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 to $\mathcal{D}(\mathsf{A}') = \mathsf{C}_0^2(\mathbf{R})$.

We will use Lemma 3.4.20 to show that $\mathsf{A}' = \mathsf{A}$. The lemma implies that this is true if, given $\lambda > 0$, say $\lambda = 1$, there exists no function $h \in \mathcal{D}(\mathsf{A}') = \mathsf{C}_0^2(\mathbf{R})$ with $h''/2 = h$.

Let us assume that such a function exists. Then there exists $x \in \mathbf{R}$ with $h(x) \geq h(y)$ for all $y \in \mathbf{R}$. This implies that $h'(x) = 0$ and $h''(x) \leq 0$ by a second order Taylor expansion.

Hence $h(x) = h''(x)/2 \leq 0$. Consequently $0 \geq h(y) = h''(y)/2$ for all $y \in \mathbf{R}$. It follows that $h'(y) \leq 0$ for $y \geq x$.

By the assumption that h is non-trivial, there must be some $x' \geq x$ with $h(x') < 0$. Since $h \in C_0(\mathbf{R})$, $\lim_{y\to\infty} h(y) = 0$. Hence, there exists $y > x$ with $h'(y) > 0$. Contradiction.

We are left to show that **A1** and **A2** hold. Let $h \in \mathcal{D}(\mathsf{A})$. Then there exists $f \in \mathsf{C}_0(\mathbf{R})$ such that $h = \mathsf{R}_\lambda f$, i.e.

$$\begin{aligned} h(x) = R_\lambda f(x) &= \int_{\mathrm{R}} f(y) r_\lambda(x,y)dy \\ &= \int_{\mathrm{R}} f(y)\frac{1}{\sqrt{2\lambda}} e^{-\sqrt{2\lambda}|y-x|}dy, \end{aligned}$$

which integral is bounded and differentiable to x. Hence,

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

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

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

Clearly,

$$\int_{y<x} \cdots + \int_{y>x+\delta} \cdots \to 2\lambda \int_{\mathrm{R}} f(y) r_\lambda(x,y)dy = 2\lambda h(x), \quad \delta \to 0.$$

Further,

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

by dominated convergence. Combining yields

$$\frac{h'(x+\delta) - h'(x)}{\delta} \to 2\lambda h(x) - 2f(x), \quad \delta \downarrow 0.$$

The same holds for $\delta \uparrow 0$, and so we find that h is twice differentiable with

$$h'' = 2\lambda h - 2f.$$

Hence,

$$\lambda h - \tfrac{1}{2}h'' = f = (\lambda \mathbf{I} - \mathsf{A})h = \lambda h - \mathsf{A}h,$$

and so **A1,2** hold.

The first observation is that this is a nasty computation that is hardly amenable to use for more complicated situations. The second is that we derive the domain directly from Lemma 3.4.20. What we implicitly use is the so-called *maximum principle* that we will not further discuss. We believe that the essential feature validating an extension B to equal the generator A, is the fact that $\lambda\mathbf{I} - B$ is 1-1.

Before further investigating ways to compute the generator, we first generalise Lemma 3.1.9, showing conditions under which a function ϕ of a Feller-Dynkin process is Feller-Dynkin, and it provides a relation between the corresponding generators.

Lemma 3.5.6 *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 3.5.7 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 3.24.

3.5.2 Applications of the generator and alternative computation

Generators provide an important link between Feller-Dynkin processes and martingales.

Theorem 3.5.8 *Let X be a Feller-Dynkin process, defined on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_t)$. For every $f \in \mathcal{D}(\mathsf{A})$ and initial 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 $s \leq t$

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

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 Theorem 3.4.9 (iii). QED

This gives rise to the following extremely convenient formula.

Corollary 3.5.9 (Dynkin's formula) *Let X be a canonical, right-continuous Feller-Dynkin process. For every $f \in \mathcal{D}(\mathsf{A})$ and every $\{\mathcal{F}_t^X\}_t$-stopping time τ with $\mathsf{E}_x\tau < \infty$, we have*

$$\mathsf{E}_x f(X_\tau) = f(x) + \mathsf{E}_x \int_0^\tau \mathsf{A}f(X_s)ds, \quad x \in E.$$

Proof. By Theorem 3.5.8 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 $\mathsf{A}f \in C_0(E)$, we have $\|\mathsf{A}f\|_\infty < \infty$ and so

$$\Big| \int_0^{\tau\wedge n} \mathsf{A}f(X_s)ds \Big| \leq \|\mathsf{A}f\|\tau.$$

By the fact that $\mathsf{E}_x\tau < \infty$ and by dominated convergence, the integral on the right-handside converges to

$$\mathsf{E}_x \int_0^\tau \mathsf{A}f(X_s)ds.$$

QED

This lemma is particularly useful.

Example 3.5.10 Consider the (canonical continuous) BM process $X_t = X_0 + W_t$, $t \geq 0$, where X_0 and $(W_t)_t$ are independent, and $(W_t)_t$ a standard BM.

Let $(a, b) \subset \mathbf{R}$, $a < b$. The problem is to determine a function $f : [a, b] \to \mathbf{R}$, $f \in C_0^2[a, b]$, with $f''(x) = 0$, $x \in (a, b)$ and $f(a) = c_1$, $f(b) = c_2$ for given constants c_1, c_2. Clearly, in dimension 1 this is a simple problem - f is a linear function. However, we would like to use it as an illustration of our theory.

Suppose such a function f exists. Then f can be extended as a $C_0^2(\mathbf{R})$ function. Then $f \in \mathcal{D}(\mathsf{A})$ for our process X. Let $\nu = \delta_x$ be the initial distribution of X for some $x \in (a, b)$. We have seen that $\tau_{a,b} = \inf\{t > 0 \,|\, X_t \in \{a, b\}\}$ is a finite stopping time with finite expectation. Consequently, Dynkin's formula applies and so

$$\mathsf{E}_x f(X_{\tau_{a,b}}) = f(x) + \mathsf{E}_x \int_0^{\tau_{a,b}} \mathsf{A}f(X_s)ds = f(x) + \mathsf{E}_x \int_0^{\tau_{a,b}} \tfrac{1}{2}f''(X_s)ds = f(x).$$

The left-handside equals

$$c_1\frac{b-x}{b-a} + c_2\frac{x-a}{b-a}$$

(cf. Exercise 2.28).

Characteristic operator We will now give a probabilistic interpretation of the generator.

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 3.10).

Let X be a canonical Feller-Dynkin process For $r > 0$, define the $\{\mathcal{F}_t^X\}_t$-stopping time

$$\eta_r = \inf\{t \geq 0 \mid \|X_t - X_0\| > r\}. \tag{3.5.2}$$

If x is absorbing, then P_x-a.s. we have $\eta_r = \infty$ for all $r > 0$. For non-absorbing points however, the *escape time* η_r is a.s. finite and has finite mean provided r is small enough.

Lemma 3.5.11 *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 $P_t(x, B_x(\epsilon)) < p < 1$ for some $t, \epsilon > 0$.

By the Feller-Dynkin property of the semi-group P_t we have that $P_t(y, \cdot) \xrightarrow{\mathsf{w}} P_t(x, \cdot)$ as $y \to x$. Hence, the Portmanteau theorem, and the fact that $B_x(\epsilon)$ is closed imply that

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

Let $\hat{p} \in (p, 1)$. It follows that for all y sufficiently close to x, say $y \in B_x(r)$ for some $r \in (0, \epsilon)$, we have $P_t(y, B_x(r)) < \hat{p}$. Using the Markov property it is easy to show (cf. Exercise 3.25) that $\mathsf{P}_x(\eta_r \geq nt\} \leq \hat{p}^n$, $n = 0, 1, \ldots$. Hence,

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

This completes the proof. QED

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

Theorem 3.5.12 *Let X be a right-continuous canonical Feller-Dynkin process. For $f \in \mathcal{D}(\mathsf{A})$ we have $\mathsf{A}f(x) = 0$ if x is absorbing, and otherwise*

$$\mathsf{A}f(x) = \lim_{r \downarrow 0} \frac{\mathsf{E}_x f(X_{\eta_r}) - f(x)}{\mathsf{E}_x \eta_r}. \tag{3.5.3}$$

Proof. If x is absorbing, we have $P_t f(x) = f(x)$ 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

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

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

The operator defined by the right-handside of (3.5.3) is called the *characteristic operator* of the Markov process X. Its domain is simply the collection of all functions $f \in C_0(E)$ for which the limit in (3.5.3) exists as a $C_0(E)$-function. The theorem states that for right-continuous, canonical Feller-Dynkin processes the characteristic operator extends the infinitesimal generator. We will check the conditions of Lemma 3.4.20 to show that the characteristic operator is the generator. To this end, denote the characteristic operator by B. Suppose that generator and characteristic operator are not equal. Then there exists $f \in C_0(E)$, with $\lambda f = Bf$. We may assume $\lambda = 1$. Then this implies that

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

Suppose that f has a maximum at x: $f(x) \geq f(y)$ for all $y \in E$. Then the above implies that $f(x) \leq 0$ and hence $f(y) \leq 0$ for all $y \in E$. On the other hand, $g = -f$ satisfies $Bg = g$. Let x' be a point where g is maximum. Then, similarly, $g(x') \leq 0$ and so $g(y) \leq 0$ for all $y \in E$. Consequently $f = -f \equiv 0$.

Example 3.5.13 Let X be a canonical right-continuous Feller-Dynkin process on the countable state space E, equipped with the σ-algebra $\mathcal{E}$, where $\mathcal{E}$ is generated by the one-point sets of E. Define $\sigma_x = \inf\{t > 0 \mid X_t \neq x\}$. In Exercises 3.10 and 3.11 you are asked to show that $\mathsf{P}_x\{\sigma_x > t\} = e^{-a(x)t}$, $t \geq 0$, for some $a(x) \in [0, \infty]$. Further, if $a(x) \in (0, \infty)$, then x is a *holding point*, and the holding time (in state x) is a (negative) exponentially distributed random variable with parameter $a(x)$. Moreover, $\mathsf{P}_x\{X_{\sigma_x} = x\} = 0$, so that a holding point x can only be left by a jump. Regular points cannot exist, because of required continuity of $t \mapsto P_t f(x)$ for all $f \in C_0(E)$.

The first observation is that $\mathbf{1}_{\{y\}} \in \mathcal{D}(\mathsf{A})$ for all $y \in E$. This is shown by using the characteristic operator. If x is absorbing, then $\mathsf{A}\mathbf{1}_{\{y\}}(x) = 0$. Suppose that x is not absorbing. If $x \neq y$, then

$$\mathsf{A}\mathbf{1}_{\{y\}}(x) = \lim_{r \downarrow 0} \frac{\mathsf{E}_x(X_{\eta_r})}{\mathsf{E}_x \sigma_x} = \frac{\mathsf{P}_x(X_{\sigma_x} = y)}{\mathsf{E}_x \sigma_x} = a(x)\mathsf{P}_x(X_{\sigma_x} = y);$$

and if $x = y$

$$\mathsf{A}\mathbf{1}_{\{y\}}(x) = \frac{-1}{\mathsf{E}_x \sigma_x} = a(x).$$

This implies that the generator can be represented as an $E \times E$ matrix, with elements $\mathsf{A}(x, y)$, $x, y \in E$, given by

$$\mathsf{A}(x, y) = \begin{cases} 0, & x \text{ absorbing} \\ a(x)\mathsf{P}_x(X_{\sigma_x} = y), & x \text{ holding point }, y \neq x \\ a(x), & x \text{ holding point }, y = x. \end{cases}$$

Finally we turn to the regularisation problem for Feller-Dynkin processes.

3.6 Regularisation of Feller-Dynkin processes

3.6.1 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 $(\Omega, \mathcal{F}) = (E_\delta^{\mathbb{R}_+}, \mathcal{E}_\delta^{\mathbb{R}_+})$, 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 3.6.1 *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

$$\mathsf{E}_\nu(e^{-\lambda t} R_\lambda f(X_t) \mid \mathcal{F}_s^X) = e^{-\lambda t} P_{t-s} R_\lambda f(X_s) \quad \mathsf{P}_\nu - \text{a.s.}$$

(see Theorem 3.2.4). Hence, to prove the statement of the lemma it suffices to prove that

$$e^{-\lambda t} P_{t-s} R_\lambda f(x) \leq e^{-\lambda} R_\lambda f(x), \quad x \in E. \tag{3.6.1}$$

This is a straightforward calculation (cf. Exercise 3.30). QED

Theorem 3.6.2 *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 $t \geq 0$ 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 3.31). By the second statement of Corollary **??**, the class

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

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

$$\Omega_{h'} = \{\omega \mid \lim_{q \downarrow t} h'(X_r)(\omega), \lim_{q \uparrow t} h'(X_r)(\omega) \text{ exist as finite limits for all } t > 0\} \tag{3.6.2}$$

is $\mathcal{F}_\infty^X$-measurable. By virtue of Lemma 3.6.1 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 3.32, it follows that on Ω' the limits

$$\lim_{q \downarrow t} X_q(\omega), \quad \lim_{q \uparrow t} X_q(\omega)$$

exist in E_δ, for all $t \geq 0$, $\omega \in \Omega'$.

Now fix an arbitrary point $x_0 \in E$ and define a new process $Y = (Y_t)$ as follows. For $\omega \notin \Omega'$, put $Y_t(\omega) = x_0$. For $\omega \in \Omega'$ and $t \geq 0$ define

$$Y_t(\omega) = \lim_{q \downarrow t} X_q(\omega).$$

We claim that for every initial distribution ν and $t \geq 0$, we have $X_t = Y_t$ P_ν-a.s. To prove this, let f and g be two functions on $C_0(E_\delta)$. By dominated convergence, and the Markov property

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

By strong continuity, $P_{q-t} g(X_t) \to g(X_t)$, $q \downarrow t$, P_ν-a.s.. By dominated convergence, it follows that $\mathsf{E}_\nu f(X_t) g(Y_t) = \mathsf{E}_\nu f(X_t) g(X_t)$. By Exercise 3.33 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 3.32, 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 $*$).

3.6.2 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 3.6.3 *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

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

It therefore suffices to show (explain!) that

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

To this end, define

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

This is a d-system, and so by BN Lemma 3.7 it suffices to show that $\mathcal{S}$ contains a π-system generating $\mathcal{F}_\infty^X$. The appropriate π-system 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

$$\mathbf{1}_{\{A\}} = \prod_{k=1}^n \mathbf{1}_{\{A_k\}}(X_{t_k}). \tag{3.6.3}$$

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

$$\mathsf{E}(Z \mid \mathcal{F}_t^\nu) = \mathsf{E}(Z \mid \mathcal{F}_{t+}^\nu), \quad \mathsf{P}_\nu - \text{a.s.} \tag{3.6.4}$$

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, $\|g^h - g^0\|_\infty \to 0$, 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 (3.6.4).

The only thing left to prove is that we can replace Z by $\mathbf{1}_{\{A\}}$ from (3.6.3). 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 $Z = \prod_{i=1}^n f_i^m(X_{t_i})$ (3.6.4) holds. Use monotone convergence, to obtain that (3.6.4) holds for $Z = \mathbf{1}_{\{A\}}$ given in (3.6.3). 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 3.6.4 *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 3.6.5 *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 3.6.6 *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 3.6.7 (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 3.6.4 and 3.6.5 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

$$\int_A Z \circ \theta_t d\mathsf{P}_\nu = \int_A \mathsf{E}_{X_t} Z d\mathsf{P}_\nu. \tag{3.6.5}$$

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 (3.6.5) 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 (3.6.5) 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 again a Feller-Dynkin canonical cadlag process X with state space $(E_\delta, \mathcal{E}_\delta)$, where $E \subseteq \mathbf{R}^d, \mathbf{Z}_+^d$. This is a Markov process with respect to the usual augmentation $(\mathcal{F}_t)_t$ of the natural filtration of the canonical process on the compactified state space E_δ. As before, we denote shift operators by θ_t.

In this section we will prove that for Feller-Dynkin processes the Markov property of Theorem 3.6.7 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 3.6.8 (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.*

$$\mathsf{E}_\nu(Z \circ \theta_\tau \,|\, \mathcal{F}_\tau) = \mathsf{E}_{X_\tau}(Z). \tag{3.6.6}$$

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 (3.6.6) holds for Z an $\mathcal{F}^X_\infty$-measurable random variable. A similar reasoning as in the proof of Theorem 3.6.3 shows (check yourself) that we can restrict to showing (3.6.6) 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,

$$\mathsf{E}_\nu\Big(\prod_i f_i(X_{t_i}) \circ \theta_{\tau_n} \,|\, \mathcal{F}_{\tau_n}\Big) = \mathsf{E}_{X_{\tau_n}} \prod_i f_i(X_{t_i}) = \mathbf{1}_{\{\tau_n < \infty\}} g(X_{\tau_n}),$$

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 (3.6.6) 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 3.6.7.

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

3.7 Countable state Markov processes

3.8 Regenerative processes and killed Feller processes

3.8.1 Regenerative processes

3.8.2 Killed Feller processes

3.9 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. 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 3.3.11 and Theorem 3.3.12.

Exercise 3.10 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 3.11 Consider the situation of Exercise 3.10. 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 $\{X_{\sigma_x} = x, \sigma_x < \infty\} \subseteq \{\sigma_x \circ \theta_{\sigma_x} = 0, \sigma_x < \infty\}$.

ii) Using the strong Markov property, show that

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

iii) Conclude that $\mathsf{P}_x\{X_{\sigma_x} = x, \sigma_x < \infty\} = 0$, i.e. a canonical Markov process with right-continuous paths, satisfying the strong Markov property can only leave a holding point by a jump.

Exercise 3.12 Prove Corollary 3.3.7 and Lemma 3.3.8.

Exercise 3.13 Show for Example 3.3.15 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 3.14 Show that the maps ϕ and ψ in the proof of Theorem 3.3.13 are Borel measurable.

Exercise 3.15 Derive the expression for the joint density of BM and its running maximum given in Corollary 3.3.14.

Exercise 3.16 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 3.17 Prove Corollary 3.3.16

Exercise 3.18 Consider the Poisson process. Define an appropriate Banach space such that the associated transition function is a strongly continuous semigroup. Show that the generator of the Poisson process is given by

$$\mathsf{A}f(x) = \lambda f(x+1) - \lambda f(x).$$

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

Exercise 3.19 Show that the generator of the Ornstein-Uhlenbeck process (cf. Example 3.1.7 (B) and 3.4.7) 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 3.5.5. 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 3.20 Prove the Integration Lemma.

Exercise 3.21 Prove the claim made in Example 3.4.13. Hint: to derive the explicit expression for the resolvent kernel it is needed to calculate integrals of the form

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

To this end, first perform the substitution $t = (b/a)s^2$. Next, make a change of variables $u = s - 1/s$ and observe that $u(s) = s - 1/s$ is a continuously differentiable bijective function from $(0, \infty)$ to $\mathbf{R}$, the inverse $u^{-1} : \mathbf{R} \to (0, \infty)$ of which satisfies $u^{-1}(t) - u^{-1}(-t) = t$, whence $(u^{-1})'(t) + (u^{-1})'(t) = 1$.

Exercise 3.22 Prove the validity of the expression for the resolvent of the Markov jump process in Example 3.4.14.

Exercise 3.23 Show that the Markov process from Example 3.2.5 is a Feller-Dynkin process if $PC_0(E) \subset C_0(E)$. Give an example of a Markov jump process that is not a Feller-Dynkin process.

Exercise 3.24 Prove Lemma 3.5.6. 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 3.5.7.

Exercise 3.25 In the proof of Lemma 3.5.11, show that $\mathsf{P}\{\eta_r \geq nt\} \leq \hat{p}^n$ for $n = 0, 1, \ldots$.

Exercise 3.26 **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 3.27 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 3.28 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

$$P_t f(x) = \frac{1}{(2\pi t)^{d/2}} \int_{\mathrm{R}^d} f(y) e^{-\|y-x\|^2/2t} dy,$$

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 3.29 (Continuation of Exercise 3.28) Let X be an $\mathbf{R}^d$-valued canonical continuous Feller-Dynkin process, such that $X_t = X_0 + W_t$, where X_0 is an $\mathbf{R}^d$-valued r.v. and $(W_t)_t$ a standard d-dimensional BM that is independent of X_0. Notice that X is strong Markov.

We would like to show that the generator is defined by

$$\mathsf{A} f(x) = \tfrac{1}{2} \Delta f(x), \tag{3.9.1}$$

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

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

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 $\{y \mid \|y - x\| = r\}$.

ii) Show the validity of (3.9.1) for $f \in C_0^2(\mathbf{R}^d)$ (use the characteristic operator). Argue that this implies $\mathcal{D}(\mathsf{A}) = C_0^2(\mathbf{R}^d)$.

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 3.27.

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

Exercise 3.30 Prove (3.6.1) in the proof of Lemma 3.6.1.

Exercise 3.31 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 3.32 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 3.33 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 $\mathcal{H} = \{h : E \times E \to \mathbf{R} \,|\, h \ \mathcal{E} \times \mathcal{E} -$ measurable, $\|h\|_\infty < \infty, \mathsf{E}h(X, Y) = \mathsf{E}h(X, X)\}$.

Exercise 3.34 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 3.35 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 3.33.