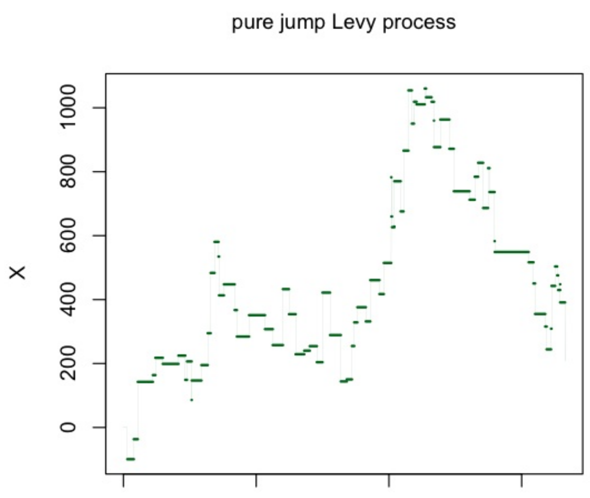
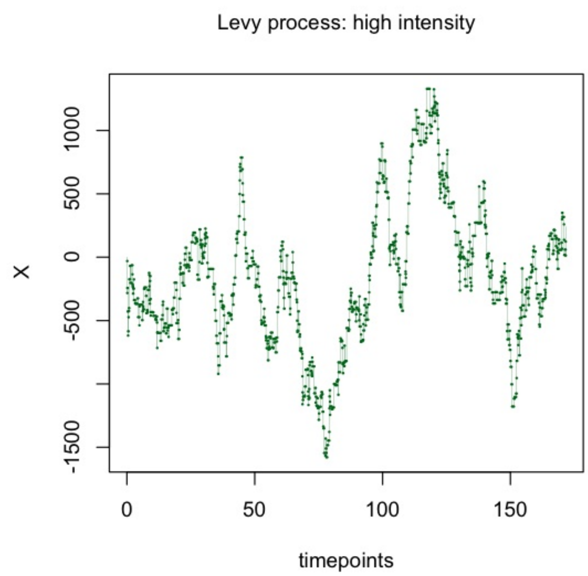
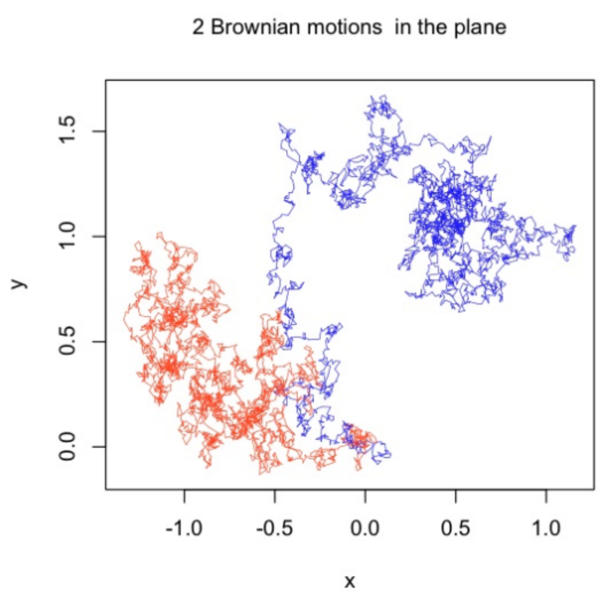
Markov- prosessien jatkokurssi
Markov processes







C. Geiss and S. Geiss

December 8, 2021

Contents

1 Introduction

Why should one study Markov processes? The class of Markov processes contains the

- Brownian motion,
- Lévy process,
- Feller processes,

where these classes are contained in each other, the class of Brownian motions is the smallest class. Moreover,

- solutions to certain SDEs are Markov processes.

Looking from another perspective we will see useful relations between Markov processes and

- martingale problems,
- diffusions,
- second order differential and integral operators.

The Markov processes are named after the Russian mathematician Andrey Andreyevich Markov (14 June 1856 – 20 July 1922).

2 Definition of a Markov process

For the following we let

(1) $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space,

(2) $(E, \mathcal{E})$ be a measurable space,

(3) $\mathbf{T} \subseteq \mathbb{R} \cup \{\infty\} \cup \{-\infty\}$ with $T \neq \emptyset$.

Let us fix some notation:

- We call $X = \{X_t; t \in \mathbf{T}\}$ a stochastic process if

$$X_t : (\Omega, \mathcal{F}) \to (E, \mathcal{E}) \quad \text{for all} \quad t \in \mathbf{T}.$$

- The map $t \mapsto X_t(\omega)$ is called a path of X.
- We say that $\mathbb{F} = \{\mathcal{F}_t; t \in \mathbf{T}\}$ is a *filtration* if $\mathcal{F}_t \subseteq \mathcal{F}$ is a sub-σ-algebra for any $t \in \mathbf{T}$ and it holds $\mathcal{F}_s \subseteq \mathcal{F}_t$ for $s \leq t$.
- The process X is *adapted* to $\mathbb{F}$ if X_t is $\mathcal{F}_t$ measurable for all $t \in \mathbf{T}$.
- The *natural filtration* $\mathbb{F}^X = \{\mathcal{F}_t^X; t \in \mathbf{T}\}$ of $X = \{X_t; t \in \mathbf{T}\}$ is given by $\mathcal{F}_t^X := \sigma(X_s; s \leq t, s \in \mathbf{T})$.

Obviously, X is always adapted to its natural filtration $\mathbb{F}^X = \{\mathcal{F}_t^X; t \in \mathbf{T}\}$. Now we turn to our main definition:

Definition 2.1 (Markov process)**.** The stochastic process X is called a *Markov process* w.r.t. $\mathbb{F}$ if and only if

(1) X is adapted to $\mathbb{F}$,

(2) for all $t \in \mathbf{T}$, $A \in \mathcal{F}_t$, and $B \in \sigma(X_s; s \geq t)$ one has

$$\mathbb{P}(A \cap B | X_t) = \mathbb{P}(A|X_t)\mathbb{P}(B|X_t) \text{ a.s.,}$$

i.e. the σ-algebras $\mathcal{F}_t$ and $\sigma(X_s; s \geq t, s \in \mathbf{T})$ are conditionally independent given X_t.

Remark 2.2.

(1) We recall that we define the conditional probability using conditional expectation as

$$\mathbb{P}(C|X_t) := \mathbb{P}(C|\sigma(X_t)) = \mathbb{E}[\mathbb{1}_C|\sigma(X_t)].$$

(2) If X is a Markov process w.r.t. $\mathbb{F}$, then X is a Markov process w.r.t. $\mathbb{F}^X$.

(3) If X is a Markov process w.r.t. its natural filtration $\mathbb{F}^X$, then the Markov property is preserved if one reverses the order in $\mathbf{T}$.

The following result is our first main result:

Theorem 2.3. *Let X be $\mathbb{F}$-adapted. Then the following conditions are equivalent:*

(1) *X is a* MARKOV *process w.r.t. $\mathbb{F}$.*

(2) *For each $t \in \mathbf{T}$ and each bounded $\sigma(X_s; s \geq t, s \in \mathbf{T})$-measurable $Y : \Omega \to \mathbb{R}$ one has*

$$\mathbb{E}[Y|\mathcal{F}_t] = \mathbb{E}[Y|X_t] \text{ a.s.} \tag{2.1}$$

(3) *If $s, t \in \mathbf{T}$ and $t \leq s$, then*

$$\mathbb{E}[f(X_s)|\mathcal{F}_t] = \mathbb{E}[f(X_s)|X_t] \text{ a.s.}$$

for all bounded $f : (E, \mathcal{E}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Proof. (1) $\implies$ (2) We can decompose $Y = Y^+ - Y^-$ into the positive and negative part, and each part can be approximated from below point-wise by $\sigma(X_s; s \geq t, s \in \mathbf{T})$-measurable simple functions. Therefore it suffices to show (2.1) for $Y = \mathbb{1}_B$ where $B \in \sigma(X_s; s \geq t, s \in \mathbf{T})$. In fact, for $A \in \mathcal{F}_t$ we have, a.s.,

$$\begin{aligned}
\mathbb{E}(\mathbb{E}[Y|\mathcal{F}_t]\mathbb{1}_A) &= \mathbb{E}\mathbb{1}_A\mathbb{1}_B \\
&= \mathbb{P}(A \cap B) \\
&= \mathbb{E}\mathbb{P}(A \cap B|X_t) \\
&= \mathbb{E}\mathbb{P}(A|X_t)\mathbb{P}(B|X_t)
\end{aligned}$$

$$
\begin{aligned}
&= \mathbb{E}\mathbb{E}[\mathbb{1}_A|X_t]\mathbb{P}(B|X_t) \\
&= \mathbb{E}\mathbb{1}_A\mathbb{P}(B|X_t) \\
&= \mathbb{E}(\mathbb{E}[Y|X_t]\mathbb{1}_A)
\end{aligned}
$$

which implies (2).
(2) $\Longrightarrow$ (1) If $A \in \mathcal{F}_t$ and $B \in \sigma(X_s; s \geq t, s \in \mathbf{T})$, then, a.s.,

$$
\begin{aligned}
\mathbb{P}(A \cap B|X_t) &= \mathbb{E}[\mathbb{1}_{A\cap B}|X_t] \\
&= \mathbb{E}[\mathbb{E}[\mathbb{1}_{A\cap B}|\mathcal{F}_t]|X_t] \\
&= \mathbb{E}[\mathbb{1}_A\mathbb{E}[\mathbb{1}_B|\mathcal{F}_t]|X_t] \\
&= \mathbb{E}[\mathbb{1}_A|X_t]\mathbb{E}[\mathbb{1}_B|X_t],
\end{aligned}
$$

which implies (1).
(2) $\Longrightarrow$ (3) is trivial. (3) $\Longrightarrow$ (2) To apply the Monotone Class Theorem for functions we let

$$
\begin{aligned}
\mathcal{H} := \{Y; \quad & Y \text{ is bounded and } \sigma(X_s; s \geq t, s \in \mathbf{T}) - \text{measurable} \\
& \text{such that (2.1) holds}\}.
\end{aligned}
$$

Then $\mathcal{H}$

- is a vector space,
- contains the constants,
- is closed under bounded and monotone limits.

(a) For bounded $f_i : (E, \mathcal{E}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ and $t \leq s_1 < ... < s_n$, $n \geq 1$, we show that

$$
Y = \Pi_{i=1}^n f_i(X_{s_i}) \in \mathcal{H}. \tag{2.2}
$$

We show (2.2) by induction over n. The case $\underline{n = 1}$ is assertion (3).
$\underline{n > 1}$: Assume that the statement is true for $n - 1$. Then we get, a.s.,

$$
\begin{aligned}
\mathbb{E}[Y|\mathcal{F}_t] &= \mathbb{E}[\mathbb{E}[Y|\mathcal{F}_{s_{n-1}}]|\mathcal{F}_t] \\
&= \mathbb{E}[\Pi_{i=1}^{n-1} f_i(X_{s_i})\mathbb{E}[f_n(X_{s_n})|\mathcal{F}_{s_{n-1}}]|\mathcal{F}_t] \\
&= \mathbb{E}[\Pi_{i=1}^{n-1} f_i(X_{s_i})\mathbb{E}[f_n(X_{s_n})|X_{s_{n-1}}]|\mathcal{F}_t].
\end{aligned}
$$

By the Factorization Lemma A.1 there exists a $h : (E, \mathcal{E}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that $\mathbb{E}[f_n(X_{s_n})|X_{s_{n-1}}] = h(X_{s_{n-1}})$ a.s. By the induction hypothesis we get, a.s.,

$$
\mathbb{E}[\Pi_{i=1}^{n-1} f_i(X_{s_i})h(X_{s_{n-1}})|\mathcal{F}_t] = \mathbb{E}[\Pi_{i=1}^{n-1} f_i(X_{s_i})h(X_{s_{n-1}})|X_t].
$$

And finally, by the tower property, since $\sigma(X_t) \subseteq \mathcal{F}_{s_{n-1}}$, a.s.,

$$\begin{aligned}
\mathbb{E}[\Pi_{i=1}^{n-1} f_i(X_{s_i}) h(X_{s_{n-1}})|X_t] &= \mathbb{E}[\Pi_{i=1}^{n-1} f_i(X_{s_i}) \mathbb{E}[f_n(X_{s_n})|\mathcal{F}_{s_{n-1}}]|X_t] \\
&= \mathbb{E}[\mathbb{E}[\Pi_{i=1}^{n-1} f_i(X_{s_i}) f_n(X_{s_n})|\mathcal{F}_{s_{n-1}}]|X_t] \\
&= \mathbb{E}[\Pi_{i=1}^{n} f_i(X_{s_i})|X_t].
\end{aligned}$$

(b) Now we apply the Monotone Class Theorem A.2. From step (a) we know that $\mathbb{1}_A \in \mathcal{H}$ for any $A \in \mathcal{A}$ with

$$\mathcal{A} = \{\{\omega \in \Omega; X_{s_1}(\omega) \in I_1, ..., X_{s_n}(\omega) \in I_n\} : I_k \in \mathcal{B}(\mathbb{R}), s_k \in \mathbf{T}, s_k \geq t, n \geq 1\}$$

where $\sigma(\mathcal{A}) = \sigma(X_s; s \geq t, s \in \mathbf{T})$. Therefore

$$\{Y; Y \text{ is bounded and } \sigma(X_s; s \geq t, s \in \mathbf{T}) - \text{measurable}\} \subseteq \mathcal{H}.$$

$\square$

3 Transition functions

In this section we assume that $\mathbf{T} = [0, \infty)$.

Definition 3.1 (Markov transition function)**.**

(1) A family $(P_{t,s})_{0 \le t \le s < \infty}$ is called Markov *transition function on* $(E, \mathcal{E})$ if all $P_{s,t} : E \times \mathcal{E} \to [0,1]$ satisfy that

  (a) $A \mapsto P_{t,s}(x, A)$ is a probability measure on $(E, \mathcal{E})$ for each (t, s, x),

  (b) $x \mapsto P_{t,s}(x, A)$ is $\mathcal{E}$-measurable for each (t, s, A),

  (c) $P_{t,t}(x, A) = \delta_x(A)$,

  (d) if $0 \le t < s < u$, then the Chapman-Kolmogorov equation

$$P_{t,u}(x, A) = \int_E P_{s,u}(y, A) P_{t,s}(x, dy)$$

  holds for all $x \in E$ and $A \in \mathcal{E}$.

(2) The Markov transition function $(P_{t,s})_{s \le t}$ is *homogeneous* if and only if $P_{t,s} = P_{0,s-t}$ for all $0 \le t \le s < \infty$.

(3) We say that a Markov process X w.r.t. $\mathbb{F}$ is associated to the Markov transition function $(P_{t,s})_{0 \le t \le s < \infty}$ provided that

$$\mathbb{E}[f(X_s) | \mathcal{F}_t] = \int_E f(y) P_{t,s}(X_t, dy) \text{ a.s.} \tag{3.1}$$

for all $0 \le t \le s < \infty$ and all bounded $f : (E, \mathcal{E}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

(4) Let μ be a probability measure on $(E, \mathcal{E})$ such that $\mu(A) = \mathbb{P}(X_0 \in A)$. Then μ is called initial distribution of X.

Remark 3.2.

(1) There exist Markov processes which do not possess transition functions (see [2, Remark 1.11, page 446]).

(2) Using monotone convergence one can check that the map

$$x \mapsto \int_E f(y) P_{t,s}(x, dy)$$

is $(\mathcal{E}, \mathcal{B}(\mathbb{R}))$-measurable for a bounded $f : (E, \mathcal{E}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Proposition 3.3. *A* Markov *process w.r.t.* $\mathbb{F}$ *having* $(P_{t,s})_{t\le s}$ *as transition function satisfies for* $0 \le t_1 < t_2 < ... < t_n$ *and bounded* $f : (E^n, \mathcal{E}^{\otimes n}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ *the relation*

$$\mathbb{E}f(X_{t_1}, ..., X_{t_n}) = \int_E \mu(dx_0) \int_E P_{0,t_1}(x_0, dx_1) ... \int_E P_{t_{n-1},t_n}(x_{n-1}, dx_n) f(x_1, ..., x_n).$$

4 Existence of Markov processes

Given a distribution μ and MARKOV transition functions $\{P_{t,s}(x,A)\}$, does there always exist a MARKOV process with initial distribution μ and transition function $\{P_{t,s}(x,A)\}$?

Definition 4.1.

(1) For a measurable space $(E,\mathcal{E})$ and a non-empty index set $\mathbf{T}$ we let

$$\Omega := E^{\mathbf{T}}, \quad \mathcal{F} := \mathcal{E}^{\mathbf{T}} := \sigma(X_t; t \in \mathbf{T}),$$

where $X_t : \Omega \to E$ is the coordinate map

$$X_t(\omega) = \omega(t) \quad \text{where} \quad \omega = (\omega(t))_{t\in\mathbf{T}} \in \Omega.$$

(2) Let $\mathrm{Fin}(\mathbf{T}) := \{\mathrm{J} \subseteq \mathbf{T}; 0 < |\mathrm{J}| < \infty\}$ where in J all elements are pairwise distinct.

(3) For $J = \{t_1, ..., t_n\} \in \mathrm{Fin}(\mathbf{T})$ we define the projections $\pi_J : \Omega \to E^J$ by

$$\pi_J(\omega) := (\omega(t_1), ..., \omega(t_n)) = (X_{t_1}, ..., X_{t_n}) \in E^J.$$

(4) A set $\{\mathbf{P}_J : \mathbf{P}_J$ is a probability measure on $(E^J, \mathcal{E}^J), J \in \mathrm{Fin}(\mathbf{T})\}$ is called a set of *finite-dimensional distributions*.

(5) A set of of finite-dimensional distributions $\{\mathbf{P}_J : J \in \mathrm{Fin}(\mathbf{T})\}$ is called *Kolmogorov consistent* (or compatible or projective) provided that the following holds.

(a) **Symmetry:** One has

$$\mathbf{P}_{t_{\sigma(1)},...,t_{\sigma(n)}}(A_{\sigma(1)} \times ... \times A_{\sigma(n)}) = \mathbf{P}_{t_1,...,t_n}(A_1 \times ... \times A_n)$$

for any permutation $\sigma : \{1, ..., n\} \to \{1, ..., n\}$.

(b) **Projection property:** One has

$$\mathbf{P}_J = \mathbf{P}_K \circ (\pi_J \mid_{E^K})^{-1}$$

for all $J \subseteq K$ with $J, K \in \mathrm{Fin}(\mathbf{T})$.

Theorem 4.2 (KOLMOGOROV's extension theorem, DANIELL-KOLMOGOROV Theorem)**.** *Let E be a complete, separable metric space and $\mathcal{E} = \mathcal{B}(E)$. Let $\mathbf{T}$ be a non-empty set. Suppose that for each $J \in \text{Fin}(\mathbf{T})$ there exists a probability measure P_J on $(E^J, \mathcal{E}^J)$ and that*

$$\{\mathbf{P}_J; J \in \text{Fin}(\mathbf{T})\}$$

is Kolmogorov consistent. Then there exists a unique probability measure $\mathbb{P}$ on $(E^{\mathbf{T}}, \mathcal{E}^{\mathbf{T}})$ such that

$$\mathbf{P}_J = \mathbb{P} \circ \pi_J^{-1} \quad \text{on} \quad (E^J, \mathcal{E}^J).$$

For the proof see, for example [5, Theorem 2.2 in Chapter 2]. The main result of this section is the following existence theorem that will be deduced from Theorem 4.2.

Theorem 4.3 (Existence of MARKOV processes)**.** *Let $E = \mathbb{R}^d$, $\mathcal{E} = \mathcal{B}(\mathbb{R}^d)$, and $\mathbf{T} \subseteq [0, \infty)$. Assume that μ is a probability measure on $(E, \mathcal{E})$ and that*

$$\{P_{t,s}(x, A); 0 \le t \le s < \infty, \, x \in E, \, A \in \mathcal{E}\}$$

is a MARKOV transition function (Definition 3.1). If $J = \{t_1, ..., t_n\} \subseteq \mathbf{T}$ and $\{s_1, ..., s_n\} = \{t_1, ..., t_n\}$ with $s_1 < ... < s_n$, i.e. the $t_k{}'s$ are re-arranged according to their size, we define

$$\begin{aligned} \mathbf{P}_J(A_1 \times ... \times A_n) \;:=\; & \int_E ... \int_E \mathbb{1}_{A_1 \times ... \times A_n}(x_1, .., x_n) \mu(dx_0) P_{0,s_1}(x_0, dx_1) \\ & ... P_{s_{n-1}, s_n}(x_{n-1}, dx_n). \end{aligned} \tag{4.1}$$

Then there exists a probability measure $\mathbb{P}$ on $(E^{\mathbf{T}}, \mathcal{E}^{\mathbf{T}})$ such that the coordinate mappings, i.e.

$$X_t : E^{\mathbf{T}} \to \mathbb{R}^d : \omega \mapsto \omega(t),$$

form a Markov process w.r.t. $\mathbb{F}^X$ with the MARKOV transition function $(P_{t,s})_{0 \le t \le s < \infty}$.

Remark 4.4. Using the monotone convergence one can show that (4.1) implies that for any bounded $f : (E^n, \mathcal{E}^n) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ it holds

$$\begin{aligned} \mathbb{E} f(X_{s_1}, ..., X_{s_n}) \;=\; & \int_E ... \int_E f(x_1, .., x_n) \mu(dx_0) P_{0,s_1}(x_0, dx_1) \\ & ... P_{s_{n-1}, s_n}(x_{n-1}, dx_n). \end{aligned} \tag{4.2}$$

Proof of Theorem 4.3. (a) By construction, $\mathbf{P}_J$ is a probability measure on $(E^J, \mathcal{E}^J)$. We show that the set $\{\mathbf{P}_J; J \in \mathrm{Fin}(\mathbf{T})\}$ is Kolmogorov consistent. The symmetry follows by construction, we only need to verify the projection property. Consider $K \subseteq J$ with

$$K = \{s_{i_1} < \cdots < s_{i_k}\} \subseteq J = \{s_1 < ... < s_n\}$$

and $1 \leq k < n$, and

$$\mathbf{P}_{J,K} : E^J \to E^K : (x_1, ..., x_n) \mapsto (x_{i_1}, ..x_{i_k}).$$

We have $\mathbf{P}_{J,K}^{-1}(B_1 \times ... \times B_k) = A_1 \times ... \times A_n$ with $A_i \in \{B_1, ..., B_k, E\}$. Let us assume, for example, that $k = n - 1$ and

$$A_1 \times ... \times A_n = B_1 \times ... \times B_{n-2} \times E \times B_n.$$

Then

$$\begin{aligned}
&\mathbf{P}_J(A_1 \times ... \times A_n) \\
&= \int_E ... \int_E \mathbb{1}_{B_1 \times ... \times B_{n-2} \times E \times B_n}(x_1, ..., x_n) \\
&\qquad\qquad \mu(dx_0) P_{0,s_1}(x_0, dx_1) ... P_{s_{n-1},s_n}(x_{n-1}, dx_n) \\
&= \mathbf{P}_{\{s_1,...,s_{n-2},s_n\}}(B_1 \times ... \times B_{n-2} \times B_n)
\end{aligned}$$

since, by the Chapman-Kolmogorov equation, we have

$$\int_E P_{s_{n-2},s_{n-1}}(x_{n-2}, dx_{n-1}) P_{s_{n-1},s_n}(x_{n-1}, dx_n) = P_{s_{n-2},s_n}(x_{n-2}, dx_n).$$

(b) Now we check that the process is a Markov process. According to Definition 2.1 we need to show that

$$\mathbb{P}(A \cap B | X_t) = \mathbb{P}(A | X_t) \mathbb{P}(B | X_t) \text{ a.s.} \tag{4.3}$$

for $A \in \mathcal{F}_t^X = \sigma(X_u; u \leq t)$ and $B \in \sigma(X_s; s \geq t)$. We only prove the special case

$$\mathbb{P}(X_u \in B_1, X_s \in B_3, | X_t) = \mathbb{P}(X_u \in B_1 | X_t) \mathbb{P}(X_s \in B_3 | X_t) \text{ a.s.}$$

for $u < t < s$ and $B_i \in \mathcal{E}$. For this we show that it holds

$$\mathbb{E}\left[\mathbb{1}_{B_1}(X_u) \mathbb{1}_{B_3}(X_s) \mathbb{1}_{B_2}(X_t)\right] = \mathbb{E}\left[\mathbb{P}(X_u \in B_1 | X_t) \mathbb{P}(X_s \in B_3 | X_t) \mathbb{1}_{B_2}(X_t)\right].$$

Indeed, by (4.1),

$$\mathbb{E}\mathbb{1}_{B_1}(X_u)\mathbb{1}_{B_3}(X_s)\mathbb{1}_{B_2}(X_t) = \int_E\int_E\int_E\int_E \mathbb{1}_{B_1\times B_2\times B_3}(x_1,x_2,x_3) \\ \mu(dx_0)P_{0,u}(x_0,dx_1)P_{u,t}(x_1,dx_2)P_{t,s}(x_2,dx_3).$$

Using the tower property we get

$$\begin{aligned}&\mathbb{E}\left[\mathbb{P}(X_s\in B_3|X_t)\mathbb{P}(X_u\in B_1|X_t)\mathbb{1}_{B_2}(X_t)\right]\\ &=\mathbb{E}\left[(\mathbb{E}[\mathbb{1}_{B_3}(X_s)|X_t])\mathbb{1}_{B_1}(X_u)\mathbb{1}_{B_2}(X_t)\right]\\ &=\mathbb{E}\left[P_{t,s}(X_t,B_3)\mathbb{1}_{B_1}(X_u)\mathbb{1}_{B_2}(X_t)\right].\end{aligned}$$

To see that $\mathbb{E}[\mathbb{1}_{B_3}(X_s)|X_t]) = P_{t,s}(X_t,B_3)$ we write

$$\begin{aligned}\mathbb{E}\mathbb{1}_{B_3}(X_s)\mathbb{1}_{B}(X_t) &= \int_E\int_E\int_E \mathbb{1}_{B_3}(x_2)\mathbb{1}_B(x_1)\mu(dx_0)P_{0,t}(x_0,dx_1)P_{t,s}(x_1,dx_2)\\ &= \int_E\int_E\int_E \mathbb{1}_B(x_1)\mu(dx_0)P_{0,t}(x_0,dx_1)P_{t,s}(x_1,B_3)\\ &= \mathbb{E}P_{t,s}(X_t,B_3)\mathbb{1}_B(X_t).\end{aligned}$$

where we used (4.2) for $f(x_1) = \mathbb{1}_B(x_1)P_{t,s}(x_1,B_3)$. Again by (4.2), now for $f(X_u,X_t) := P_{t,s}(X_t,B_3)\mathbb{1}_{B_1}(X_u)\mathbb{1}_{B_2}(X_t)$, we get that

$$\begin{aligned}&\mathbb{E}P_{t,s}(X_t,B_3)\mathbb{1}_{B_1}(X_u)\mathbb{1}_{B_2}(X_t)\\ &=\int_E\int_E\int_E P_{t,s}(x_2,B_3)\mathbb{1}_{B_1\times B_2}(x_1,x_2)\mu(dx_0)P_{0,u}(x_0,dx_1)P_{u,t}(x_1,dx_2)\\ &=\int_E\int_E\int_E\int_E \mathbb{1}_{B_1\times B_2\times B_3}(x_1,x_2,x_3)\\ &\qquad\qquad \mu(dx_0)P_{0,u}(x_0,dx_1)P_{u,t}(x_1,dx_2)P_{t,s}(x_2,dx_3).\end{aligned}$$

□

5 A reminder on stopping and optional times

For $(\Omega, \mathcal{F})$ we assume a filtration $\mathbb{F} = \{\mathcal{F}_t; t \in \mathbf{T}\}$ where $\mathbf{T} = [0, \infty) \cup \{\infty\}$ and $\mathcal{F} = \mathcal{F}_\infty = \sigma\left(\bigcup_{s \in [0,\infty)} \mathcal{F}_s\right)$. Moreover, we set

$$\mathcal{F}_{t+} := \bigcap_{s>t} \mathcal{F}_s, \quad t \in [0, \infty), \qquad\qquad \mathcal{F}_{\infty+} := \mathcal{F}_\infty,$$

$$\mathcal{F}_{t-} := \sigma\left(\bigcup_{0 \le s < t} \mathcal{F}_s\right), \quad t \in (0, \infty], \quad \mathcal{F}_{0-} := \mathcal{F}_0.$$

Therefore, for all $t \in \mathbf{T}$ one has that

$$\mathcal{F}_{t-} \subseteq \mathcal{F}_t \subseteq \mathcal{F}_{t+}.$$

Definition 5.1.

(1) A map $\tau : \Omega \to \mathbf{T}$ is called a *stopping time w.r.t.* $\mathbb{F}$ provided that

$$\{\tau \le t\} \in \mathcal{F}_t \quad \text{for all} \quad t \in [0, \infty).$$

(2) The map $\tau : \Omega \to \mathbf{T}$ is called an *optional time w.r.t* $\mathbb{F}$ provided that

$$\{\tau < t\} \in \mathcal{F}_t \quad \text{for all} \quad t \in [0, \infty).$$

(3) For a stopping time $\tau : \Omega \to \mathbf{T}$ w.r.t. $\mathbb{F}$ we define

$$\mathcal{F}_\tau := \{A \in \mathcal{F} : A \cap \{\tau \le t\} \in \mathcal{F}_t \quad \forall t \in [0, \infty)\}.$$

(4) For an optional time $\tau : \Omega \to \mathbf{T}$ w.r.t. $\mathbb{F}$ we define

$$\mathcal{F}_{\tau+} := \{A \in \mathcal{F} : A \cap \{\tau < t\} \in \mathcal{F}_t \quad \forall t \in [0, \infty)\}.$$

Remark 5.2.

(1) For a stopping time we have that $\{\tau = \infty\} = \{\tau < \infty\}^c \in \mathcal{F}_\infty$ because

$$\{\tau < \infty\} = \bigcup_{n \in \mathbb{N}} \{\tau \le n\} \in \mathcal{F}_\infty.$$

(2) For an optional time we have that $\{\tau < \infty\} \in \mathcal{F}_\infty$.

(3) $\mathcal{F}_\tau$ and $\mathcal{F}_{\tau+}$ are σ algebras.

Definition 5.3. The filtration $\{\mathcal{F}_t; t \in \mathbf{T}\}$ is called *right-continuous* if $\mathcal{F}_t = \mathcal{F}_{t+}$ for all $t \in [0, \infty)$.

Lemma 5.4. *If τ and σ are stopping times w.r.t. $\mathbb{F}$, then*

(1) $\tau + \sigma$,

(2) $\tau \wedge \sigma = \min\{\tau, \sigma\}$,

(3) $\tau \vee \sigma = \max\{\tau, \sigma\}$,

are stopping times w.r.t. $\mathbb{F}$.

Lemma 5.5.

(1) *For $t_0 \in \mathbf{T}$ the map $\tau(\omega) \equiv t_0$ for all $\omega \in \Omega$ is a stopping time and one has $\mathcal{F}_{t_0} = \mathcal{F}_\tau$.*

(2) *Every stopping time is an optional time.*

(3) *If $\{\mathcal{F}_t; t \in \mathbf{T}\}$ is right-continuous, then every optional time is a stopping time.*

(4) *The map τ is an $\{\mathcal{F}_t; t \in \mathbf{T}\}$ optional time if and only if τ is an $\{\mathcal{F}_{t+}; t \in \mathbf{T}\}$ stopping time.*

Proof. (1) follows from

$$\{\tau \leq t\} = \begin{cases} \Omega; & t_0 \leq t \\ \emptyset; & t_0 > t \end{cases} .$$

(2) Let τ be a stopping time. Then

$$\{\tau < t\} = \bigcup_{n=1}^{\infty} \underbrace{\left\{\tau \leq t - \frac{1}{n}\right\}}_{\in \mathcal{F}_{t-\frac{1}{n}} \subseteq \mathcal{F}_t} \in \mathcal{F}_t.$$

(3) We have that $\{\tau \le t\} = \bigcap_{n=1}^{\infty} \underbrace{\left\{\tau < t + \frac{1}{n}\right\}}_{\in \mathcal{F}_{t+\frac{1}{n}}}$. Because of

$$\bigcap_{n=1}^{M} \left\{\tau < t + \frac{1}{n}\right\} = \left\{\tau < t + \frac{1}{M}\right\} \in \mathcal{F}_{t+\frac{1}{M}}$$

we get that $\{\tau \le t\} \in \mathcal{F}_{t+\frac{1}{M}} \quad \forall M \in \mathbb{N}^*$ and hence $\{\tau \le t\} \in \mathcal{F}_{t+} = \mathcal{F}_t$ since $\{\mathcal{F}_t; t \in \mathbf{T}\}$ is right-continuous.

(4) $\implies$ If τ is an $\{\mathcal{F}_t; t \in \mathbf{T}\}$ optional time, then $\{\tau < t\} \in \mathcal{F}_t$ implies $\{\tau < t\} \in \mathcal{F}_{t+}$ because $\mathcal{F}_t \subseteq \bigcap_{s>t} \mathcal{F}_s = \mathcal{F}_{t+}$. This means that τ is an $\{\mathcal{F}_{t+}; t \in \mathbf{T}\}$ optional time. Since $\{\mathcal{F}_{t+}; t \in \mathbf{T}\}$ is right-continuous, we conclude from (3) that τ is an $\{\mathcal{F}_{t+}; t \in \mathbf{T}\}$ stopping time.

$\impliedby$ If τ is an $\{\mathcal{F}_{t+}; t \in \mathbf{T}\}$ stopping time, then

$$\{\tau < t\} = \bigcup_{n=1}^{\infty} \underbrace{\{\tau \le t - \frac{1}{n}\}}_{\in \mathcal{F}_{(t-1/n)+} = \bigcap_{s>t-1/n} \mathcal{F}_s \subseteq \mathcal{F}_t} \in \mathcal{F}_t.$$

□

Lemma 5.6. *For stopping times $\sigma, \tau, \tau_1, \tau_2, \ldots$ w.r.t. $\mathbb{F}$ the following holds:*

(1) *τ is $\mathcal{F}_\tau$-measurable.*

(2) *If $\tau \le \sigma$, then $\mathcal{F}_\tau \subseteq \mathcal{F}_\sigma$.*

(3) *$\mathcal{F}_{\tau+} = \{A \in \mathcal{F} : A \cap \{\tau \le t\} \in \mathcal{F}_{t+} \quad \forall t \in [0, \infty)\}$.*

(4) *The map $\sup_n \tau_n : \Omega \to \mathbf{T}$ is a stopping time w.r.t. $\mathbb{F}$.*

6 Strong Markov processes

6.1 Strong Markov property

Definition 6.1 (progressively measurable)**.** Let E be a complete, separable metric space and $\mathcal{E} = \mathcal{B}(E)$.

(1) A process $X = \{X_t; t \in [0, \infty)\}$, with $X_t : \Omega \to E$ is called $\mathbb{F}$-progressively measurable if for all $t \geq 0$ it holds

$$X : ([0,t] \times \Omega, \mathcal{B}([0,t]) \otimes \mathcal{F}_t) \to (E, \mathcal{E}).$$

(2) We will say that a stochastic process X is right-continuous (left-continuous), if for all $\omega \in \Omega$ the functions

$$[0, \infty) \ni t \mapsto X_t(\omega) \in E$$

are right-continuous (left-continuous).

We will start with a technical lemma:

Lemma 6.2.

(1) *If X is $\mathbb{F}$-progressively measurable, then X is $\mathbb{F}$-adapted,*

(2) *If X is $\mathbb{F}$-adapted and right-continuous (or left-continuous), then X is $\mathbb{F}$-progressively measurable.*

(3) *If τ is an $\mathbb{F}$-stopping time and X is $\mathbb{F}$- progressively measurable, then $X_\tau : \{\tau < \infty\} \to E$ is $\mathcal{F}_\tau|_{\{\tau<\infty\}}$-measurable.*

(4) *For an $\mathbb{F}$-stopping time τ and a $\mathbb{F}$- progressively measurable process X the stopped process X^τ given by*

$$X_t^\tau(\omega) := X_{t\wedge\tau}(\omega)$$

is $\mathbb{F}$- progressively measurable,

(5) *If τ is an $\mathbb{F}$-optional time and X is $\mathbb{F}$- progressively measurable, then $X_\tau : \{\tau < \infty\} \to E$ is $\mathcal{F}_{\tau+}|_{\{\tau<\infty\}}$-measurable.*

Proof. The assertions (1), (2) and (5) are exercises.
(3) For $s \in [0, \infty)$ it holds

$$\{\tau \wedge t \leq s\} = \{\tau \leq s\} \cup \{t \leq s\} = \begin{cases} \Omega, & s \geq t \\ \{\tau \leq s\}, & s < t \end{cases} \in \mathcal{F}_t.$$

Hence $\tau \wedge t$ is $\mathcal{F}_t$-measurable. Next we observe that $h(\omega) := (\tau(\omega) \wedge t, \omega)$ is measurable as map

$$(\Omega, \mathcal{F}_t) \to ([0, t] \times \Omega, \mathcal{B}([0, t]) \otimes \mathcal{F}_t).$$

Also, since X is $\mathbb{F}$- progressively measurable, we have that

$$X : ([0, t] \times \Omega, \mathcal{B}([0, t]) \otimes \mathcal{F}_t) \to (E, \mathcal{E}) \tag{6.1}$$

and therefore

$$X \circ h : (\Omega, \mathcal{F}_t) \to (E, \mathcal{E}). \tag{6.2}$$

It holds that (3) is equivalent to

$$\{X_\tau \in B\} \cap \{\tau \leq t\} \in \mathcal{F}_t \quad \text{for all} \quad t \in [0, \infty).$$

Indeed this is true as

$$\{X_\tau \in B\} \cap \{\tau \leq t\} = \{X_{\tau \wedge t} \in B\} \cap \{\tau \leq t\}$$

which is in $\mathcal{F}_t$ because of (6.2) and since τ is a stopping time.
(4) It holds that the map $H(s, \omega) := (\tau(\omega) \wedge s, \omega)$ is measurable as map

$$([0, t] \times \Omega, \mathcal{B}([0, t]) \otimes \mathcal{F}_t) \to ([0, t] \times \Omega, \mathcal{B}([0, t]) \otimes \mathcal{F}_t)$$

for $t \geq 0$ since, for $r \in [0, t]$,

$$\{(s, \omega) \in [0, t] \times \Omega : \tau(\omega) \wedge s \in [0, r]\} = ([0, r] \times \Omega) \cup ((r, t] \times \{\tau \leq r\}).$$

Because of (6.1) we have for the composition $(X \circ H)(s, \omega) := X_{\tau(\omega) \wedge s}(\omega) = X_s^\tau(\omega)$ the measurability

$$X \circ H : ([0, t] \times \Omega, \mathcal{B}([0, t]) \otimes \mathcal{F}_t) \to (E, \mathcal{E}).$$

$\square$

Definition 6.3 (strong Markov process)**.** Assume that $\{X_t : t \geq 0\}$ is an $\mathbb{F}$-progressively measurable Markov process with homogeneous transition function $(P_t)_{t\geq 0}$ in the sense that $P_t = P_{0,t}$. The process X is called a *strong Markov process* if

$$\mathbb{P}(X_{\tau+t} \in A|\mathcal{F}_{\tau+}) = P_t(X_\tau, A) \text{ a.s.}$$

for all $t \geq 0$, $A \in \mathcal{E}$ and all $\mathbb{F}$-optional times $\tau : \Omega \to [0, \infty)$.

One can formulate the strong Markov property without transition functions:

Proposition 6.4. *Let X be an $\mathbb{F}$-progressively measurable process. Then, provided X is a Markov process with transition function, the following assertions are equivalent to Definition 6.3:*

(1) *For all For all $t \in \mathbf{T}$ and $A \in \mathcal{E}$ one has*

$$\mathbb{P}(X_{\tau+t} \in A|\mathcal{F}_{\tau+}) = \mathbb{P}(X_{\tau+t} \in A|X_\tau) \text{ a.s.}$$

for all $\mathbb{F}$-optional times $\tau : \Omega \to [0, \infty)$.

(2) *For all $t_1, ..., t_n \in \mathbf{T}$ and $A_1, ..., A_n \in \mathcal{E}$ one has*

$$\mathbb{P}(X_{\tau+t_1} \in A_1, ..., X_{\tau+t_n} \in A_n|\mathcal{F}_{\tau+}) = \mathbb{P}(X_{\tau+t_1} \in A_1, ..., X_{\tau+t_n} \in A_n|X_\tau) \text{ a.s.}$$

for all $\mathbb{F}$-optional times $\tau : \Omega \to [0, \infty)$.

6.2 Lévy processes are strong Markov processes

Definition 6.5. A process $X = \{X_t : t \geq 0\}$ is called Lévy process if the following holds:

(1) $X_0 \equiv 0$.

(2) The paths of X are càdlàg (i.e. they are right-continuous and have left limits).

(3) For all $0 \leq s \leq t < \infty$ one has $X_t - X_s \stackrel{d}{=} X_{t-s}$.

(4) For all $0 \leq s \leq t < \infty$ one has that $X_t - X_s$ is independent of $\mathcal{F}_s^X$.

The strong Markov property of a Lévy process will be obtained as follows:

Theorem 6.6. *Let X be a Lévy process. Assume that $\tau : \Omega \to [0,\infty)$ is an $\mathbb{F}^X$-optional time. Define the process $\tilde{X} = \{\tilde{X}_t; t \geq 0\}$ by*

$$\tilde{X}_t = (X_{t+\tau} - X_\tau), \quad t \geq 0.$$

Then the process $\tilde{X}$ is independent of $\mathcal{F}^X_{\tau+}$ and $\tilde{X}$ has the same distribution as X.

Proof. The finite dimensional distributions determine the law of a stochastic process. Hence it is sufficient to show for arbitrary $0 = t_0 < t_1 < ... < t_m < \infty$ that

$$\tilde{X}_{t_m} - \tilde{X}_{t_{m-1}}, ..., \tilde{X}_{t_1} - \tilde{X}_{t_0} \quad \text{and} \quad \mathcal{F}_{\tau+} \quad \text{are independent.}$$

Let $G \in \mathcal{F}_{\tau+}$. We define a sequence of random times

$$\tau^{(n)} = \sum_{k=1}^{\infty} \frac{k}{2^n} \mathbb{1}_{\{\frac{k-1}{2^n} \leq \tau < \frac{k}{2^n}\}}.$$

We have that $\tau^{(n)} < \infty$. Then for $\theta_1, ..., \theta_m \in \mathbb{R}$, using tower property,

$$\begin{aligned}
& \mathbb{E} \exp\left\{ i \sum_{l=1}^{m} \theta_l (X_{\tau^{(n)}+t_l} - X_{\tau^{(n)}+t_{l-1}}) \right\} \mathbb{1}_G \\
= & \sum_{k=1}^{\infty} \mathbb{E} \exp\left\{ i \sum_{l=1}^{m} \theta_l (X_{\tau^{(n)}+t_l} - X_{\tau^{(n)}+t_{l-1}}) \right\} \mathbb{1}_{G \cap \{\tau^{(n)} = \frac{k}{2^n}\}} \\
= & \sum_{k=1}^{\infty} \mathbb{E} \exp\left\{ i \sum_{l=1}^{m} \theta_l (X_{\frac{k}{2^n}+t_l} - X_{\frac{k}{2^n}+t_{l-1}}) \right\} \mathbb{1}_{G \cap \{\tau^{(n)} = \frac{k}{2^n}\}} \\
= & \sum_{k=1}^{\infty} \mathbb{E} \mathbb{1}_{G \cap \{\tau^{(n)} = \frac{k}{2^n}\}} \mathbb{E}\left[\exp\left\{ i \sum_{l=1}^{m} \theta_l (X_{\frac{k}{2^n}+t_l} - X_{\frac{k}{2^n}+t_{l-1}}) \right\} \middle| \mathcal{F}_{\frac{k}{2^n}} \right] \\
= & \sum_{k=1}^{\infty} \mathbb{E} \mathbb{1}_{G \cap \{\tau^{(n)} = \frac{k}{2^n}\}} \mathbb{E} \exp\left\{ i \sum_{l=1}^{m} \theta_l (X_{\frac{k}{2^n}+t_l} - X_{\frac{k}{2^n}+t_{l-1}}) \right\} \\
= & \sum_{k=1}^{\infty} \mathbb{E} \mathbb{1}_{G \cap \{\tau^{(n)} = \frac{k}{2^n}\}} \mathbb{E} \exp\left\{ i \sum_{l=1}^{m} \theta_l (X_{t_l} - X_{t_{l-1}}) \right\} \\
= & \mathbb{P}(G) \mathbb{E} \exp\left\{ i \sum_{l=1}^{m} \theta_l (X_{t_l} - X_{t_{l-1}}) \right\}
\end{aligned}$$

since $G \cap \{\tau^{(n)} = \frac{k}{2^n}\} = G \cap \left\{\frac{k-1}{2^n} \leq \tau < \frac{k}{2^n}\right\} \in \mathcal{F}_{\frac{k}{2^n}}$. Because we have $\tau^{(n)}(\omega) \downarrow \tau(\omega)$ and X is right-continuous, we get

$$\lim_{n\to\infty} X_{\tau^{(n)}(\omega)+s} = X_{\tau(\omega)+s}$$

for all $s \geq 0$ and

$$\mathbb{E}\exp\left\{i\sum_{l=1}^{m}\theta_l(X_{\tau+t_l} - X_{\tau+t_{l-1}})\right\}\mathbb{1}_G = \mathbb{P}(G)\mathbb{E}\exp\left\{i\sum_{l=1}^{m}\theta_l(X_{t_l} - X_{t_{l-1}})\right\}$$

by dominated convergence. Specialising to $\Omega = G$ yields to

$$\mathbb{E}\exp\left\{i\sum_{l=1}^{m}\theta_l(X_{\tau+t_l} - X_{\tau+t_{l-1}})\right\} = \mathbb{E}\exp\left\{i\sum_{l=1}^{m}\theta_l(X_{t_l} - X_{t_{l-1}})\right\},$$

which implies that X and $\tilde{X}$ have the same finite-dimensional distributions. In turn, this also gives

$$\mathbb{E}\exp\left\{i\sum_{l=1}^{m}\theta_l(X_{\tau+t_l}-X_{\tau+t_{l-1}})\right\}\mathbb{1}_G = \mathbb{P}(G)\mathbb{E}\exp\left\{i\sum_{l=1}^{m}\theta_l(X_{\tau+t_l}-X_{\tau+t_{l-1}})\right\}.$$

which means that $\tilde{X}$ is independent from $\mathcal{F}^X_{\tau+}$. □

Theorem 6.7. *A* LÉVY *process is a strong* MARKOV *process.*

Proof. Assume that $\tau : \Omega \to [0,\infty)$ is an $\mathbb{F}^X$-optional time. Since by Lemma 6.2 we have that X_τ is $\mathcal{F}^X_{\tau+}$ measurable and from Theorem 6.6 we have that $X_{t+\tau} - X_\tau$ is independent from $\mathcal{F}^X_{\tau+}$, we get that for any $A \in \mathcal{E}$ it holds, a.s.,

$$\begin{aligned}\mathbb{P}(X_{\tau+t} \in A|\mathcal{F}_{\tau+}) &= \mathbb{E}[\mathbb{1}_{(X_{t+\tau}-X_\tau)+X_\tau\in A\}}|\mathcal{F}_{\tau+}] \\ &= (\mathbb{E}\mathbb{1}_{(X_{t+\tau}-X_\tau)+y\in A\}}) \,|_{y=X_\tau}\end{aligned}$$

The assertion from Theorem 6.6 that $X_{t+\tau} - X_\tau \stackrel{d}{=} X_t$ allows us to write

$$\mathbb{E}\mathbb{1}_{\{(X_{t+\tau}-X_\tau)+y\in A\}} = \mathbb{E}\mathbb{1}_{\{X_t+y\in A\}} = P_t(y,A).$$

Consequently, we have shown that

$$\mathbb{P}(X_{\tau+t} \in A|\mathcal{F}_{\tau+}) = P_t(X_\tau, A) \text{ a.s.}$$ □

7 The semi-group and infinitesimal generator approach

7.1 Contraction semi-groups

Definition 7.1 (semi-group)**.**

(1) Let $\mathcal{B}$ be a real Banach space with norm $\|\cdot\|$. A one-parameter family $\{T(t); t \geq 0\}$ of bounded linear operators $T(t) : \mathcal{B} \to \mathcal{B}$ is called a *semi-group* if

   (a) $T(0) = Id$,

   (b) $T(s+t) = T(s)T(t)$ for all $s, t \geq 0$.

(2) A semi-group $\{T(t); t \geq 0\}$ is called *strongly continuous* (or C_0 semi-group) if, for all $f \in \mathcal{B}$,
$$\lim_{t \downarrow 0} T(t)f = f.$$

(3) The semi-group $\{T(t); t \geq 0\}$ is a *contraction semi-group* if, for all $t \geq 0$,
$$\|T(t)\| = \sup_{\|f\|=1} \|T(t)f\| \leq 1.$$

Example 7.2. Let $\mathcal{B} := \mathbb{R}^d$ and let A be a $d \times d$ matrix. For $t \geq 0$ define

$$T(t) := e^{tA} := \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k$$

with A^0 being the identity matrix. As norm we take the operator norm of A, i.e.
$$\|A\| := \sup\{|Ax| : |x| \leq 1\},$$
where $|(x_1, \ldots, d)| := (x_1^2 + \cdots + x_d^2)^{\frac{1}{2}}$. Then one has that

(1) $e^{(s+t)A} = e^{sA}e^{tA}$ for all $s, t \geq 0$,

(2) $\{e^{tA}; t \geq 0\}$ is strongly continuous, and

(3) $\|e^{tA}\| \leq e^{t\|A\|}$ for $t \geq 0$.

Definition 7.3. Let E be a complete separable metric space and let $\mathcal{B}(E)$ be the BOREL-σ-algebra generated by the open sets of E. By $\mathcal{B}_E$ we denote the space of bounded measurable functions

$$f : (E, \mathcal{B}(E)) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

and equip this space with the norm $\|f\| := \sup_{x\in E} |f(x)|$.

Theorem 7.4. *Let E be a complete separable metric space and X be a homogeneous Markov process with transition function $\{P_t(x, A)\}$. Then the following is true:*

(1) *The space $\mathcal{B}_E$ defined in Definition 7.3 is a Banach space.*

(2) *The family of operators $\{T(t); t \geq 0\}$ with*

$$T(t)f(x) := \int_E f(y)P_t(x, dy), \quad f \in \mathcal{B}_E,$$

is a contraction semi-group.

Proof. (1) We realise that $\mathcal{B}_E$ is indeed a Banach space:

- Measurable and bounded functions form a vector space.
- $\|f\| := \sup_{x\in E} |f(x)|$ is a norm.
- $\mathcal{B}_E$ is complete w.r.t. this norm.

(2) We show that $T(t) : \mathcal{B}_E \to \mathcal{B}_E$: To verify that

$$T(t)f : (E, \mathcal{B}(E)) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

we can restrict ourself to $f \geq 0$ and find simple (measurable!) functions $f_n = \sum_{k=1}^{N_n} a_k^n \mathbb{1}_{A_k^n}$, $A_k^n \in \mathcal{B}(E), a_k^n \geq 0$ such that $f_n \uparrow f$. Then

$$\begin{aligned} T(t)f_n(x) &= \int_E \sum_{k=1}^{N_n} a_k^n \mathbb{1}_{A_k^n}(y)P_t(x, dy) \\ &= \sum_{k=1}^{N_n} a_k^n \int_E \mathbb{1}_{A_k^n}(y)P_t(x, dy) \end{aligned}$$

$$= \sum_{k=1}^{N_n} a_k^n P_t(x, A_k^n).$$

Since

$$P_t(\cdot, A_k^n) : (E, \mathcal{B}(E)) \to (\mathbb{R}, \mathcal{B}(\mathbb{R})),$$

we have this measurability for $T(t)f_n$, and by dominated convergence also for $T(t)f$. Moreover, we have

$$\begin{aligned} \|T(t)f\| &= \sup_{x \in E} |T(t)f(x)| \\ &\leq \sup_{x \in E} \int_E |f(y)| P_t(x, dy) \\ &\leq \sup_{x \in E} \|f\| P_t(x, E) = \|f\|. \end{aligned} \tag{7.1}$$

Hence $T(t)f \in \mathcal{B}_E$.

(c) $\{T(t); t \geq 0\}$ is a semi-group: We first observe that

$$T(0)f(x) = \int_E f(y) P_0(x, dy) = \int_E f(y) \delta_x(dy) = f(x)$$

which implies that $T(0) = Id$. From the Chapman-Kolmogorov equation we derive

$$\begin{aligned} T(s)T(t)f(x) &= T(s)(T(t)f)(x) \\ &= T(s)\left(\int_E f(y) P_t(\cdot, dy)\right)(x) \\ &= \int_E \int_E f(y) P_t(z, dy) P_s(x, dz) \\ &= \int_E f(y) P_{t+s}(x, dy) = T(t+s)f(x). \end{aligned}$$

(d) We have already seen in (7.1) that $\{T(t); t \geq 0\}$ is a contraction. $\square$

7.2 Infinitesimal generator

Definition 7.5 (infinitesimal generator)**.** Let $\{T(t); t \geq 0\}$ be a contraction semi-group on $\mathcal{B}_E$. Define $D(A)$ to be the set of all $f \in \mathcal{B}_E$ such that there exists a $g \in \mathcal{B}_E$ such that

$$\lim_{t \downarrow 0} \left\| \frac{T(t)f - f}{t} - g \right\| = 0 \tag{7.2}$$

and

$$A : D(A) \to \mathcal{B}_E \quad \text{by} \quad Af := \lim_{t \downarrow 0} \frac{T(t)f - f}{t}.$$

The operator A is called *infinitesimal generator* of $\{T(t); t \geq 0\}$ and $D(A)$ the *domain* of A.

Example 7.6. If $W = (W_t)_{t \geq 0}$ is the one-dimensional Brownian motion and

$$C_u^2(\mathbb{R}) := \{f : \mathbb{R} \to \mathbb{R} : \text{ twice continuously differentiable and } f'' \text{ is uniformly continuous and bounded}\},$$

then $C_u^2(\mathbb{R}) \subseteq D(A)$ and for $f \in C_u^2(\mathbb{R})$ we have that $Af = \frac{1}{2} \frac{d^2}{dx^2} f$.

Proof. We have $P_t(x, A) = \mathbb{P}(x + W_t \in A)$ and

$$T(t)f(x) = \mathbb{E} f(x + W_t).$$

By Itô's formula,

$$f(x + W_t) = f(x) + \int_0^t f'(x + W_s) dW_s + \frac{1}{2} \int_0^t f''(x + W_s) ds.$$

Since f' is bounded, we have $\mathbb{E} \int_0^t (f'(x + W_s))^2 ds < \infty$ and therefore

$$\mathbb{E} \int_0^t f'(x + W_s) dW_s = 0.$$

For $t > 0$ this implies

$$\begin{aligned} \lim_{t \downarrow 0} \frac{\mathbb{E} f(x + W_t) - f(x)}{t} &= \frac{1}{2} \lim_{t \downarrow 0} \mathbb{E} \frac{1}{t} \int_0^t f''(x + W_s) ds \\ &= \frac{1}{2} \mathbb{E} \lim_{t \downarrow 0} \frac{1}{t} \int_0^t f''(x + W_s) ds \end{aligned}$$

$$= \frac{1}{2} f''(x)$$

where we use dominated convergence as

$$\left| \frac{1}{t} \int_0^t f''(x + W_s) ds \right| \le \sup_y |f''(y)| < \infty$$

and the continuity of the paths $s \mapsto W_s(\omega)$. It remains to estimate uniformly in x the expression

$$\left| \frac{1}{2} \frac{1}{t} \mathbb{E} \int_0^t f''(x + W_s) ds - \frac{1}{2} f''(x) \right|.$$

Given $\varepsilon > 0$ we find an $\eta > 0$ such that $|x - y| < \eta$ implies that $|f''(x) - f''(y)| < \varepsilon$. Then

$$\begin{aligned}
&\left| \frac{1}{t} \mathbb{E} \int_0^t f''(x + W_s) ds - f''(x) \right| \\
&\le \left| \mathbb{E} 1_{\{\sup_{s \in [0,t]} |W_s| < \eta\}} \left[\frac{1}{t} \int_0^t f''(x + W_s) ds - f''(x) \right] \right| \\
&+ 2\mathbb{P}(\sup_{s \in [0,t]} |W_s| \ge \eta) \sup_x |f''(x)| \\
&\le \varepsilon + \frac{2}{\eta^2} \mathbb{E} \sup_{s \in [0,t]} |W_s|^2 \sup_x |f''(x)| \\
&\le \varepsilon + \frac{8}{\eta^2} \mathbb{E} |W_t|^2 \sup_x |f''(x)| \\
&\le \varepsilon + \frac{8t}{\eta^2} \sup_x |f''(x)|
\end{aligned}$$

where we applied Doob's maximal inequality. Therefore, given $\varepsilon > 0$, we can take $t_0 > 0$ small enough such that, for $t \in (0, t_0]$, we have

$$\varepsilon + \frac{4t}{\eta^2} \sup_x |f''(x)| \le 2\varepsilon.$$

□

Theorem 7.7. *Let $\{T(t); t \ge 0\}$ be a contraction semi-group and A its infinitesimal generator with domain $D(A)$. Then*

(1) *If* $f \in \mathcal{B}_E$ *is such that* $\lim_{t\downarrow 0} T(t)f = f$, *then for* $t \geq 0$ *it holds*

$$\int_0^t T(s)f ds \in D(A) \quad \text{and} \quad A\left(\int_0^t T(s)f ds\right) = T(t)f - f.$$

(2) *If* $f \in D(A)$ *and* $t \geq 0$, *then* $T(t)f \in D(A)$ *and*

$$\lim_{s\downarrow 0} \frac{T(t+s)f - T(t)f}{s} = AT(t)f = T(t)Af.$$

(3) *If* $f \in D(A)$ *and* $t \geq 0$, *then* $\int_0^t T(s)f ds \in D(A)$ *and*

$$T(t)f - f = A\int_0^t T(s)f ds = \int_0^t AT(s)f ds = \int_0^t T(s)Af ds.$$

Proof. (1) If $\lim_{t\downarrow 0} T(t)f = f$, then

$$\lim_{s\downarrow u} T(s)f = \lim_{t\downarrow 0} T(u+t)f = \lim_{t\downarrow 0} T(u)T(t)f = T(u)\lim_{t\downarrow 0} T(t)f = T(u)f,$$

where we used the continuity of $T(u) : \mathcal{B}_E \to \mathcal{B}_E$. This continuity from the right also implies that the Riemann integral

$$\int_0^t T(s+u)f du$$

exists for all $t, s \geq 0$ if we use in the discretizations the right-hand end point: for example if we set $t_i^n := \frac{ti}{n}$, then

$$\sum_{i=1}^n T(t_i^n)f(t_i^n - t_{i-1}^n) \to \int_0^t T(u)f du, \quad n \to \infty,$$

and therefore

$$\begin{aligned} T(s)\int_0^t T(u)f du &= T(s)\left(\int_0^t T(u)f du - \sum_{i=1}^n T(t_i^n)f(t_i^n - t_{i-1}^n)\right) \\ &\quad + \sum_{i=1}^n T(s)T(t_i^n)f(t_i^n - t_{i-1}^n) \end{aligned}$$

$$\to \int_0^t T(s+u) f du.$$

This implies

$$\begin{aligned}
\frac{T(s)-I}{s}\int_0^t T(u) f du &= \frac{1}{s}\left(\int_0^t T(s+u) f du - \int_0^t T(u) f du\right) \\
&= \frac{1}{s}\left(\int_s^{t+s} T(u) f du - \int_0^t T(u) f du\right) \\
&= \frac{1}{s}\left(\int_t^{t+s} T(u) f du - \int_0^s T(u) f du\right) \\
&\to T(t) f - f, \quad s \downarrow 0.
\end{aligned}$$

Since the RHS converges to $T(t)f - f \in \mathcal{B}_E$ we get $\int_0^t T(u) f du \in D(A)$ and

$$A \int_0^t T(u) f du = T(t) f - f.$$

(2) If $f \in D(A)$, then

$$\frac{T(s)T(t)f - T(t)f}{s} = \frac{T(t)(T(s)f - f)}{s} \to T(t)Af, \quad s \downarrow 0.$$

Hence $T(t)f \in D(A)$ and $AT(t)f = T(t)Af$.

(3) If $f \in D(A)$, then $\frac{T(s)f-f}{s} \to Af$ and therefore $T(s)f - f \to 0$ for $s \downarrow 0$. Then, by (1), we get $\int_0^t T(u) f du \in D(A)$. From (2) we get by integrating

$$\int_0^t \lim_{s\downarrow 0} \frac{T(s+u)f - T(u)f}{s} du = \int_0^t AT(u) f du = \int_0^t T(u) A f du.$$

On the other hand, in the proof of (1) we have shown that

$$\int_0^t \frac{T(s+u)f - T(u)f}{s} du = \frac{T(s)-I}{s} \int_0^t T(u) f du.$$

Since $\frac{T(s+u)f-T(u)f}{s}$ converges in $\mathcal{B}_E$ we may interchange limit and integral:

$$\begin{aligned}
\int_0^t \lim_{s\downarrow 0} \frac{T(s+u)f - T(u)f}{s} du &= \lim_{s\downarrow 0} \frac{T(s)-I}{s} \int_0^t T(u) f du \\
&= A \int_0^t T(u) f du.
\end{aligned}$$

□

7.3 Martingales and Dynkin's formula

Definition 7.8 (martingale)**.** An $\mathbb{F}$-adapted stochastic process $X = \{X_t; t \geq 0\}$ such that $\mathbb{E}|X_t| < \infty$ for all $t \geq 0$ is called $\mathbb{F}$-martingale (submartingale, supermartingale) if for all $0 \leq s \leq t < \infty$ it holds

$$\mathbb{E}[X_t|\mathcal{F}_s] = (\geq, \leq) X_s \quad a.s.$$

Theorem 7.9 (Dynkin's formula)**.** *Let X be a homogeneous Markov process with càdlàg paths for all $\omega \in \Omega$ and transition function $\{P_t(x, A)\}$. Let $\{T(t); t \geq 0\}$ denote its semi-group*

$$T(t)f(x) := \int_E f(y)P_t(x, dy) \quad for \quad f \in \mathcal{B}_E$$

and $(A, D(A))$ its generator. Then, for each $g \in D(A)$ the stochastic process $\{M_t; t \geq 0\}$ is an $\{\mathcal{F}_t^X; t \geq 0\}$ martingale, where

$$M_t := g(X_t) - g(X_0) - \int_0^t Ag(X_s)ds. \tag{7.3}$$

Remark 7.10. The integral $\int_0^t Ag(X_s)ds$ is understood as a Lebesgue-integral where for each $\omega \in \Omega$, i.e.

$$\int_0^t Ag(X_s)(\omega)ds := \int_0^t Ag(X_s)(\omega)\lambda(ds),$$

where λ denotes the Lebesgue measure.

Proof. Since by Definition 7.5 we have $A : D(A) \to \mathcal{B}_E$, it follows $Ag \in \mathcal{B}_E$, which means especially

$$Ag : (E, \mathcal{B}(E)) \to (\mathbb{R}, \mathcal{B}(\mathbb{R})).$$

Since X has càdlàg paths and is adapted, it is (see Lemma 6.2) progressively measurable, i.e.

$$X : ([0, t] \times \Omega, \mathcal{B}([0, t]) \otimes \mathcal{F}_t) \to (E, \mathcal{B}(E)).$$

Hence for the composition we have

$$Ag(X_\cdot) : ([0,t] \times \Omega, \mathcal{B}([0,t]) \otimes \mathcal{F}_t) \to (\mathbb{R}, \mathcal{B}(\mathbb{R})).$$

Moreover, Ag is bounded as it is from $\mathcal{B}_E$. So the integral

$$\int_0^t Ag(X_s(\omega))\lambda(ds),$$

w.r.t. the Lebesgue measure λ is well-defined for $\omega \in \Omega$. Fubini's theorem implies that M_t is $\mathcal{F}_t^X$ - measurable. Since g and Ag are bounded we have that $\mathbb{E}|M_t| < \infty$. From (7.3) we get, a.s.,

$$\begin{aligned}
& \mathbb{E}[M_{t+h}|\mathcal{F}_t^X] + g(X_0) \\
= \; & \mathbb{E}\left[g(X_{t+h}) - \int_0^{t+h} Ag(X_s)ds|\mathcal{F}_t^X\right] \\
= \; & \mathbb{E}\left[\left(g(X_{t+h}) - \int_t^{t+h} Ag(X_s)ds\right)\middle|\mathcal{F}_t^X\right] - \int_0^t Ag(X_s)ds.
\end{aligned}$$

The Markov property from Definition 3.1 (equation (3.1)) implies that

$$\mathbb{E}\left[g(X_{t+h})|\mathcal{F}_t^X\right] = \int_E g(y)P_h(X_t, dy).$$

We show next that $\mathbb{E}\left[\int_t^{t+h} Ag(X_s)ds\middle|\mathcal{F}_t^X\right] = \int_t^{t+h} \mathbb{E}[Ag(X_s)|\mathcal{F}_t^X]ds$, where we take as version for $\mathbb{E}[Ag(X_s)|\mathcal{F}_t^X]$ the expression $\int_E Ag(y)P_{s-t}(X_t, dy)$ which is possible due to the Markov property of X. Since $g \in D(A)$ we know that Ag is a bounded function so that we can use Fubini's theorem to show that for any $G \in \mathcal{F}_t^X$ it holds

$$\begin{aligned}
\int_\Omega \int_t^{t+h} Ag(X_s)ds \mathbb{1}_G d\mathbb{P} &= \int_t^{t+h} \int_\Omega Ag(X_s)\mathbb{1}_G d\mathbb{P}ds \\
&= \int_t^{t+h} \int_\Omega \int_E Ag(y)P_{s-t}(X_t, dy)\mathbb{1}_G d\mathbb{P}ds
\end{aligned}$$

so that

$$\mathbb{E}\left[\left(g(X_{t+h}) - \int_t^{t+h} Ag(X_s)ds\right)\middle|\mathcal{F}_t^X\right] - \int_0^t Ag(X_s)ds$$

$$= \int_E g(y)P_h(X_t,dy) - \int_t^{t+h}\int_E Ag(y)P_{s-t}(X_t,dy)ds \\ - \int_0^t Ag(X_s)ds.$$

The previous computations and relation $T(h)f(X_t) = \int_E f(y)P_h(X_t,dy)$ imply

$$\begin{aligned}
&\mathbb{E}[M_{t+h}|\mathcal{F}_t^X] + g(X_0) \\
&= \int_E g(y)P_h(X_t,dy) - \int_t^{t+h}\int_E Ag(y)dsP_{s-t}(X_t,dy)ds - \int_0^t Ag(X_s)ds \\
&= T(h)g(X_t) - \int_t^{t+h} T(s-t)Ag(X_t)ds - \int_0^t Ag(X_s)ds \\
&= T(h)g(X_t) - \int_0^h T(u)Ag(X_t)du - \int_0^t Ag(X_s)ds \\
&= T(h)g(X_t) - T(h)g(X_t) + g(X_t) - \int_0^t Ag(X_s)ds \\
&= g(X_t) - \int_0^t Ag(X_s)ds \\
&= M_t + g(X_0),
\end{aligned}$$

where we used Theorem 7.7(3). □

8 Weak solutions of SDEs and martingale problems

We recall the definition of a weak solution of an SDE.

Definition 8.1. Assume that $\sigma_{ij}, b_i : (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ are locally bounded. A *weak solution* of

$$dX_t = \sigma(X_t)dB_t + b(X_t)dt, \quad X_0 = x, \quad t \geq 0, \tag{8.1}$$

is a triplet $(X_t, B_t)_{t\geq 0}$, $(\Omega, \mathcal{F}, \mathbb{P})$, $(\mathcal{F}_t)_{t\geq 0}$, such that the following holds:

(1) $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$ satisfies the usual conditions:

- $(\Omega, \mathcal{F}, \mathbb{P})$ is complete.
- All null-sets of $\mathcal{F}$ belong to $\mathcal{F}_0$.
- The filtration is right-continuous.

(2) X is a d-dimensional continuous and $(\mathcal{F}_t)_{t\geq 0}$ adapted process.

(3) $(B_t)_{t\geq 0}$ is an m-dimensional $(\mathcal{F}_t)_{t\geq 0}$-Brownian motion.

(4) For $t \geq 0$ and $1 \leq i \leq d$ one has

$$X_t^{(i)} = x^{(i)} + \sum_{j=1}^{m} \int_0^t \sigma_{ij}(X_u)dB_u^{(j)} + \int_0^t b_i(X_u)du \text{ a.s.}$$

Let $a_{ij}(x) := \sum_{k=1}^m \sigma_{ik}(x)\sigma_{jk}(x)$, i.e. in the matrix notation $a(x) := \sigma(x)\sigma^T(x)$. Consider the differential operator

$$\begin{aligned} Af(x) &:= \frac{1}{2}\sum_{ij} a_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} f(x) + \sum_i b_i(x)\frac{\partial}{\partial x_i} f(x), \\ D(A) &:= C_c^2(\mathbb{R}^d), \end{aligned}$$

the twice continuously differentiable functions with compact support in $\mathbb{R}^d$. Then it follows from Itô's formula that

$$f(X_t) - f(X_0) - \int_0^t Af(X(s))ds = \int_0^t \nabla f(X_s)\sigma(X_s)dB_s \text{ a.s.}$$

is a martingale.

Definition 8.2 (canonical path-space)**.** (1) By $\Omega := C_{\mathbb{R}^d}([0,\infty))$ we denote the space of continuous functions $\omega : [0,\infty) \to \mathbb{R}^d$.

(2) For $\omega, \bar{\omega} \in \Omega$ we let

$$d(\omega, \bar{\omega}) := \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{\sup_{0 \le t \le n} |\omega(t) - \bar{\omega}(t)|}{1 + \sup_{0 \le t \le n} |\omega(t) - \bar{\omega}(t)|}.$$

(3) We set

$$\mathcal{F}_t^X := \sigma\{X_s, s \in [0,t]\} \quad \text{where} \quad X_s : C_{\mathbb{R}^d}([0,\infty)) \to \mathbb{R}^d : \omega \mapsto \omega(s)$$

is the coordinate mapping.

Remark 8.3. (1) $[C_{\mathbb{R}^d}([0,\infty)), d]$ is a complete separable metric space, see [5, Problem 2.4.1].

(2) For $0 \le t \le u$ we have $\mathcal{F}_t^X \subseteq \mathcal{F}_u^X \subseteq \mathcal{B}(C_{\mathbb{R}^d}([0,\infty)))$, see [5, Problem 2.4.2].

We define local martingales to introduce the concept of a martingale problem:

Definition 8.4 (local martingale)**.** For a stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t \ge 0})$ satisfying the usual conditions, a continuous $(\mathcal{F}_t)_{t \ge 0}$ adapted process $M = (M_t)_{t \ge 0}$ with $M_0 = 0$ is a **local martingale** if there exists a sequence of stopping times $\tau_n : \Omega \to [0,\infty]$ with $\tau_1 \le \tau_2 \le \tau_3 \le ... \uparrow \infty$ such that the stopped process M^{τ_n} given by $M_t^{\tau_n} := M_{\tau_n \wedge t}$ is a martingale for each $n \ge 1$.

Example 8.5 ([6])**.** Let $\alpha > 1$. Then the process which solves

$$X_t = 1 + \int_0^t X_s^\alpha dB_s$$

is a local martingale but not a martingale.

Definition 8.6 ($C_{\mathbb{R}^d}([0,\infty))$- martingale problem)**.** Given $(s,x) \in [0,\infty) \times \mathbb{R}^d$, a solution to the $C_{\mathbb{R}^d}([0,\infty))$- martingale problem for the operator A is a probability measure $\mathbb{P}$ on $(C_{\mathbb{R}^d}([0,\infty)), \overline{\mathcal{B}(C_{\mathbb{R}^d}([0,\infty)))}^{\mathbb{P}})$, where

$$\overline{\mathcal{B}(C_{\mathbb{R}^d}([0,\infty)))}^{\mathbb{P}}$$

is the $\mathbb{P}$-completion of $\mathcal{B}(C_{\mathbb{R}^d}([0,\infty)))$, satisfying

$$\mathbb{P}(\{\omega \in \Omega : \omega(t) = x, \quad 0 \leq t \leq s\}) = 1$$

such that for each $f \in C_c^\infty(\mathbb{R}^d)$ the process $\{M_t^f; t \geq s\}$ with

$$M_t^f := f(X_t) - f(X_s) - \int_s^t Af(X_u)du$$

is a $\mathbb{P}$-martingale with respect to $((\mathcal{F}_t^{X,\mathbb{P}})_+)_{t\geq s}$, where $(\mathcal{F}_t^{X,\mathbb{P}})_{t\geq 0}$ is the augmentation under $\mathbb{P}$ of $(\mathcal{F}_t^X)_{t\geq 0}$, and $(\mathcal{F}_t^{X,\mathbb{P}})_+ = \bigcap_{s>t} \mathcal{F}_s^{X,\mathbb{P}}$.

Theorem 8.7. *Given by a probability measure $\mathbb{P}$ on*

$$(C_{\mathbb{R}^d}([0,\infty)), \mathcal{B}(C_{\mathbb{R}^d}[0,\infty)))$$

the following assertions are equivalent:

(1) *$\mathbb{P}$ is a solution to the $C_{\mathbb{R}^d}([0,\infty))$- martingale problem for the operator $(A, D(A))$.*

(2) *There is an extension of the stochastic basis*

$$\left(C_{\mathbb{R}^d}([0,\infty)), \overline{\mathcal{B}(C_{\mathbb{R}^d}([0,\infty)))}^{\mathbb{P}}, \mathbb{P}, ((\mathcal{F}_t^{X,\mathbb{P}})_+)_{t\geq 0}\right)$$

such that the process $(X_t)_{t\geq 0}$ becomes a weak solution to (8.1).

Proof. (2) $\Rightarrow$ (1) follows from Itô's formula as explained above.

(1) $\Rightarrow$ (2) We will show this direction only for the case $d = m$, see [5, Proposition 5.4.6] for the general case. We assume that X is a solution of the $C_{\mathbb{R}^d}([0,\infty))$- martingale problem for the operator A.

(a) We observe that for any $i = 1, ..., d$ and $f(x) := x_i$ the process $\{M_t^i := M_t^f; t \geq 0\}$ is a continuous, local martingale. This can be seen as follows: We define the stopping times for $n > \max\{|x^{(1)}|, ..., |x^{(d)}|\}$ by

$$\tau_n := \inf\{t > 0 : \max\{|X_t^{(1)}|, ..., |X_t^{(d)}|\} = n\}.$$

Then we can find a function $g_n \in C_c^\infty(\mathbb{R}^d)$ such that

$$(M^i)^{\tau_n} = (M^{g_n})^{\tau_n}.$$

By assumption M^{g_n} is a continuous martingale and it follows from the optional sampling theorem that the stopped process $(M^{g_n})^{\tau_n}$ is also a continuous martingale. In particular we have

$$M_t^i = X_t^{(i)} - x^{(i)} - \int_0^t b_i(X_s)ds.$$

Since X is continuous and b locally bounded, it holds

$$\int_0^t |b_i(X_s(\omega))|ds < \infty \quad \text{for all} \quad \omega \in \Omega \text{ and } t \geq 0.$$

(b) Also for $f(x) := x_i x_j$ for fixed i, j the process $M_t^{(ij)} := M_t^f$, defined by

$$M_t^{ij} = X_t^{(i)} X_t^{(j)} - x^{(i)} x^{(j)} - \int_0^t X_s^{(i)} b_j(X_s) + X_s^{(j)} b_i(X_s) + a_{ij}(X_s)ds$$

is a continuous, local martingale by the same reasoning as in step (a). We notice that

$$M_t^i M_t^j - \int_0^t a_{ij}(X_s)ds = M_t^{ij} - x^{(i)} M_t^j - x^{(j)} M_t^i - R_t$$

where

$$\begin{aligned} R_t \;:=\; & \int_0^t (X_s^{(i)} - X_t^{(i)}) b_j(X_s)ds + \int_0^t (X_s^{(j)} - X_t^{(j)}) b_i(X_s)ds \\ & + \int_0^t b_i(X_s)ds \int_0^t b_j(X_s)ds. \end{aligned}$$

Indeed,

$$\begin{aligned} & M_t^i M_t^j - \int_0^t a_{ij}(X_s)ds \\ = \; & \left(X_t^{(i)} - x^{(i)} - \int_0^t b_i(X_s)ds \right) \left(X_t^{(j)} - x^{(j)} - \int_0^t b_j(X_s)ds \right) - \int_0^t a_{ij}(X_s)ds \\ = \; & X_t^{(i)} X_t^{(j)} - X_t^{(i)} \left(x^{(j)} + \int_0^t b_j(X_s)ds \right) - \left(x^{(i)} + \int_0^t b_i(X_s)ds \right) X_t^{(j)} \\ & + \left(x^{(j)} + \int_0^t b_j(X_s)ds \right) \left(x^{(i)} + \int_0^t b_i(X_s)ds \right) - \int_0^t a_{ij}(X_s)ds \end{aligned}$$

$$
\begin{aligned}
&= M_t^{ij} + \underbrace{x^{(i)}} x^{(j)} + \int_0^t X_s^{(i)} b_j(X_s) + X_s^{(j)} b_i(X_s) ds \\
&\quad -X_t^{(i)} x^{(j)} - X_t^{(j)} \underbrace{x^{(i)}} - \int_0^t X_t^{(i)} b_j(X_s) + X_t^{(j)} b_i(X_s) ds \\
&\quad +x^{(i)} x^{(j)} + x^{(j)} \int_0^t b_i(X_s) ds + \underbrace{x^{(i)}} \int_0^t b_j(X_s) ds + \int_0^t b_j(X_s) ds \int_0^t b_i(X_s) ds \\
&= M_t^{ij} + \int_0^t (X_s^{(i)} - X_t^{(i)}) b_j(X_s) + (X_s^{(j)} - X_t^{(j)}) b_i(X_s) ds \\
&\quad - \underbrace{x^{(i)}} \underbrace{\left(-x^{(j)} + X_t^{(j)} - \int_0^t b_j(X_s) ds \right)}_{M_t^j} \\
&\quad -x^{(j)} \left(-x^{(i)} + X_t^{(i)} - \int_0^t b_i(X_s) ds \right) + \int_0^t b_j(X_s) ds \int_0^t b_i(X_s) ds.
\end{aligned}
$$

Since $X_s^{(i)} - X_t^{(i)} = M_s^i - M_t^i + \int_s^t b_j(X_u) du$ it follows by Itô's formula that

$$
\begin{aligned}
R_t &= \int_0^t (X_s^{(i)} - X_t^{(i)}) b_j(X_s) ds + \int_0^t (X_s^{(j)} - X_t^{(j)}) b_i(X_s) ds \\
&\quad + \int_0^t b_i(X_s) ds \int_0^t b_j(X_s) ds \\
&= \int_0^t (M_s^i - M_t^i) b_j(X_s) ds + \int_0^t (M_s^j - M_t^j) b_i(X_s) ds \\
&= -\int_0^t \int_0^s b_j(X_u) du dM_s^i - \int_0^t \int_0^s b_i(X_u) du dM_s^j.
\end{aligned}
$$

Since R_t is a continuous, local martingale and a process of bounded variation at the same time, $R_t = 0$ a.s. for all t. Then

$$
M_t^i M_t^j - \int_0^t a_{ij}(X_s) ds
$$

is a continuous, local martingale, and

$$
\langle M^i, M^j \rangle_t = \int_0^t a_{ij}(X_s) ds.
$$

By the Martingale Representation Theorem A.3 we know that there exists an extension $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ of $(\Omega, \mathcal{F}, \mathbb{P})$ carrying a d-dimensional $(\tilde{\mathcal{F}}_t)$ Brownian motion $\tilde{B}$ such that $(\tilde{\mathcal{F}}_t)$ satisfies the usual conditions, and measurable, adapted processes $\xi^{i,j}$, $i, j = 1, ..., d$, with

$$\tilde{\mathbb{P}}\left(\int_0^t (\xi_s^{i,j})^2 ds < \infty\right) = 1$$

such that

$$M_t^i = \sum_{j=1}^d \int_0^t \xi_s^{i,j} d\tilde{B}_s^j.$$

We have now

$$X_t = x + \int_0^t b(X_s)ds + \int_0^t \xi_s d\tilde{B}_s.$$

It remains to show that there exists an d-dimensional $(\tilde{\mathcal{F}}_t)$ Brownian motion B on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ such that $\tilde{\mathbb{P}}$ a.s.

$$\int_0^t \xi_s d\tilde{B}_s = \int_0^t \sigma(X_s)dB_s, \quad t \in [0, \infty).$$

For this we will use the following lemma.

Lemma 8.8. *Let*

$$\mathcal{D} := \{(\xi, \sigma); \xi \text{ and } \sigma \text{ are } d \times d \text{ matrices with } \xi\xi^T = \sigma\sigma^T\}.$$

On $\mathcal{D}$ *there exists a Borel-measurable map* $\mathcal{R} : (\mathcal{D}, \mathcal{D} \cap \mathcal{B}(\mathbb{R}^{d^2}) \to (\mathbb{R}^{d^2}, \mathcal{B}(\mathbb{R}^{d^2}))$ *such that*

$$\sigma = \xi\mathcal{R}(\xi, \sigma), \quad \mathcal{R}(\xi, \sigma)\mathcal{R}^T(\xi, \sigma) = I; \quad (\xi, \sigma) \in \mathcal{D}.$$

We set

$$B_t = \int_0^t \mathcal{R}^T(\xi_s, \sigma(X_s))d\tilde{B}_s.$$

Then B is a continuous local martingale and

$$\langle B^{(i)}, B^{(i)}\rangle_t = \int_0^t \mathcal{R}(\xi_s, \sigma(X_s))\mathcal{R}^T(\xi_s, \sigma(X_s))ds = t\delta_{ij}.$$

Lévy's theorem (see [5, Theorem 3.3.16]) implies that B is a Brownian motion. □

Usig Theorem 8.7 one can derive the following statement, which is the second main result of this section:

Theorem 8.9 (Kolmogorov 1965, Stroock-Varadhan 1969)**.** *If* $b_i, \sigma_{ij} : \mathbb{R}^d \to \mathbb{R}$ *are continuoius and bounded and if* μ *is an initial distribution on* $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ *such that*

$$\int_{\mathbb{R}^d} |x|^p \mu(dx) < \infty \quad \textit{for some} \quad p \in (2, \infty),$$

then there is a weak solution to the SDE

$$X_t = X_0 + \int_0^t \sigma(X_s) dB_s + \int_0^t b(X_s) ds \quad \textit{with} \quad \mathrm{law}(X_0) = \mu.$$

9 Feller processes

9.1 Feller semi-groups, Feller transition functions and Feller processes

Definition 9.1.

(1) $C_0(\mathbb{R}^d) := \{f : \mathbb{R}^d \to \mathbb{R} : f \text{ continuous}, \lim_{|x|\to\infty} |f(x)| = 0\}$ is equipped with the norm $\|f\| = \|f\|_{C_0(\mathbb{R}^d)} := \sup_{x\in\mathbb{R}^d} |f(x)|$.

(2) $\{T(t); t \geq 0\}$ is a *Feller semi-group* if

  (a) $T(t) : C_0(\mathbb{R}^d) \to C_0(\mathbb{R}^d)$ is positive for all $t \geq 0$, i.e. $T(t)f(x) \geq 0$ $\forall x$ if $f : \mathbb{R}^d \to [0, \infty)$,

  (b) $\{T(t); t \geq 0\}$ is a strongly continuous contraction semi-group.

(3) A Feller semi-group is *conservative* if for all $x \in \mathbb{R}^d$ it holds

$$\sup_{f\in C_0(\mathbb{R}^d),\|f\|=1} |T(t)f(x)| = 1.$$

Remark 9.2.

(1) $[C_0(\mathbb{R}^d), \|\cdot\|_{C_0(\mathbb{R}^d)}]$ is a Banach space.

(2) The subspace $C_c(\mathbb{R}^d)$ of compactly supported functions is dense in $C_0(\mathbb{R}^d)$.

Definition 9.3. If E is a locally compact Hausdorff space, a Borel measure on $(E, \mathcal{B}(E))$ is a *Radon measure* provided that

(1) $\mu(K) < \infty$ for all compact sets K,

(2) $\mu(A) = \inf\{\mu(U) : U \supseteq A, U \text{ open}\}$ for all $A \in \mathcal{B}(E)$,

(3) $\mu(B) = \sup\{\mu(K) : K \subseteq B, K \text{ compact}\}$ for all open set B.

We recall the Riesz representation theorem (see, for example, [3, Theorem 7.2]): If E is a locally compact Hausdorff space, L a positive linear functional on $C_c(E) := \{F : E \to \mathbb{R} : \text{continuous function with compact support}\}$, then there exists a unique Radon measure μ on $(E, \mathcal{B}(E))$ such that

$$LF = \int_E F(y)\mu(dy).$$

We use this theorem to prove the following:

Theorem 9.4. *Let $\{T(t); t \geq 0\}$ be a conservative* Feller *semi-group on $C_0(\mathbb{R}^d)$. Then there exists a homogeneous transition function $\{P_t : t \geq 0\}$, $P_t : \mathbb{R}^d \times \mathcal{B}(\mathbb{R}^d) \to [0, 1]$, such that*

$$T(t)f(x) = \int_{\mathbb{R}^d} f(y)P_t(x, dy) \quad \textit{for all} \quad x \in \mathbb{R}^d \textit{ and } f \in C_0(\mathbb{R}^d).$$

Proof. By the Riesz representation theorem we get for each $x \in \mathbb{R}^d$ and each $t \geq 0$ a measure $P_t(x, \cdot)$ on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ such that

$$(T(t)f)(x) = \int_{\mathbb{R}^d} f(y)P_t(x, dy), \quad \forall f \in C_c(\mathbb{R}^d).$$

We need to show that this family of measures $\{P_t(x, \cdot); t \geq 0, x \in \mathbb{R}^d\}$ has all properties of a transition function.

(a) The map $A \mapsto P_t(x, A)$ is a probability measure: Since $\{P_t(x, \cdot)$ is a measure, we only need to check whether $P_t(x, \mathbb{R}^d) = 1$, which will be an exercise.

(b) For $A \in \mathcal{B}(\mathbb{R}^d)$ we have to show that

$$x \mapsto P_t(x, A) : (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)) \to (\mathbb{R}, \mathcal{B}(\mathbb{R})). \tag{9.1}$$

We let

$$\begin{aligned} \mathcal{H} :=& \{f : \mathbb{R}^d \to \mathbb{R} : \mathcal{B}(\mathbb{R}^d) \text{ measurable and bounded}, \\ & T(t)f \text{ is } \mathcal{B}(\mathbb{R}^d) \text{ measurable}\}, \\ \mathcal{A} :=& \{[a_1, b_1] \times ... \times [a_n, b_n]; -\infty \leq a_k \leq b_k \leq \infty\} \cup \emptyset. \end{aligned}$$

By definition we have that $\sigma(\mathcal{A}) = \mathcal{B}(\mathbb{R}^d)$. Now we check the assumptions (1), (2), and (3) of Theorem A.2.

– The assumption (2), that $\mathcal{H}$ is a linear space, is obvious.

– The assumption (3), that $\mathcal{H}$ is a monotone class, follows from monotone convergence.

– $\mathbb{1}_A \in \mathcal{H}$ for all $A \in \mathcal{A}$ is verified as follows:

First we assume that $-\infty < a_k \leq b_k < \infty$. In this case we approximate $\mathbb{1}_A$ by $f_n \in C_c(\mathbb{R}^d)$ as follows: let $f_n(x_1, ...x_n) := f_{n,1}(x_1)...f_{n,d}(x_d)$ with linear, continuous functions

$$f_{n,k}(x_k) := \begin{cases} 1 & a_k \leq x_k \leq b_k, \\ 0 & x \leq a_k - \frac{1}{n} \text{ or } x \geq b_k + \frac{1}{n}. \end{cases}$$

Then $f_n \downarrow \mathbb{1}_A$. Since $T(t)f_n \in C_0(\mathbb{R}^d)$ because $f_n \in C_c(\mathbb{R}^d) \subseteq C_0(\mathbb{R}^d)$, we get

$$T(t)f_n : (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)) \to (\mathbb{R}, \mathcal{B}(\mathbb{R})).$$

It holds

$$T(t)f_n(x) = \int_{\mathbb{R}^d} f_n(y)P_t(x, dy) \to P_t(x, A) \quad \text{for} \quad n \to \infty.$$

Hence $P_t(\cdot, A) : (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, which means $\mathbb{1}_A \in \mathcal{H}$. Furthermore, the case $a_k = -\infty$ and $b_k = \infty$ can be done by monotone convergence again. Applying Theorem A.2, we obtain that $\mathcal{H}$ contains all bounded and $\mathcal{B}(\mathbb{R}^d)$-measurable functions.

(c) The Chapman-Kolmogorov equation for $\{P_t : t \geq 0\}$ we conclude from $T(t+s) = T(t)T(s)$ for all $s, t \geq 0$, which can be again done by approximating $\mathbb{1}_A$, $A \in \mathcal{A}$ and using dominated convergence and the Monotone Class Theorem.

(d) $T(0) = Id$ gives that $P_0(x, \cdot)$ is the measure μ_0 such that

$$f(x) = (T(0)f)(x) = \int_{\mathbb{R}^d} f(y)P_0(x, dy).$$

But this implies that $P_0(x, A) = \delta_x(A)$, which will be an exercise. □

Definition 9.5.

(1) A transition function associated to a conservative Feller semi-group is called a *Feller transition function.*

(2) A MARKOV process having a FELLER transition function is called a *Feller process.*

In general we have the following implications:

Theorem 9.6.

(1) *Every càdlàg* FELLER *process is a strong* MARKOV *process.*

(2) *Every strong* MARKOV *process is a* MARKOV *process.*

Now we characterize FELLER transition functions:

Theorem 9.7. *A transition function* $\{P_t(x,A)\}$ *is* FELLER *if and only if*

(1) $\int_{\mathbb{R}^d} f(y)P_t(\cdot,dy) \in C_0(\mathbb{R}^d)$ *for* $f \in C_0(\mathbb{R}^d)$ *and all* $t \geq 0$,

(2) $\lim_{t\downarrow 0} \int_{\mathbb{R}^d} f(y)P_t(x,dy) = f(x)$ *for all* $f \in C_0(\mathbb{R}^d)$ *and* $x \in \mathbb{R}^d$.

Proof. $\Longrightarrow$ is easy to see so that we turn to $\Longleftarrow$ and will show that (1) and (2) imply that $\{T(t); t \geq 0\}$ with

$$T(t)f(x) = \int_{\mathbb{R}^d} f(y)P_t(x,dy)$$

is a FELLER semi-group.

(a) We know by Theorem 7.4 that $\{T(t);\, t \geq 0\}$ is a contraction semi-group. By (1) we have that $T(t) : C_0(\mathbb{R}^d) \to C_0(\mathbb{R}^d)$. And of course, any $T(t)$ is positive. So we only have to show that

$$\lim_{t\downarrow 0} \|T(t)f - f\| = 0 \quad \text{for all} \quad f \in C_0(\mathbb{R}^d).$$

which is the strong continuity.
Since by (1) we have that $T(t)f \in C_0(\mathbb{R}^d)$ we conclude by (2) that

$$\lim_{s\downarrow 0} T(t+s)f(x) = T(t)f(x) \quad \text{for all} \quad x \in \mathbb{R}^d.$$

Hence we have that

– $t \mapsto T(t)f(x)$ is right-continuous,

– $x \mapsto T(t)f(x)$ is continuous.

This implies (similarly to the proof of the fact that right-continuity and adaptedness implies progressive measurability) that

$$(t,x) \mapsto T(t)f(x) : ([0,\infty) \times \mathbb{R}^d, \mathcal{B}([0,\infty)) \otimes \mathcal{B}(\mathbb{R}^d)) \to (\mathbb{R}, \mathcal{B}(\mathbb{R})).$$

(b) By Fubini's theorem we have for any $p > 0$, that

$$x \mapsto \mathcal{R}_p f(x) := \int_0^\infty e^{-pt} T(t) f(x) dt : (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)) \to (\mathbb{R}, \mathcal{B}(\mathbb{R})),$$

where the map $f \mapsto \mathcal{R}_p f$ is called the resolvent of order p of $\{T(t);\, t \geq 0\}$. It holds

$$\lim_{p\to\infty} p\mathcal{R}_p f(x) = f(x).$$

Indeed, since $\{T(t); t \geq 0\}$ is a contraction semi-group, it holds $\|T(\frac{u}{p})f\| \leq \|f\|$ for $u \geq 0$. Hence we can use dominated convergence in the following expression, and it follows from (2) that

$$p\mathcal{R}_p f(x) = \int_0^\infty p e^{-pt} T(t) f(x) dt = \int_0^\infty e^{-u} T\left(\frac{u}{p}\right) f(x) du \to f(x) \quad (9.2)$$

for $p \to \infty$. Moreover, one can show that $\mathcal{R}_p f \in C_0(\mathbb{R}^d)$, so that

$$\mathcal{R}_p : C_0(\mathbb{R}^d) \to C_0(\mathbb{R}^d).$$

For $p, q > 0$ it holds

$$\begin{aligned}
(q-p)\mathcal{R}_p \mathcal{R}_q f &= (q-p)\mathcal{R}_p \int_0^\infty e^{-qt} T(t) f dt \\
&= (q-p) \int_0^\infty e^{-ps} T(s) \int_0^\infty e^{-qt} T(t) f dt ds \\
&= (q-p) \int_0^\infty e^{-(p-q)s} \int_0^\infty e^{-q(t+s)} T(t+s) f dt ds \\
&= (q-p) \int_0^\infty e^{-(p-q)s} \int_s^\infty e^{-qu} T(u) f du ds \\
&= (q-p) \int_0^\infty e^{-qu} T(u) f \int_0^u e^{-(p-q)s} ds du
\end{aligned}$$

$$
\begin{aligned}
&= (q-p)\int_0^\infty e^{-qu}T(u)f\frac{1}{q-p}(e^{-(p-q)u}-1)du \\
&= -\mathcal{R}_q f + \int_0^\infty e^{-pu}T(u)f du \\
&= \mathcal{R}_p f - \mathcal{R}_q f.
\end{aligned}
$$

This also implies that

$$(q-p)\mathcal{R}_p\mathcal{R}_q f = \mathcal{R}_p f - \mathcal{R}_q f = (q-p)\mathcal{R}_q\mathcal{R}_p f.$$

Now, let

$$\mathrm{Im}(\mathcal{R}_p) := \{\mathcal{R}_p f; f \in C_0(\mathbb{R}^d)\}.$$

If $g \in \mathrm{Im}(\mathcal{R}_p)$, then there exists $f \in C_0(\mathbb{R}^d)$ such that $g = \mathcal{R}_p f$ and we have

$$g = \mathcal{R}_p f = \mathcal{R}_q f + (q-p)\mathcal{R}_q\mathcal{R}_p f = \mathcal{R}_q(f + (q-p)\mathcal{R}_p f) \in \mathrm{Im}(\mathcal{R}_q).$$

Hence $\mathrm{Im}(\mathcal{R}_p) \subseteq \mathrm{Im}(\mathcal{R}_p)$ and by symmetry, $\mathrm{Im}(\mathcal{R}_p) = \mathrm{Im}(\mathcal{R}_p)$. Let $E := \mathrm{Im}(\mathcal{R}_p)$. By (9.2) we have

$$\|p\mathcal{R}_p f\| \leq \|f\|.$$

(c) We show that $E \subseteq C_0(\mathbb{R}^d)$ is dense. We follow [3, Section 7.3] and notice that $C_0(\mathbb{R}^d)$ is the closure of $C_c(\mathbb{R}^d)$ with respect to $\|f\| := \sup_{x\in\mathbb{R}^d}|f(x)|$.

Assume that $E \subseteq C_0(\mathbb{R}^d)$ is not dense. By the Hahn-Banach theorem there is linear and continuous functional $L : C_0(\mathbb{R}^d) \to \mathbb{R}$ such that $Lf = 0$ if $f \in E$ and positive for an $f_0 \in C_0(\mathbb{R}^d)$ which is outside the closure of E and given by

$$L(f) = \int_{\mathbb{R}^d} f(x)\mu(dx) \quad \text{for some signed measure} \quad \mu.$$

However, by dominated convergence we have

$$L(f_0) = \int_{\mathbb{R}^d} f_0(x)\mu(dx) = \lim_{p\to\infty}\int_{\mathbb{R}^d} p\mathcal{R}_p f_0(x)\mu(dx) = 0,$$

which is a contradiction so that D must be dense.

(d) Now we have

$$T(t)\mathcal{R}_p f(x) \;=\; T(t)\int_0^\infty e^{-pu}T(u)f(x)du$$

$$= e^{pt}\int_t^\infty e^{-ps}T(s)f(x)ds.$$

Now we fix $p=1$ and consider $f\in E$ so that $f=\mathcal{R}_1 g$ for some $g\in C_0(\mathbb{R}^d)$. This implies

$$\begin{aligned}
&\|T(t)\mathcal{R}_1 g-\mathcal{R}_1 g\| \\
&= \sup_{x\in\mathbb{R}^d}\left|e^t\int_t^\infty e^{-s}T(s)g(x)du-\int_0^\infty e^{-u}T(u)g(x)du\right| \\
&= \sup_{x\in\mathbb{R}^d}\left|(e^t-1)\int_t^\infty e^{-s}T(s)g(x)du-\int_0^t e^{-u}T(u)g(x)du\right| \\
&\le [(e^{pt}-1)+t]\left[\int_0^\infty e^{-s}ds\right]\|g\|\to 0, \quad t\downarrow 0.
\end{aligned}$$

So we have shown that $\{T(t);t\ge 0\}$ is strongly continuous on E. Since $D\subseteq C_0(\mathbb{R}^d)$ is dense, we have also show strong continuity on $C_0(\mathbb{R}^d)$. □

9.2 Càdlàg modifications of Feller processes

In Definition 6.5 we defined a Lévy process as a stochastic process with a.s. càdlàg paths. In Theorem 6.7 we have shown that a Lévy process (with càdlàg paths) is a strong Markov process. By the Daniell-Kolmogorov Theorem we know that Markov processes exist by Theorem 4.3. But this Theorem does not say anything about path properties.

We will proceed with the definition of a Lévy process in law (and leave it as an exercise to show that such a process is a Feller process). We will prove then that any Feller process has a càdlàg modification.

Definition 9.8 (Lévy process in law)**.** A stochastic process $X=\{X_t;t\ge 0\}$ on $(\Omega,\mathcal{F},\mathbb{P})$ with $X_t:(\Omega,\mathcal{F})\to(\mathbb{R}^d,\mathcal{B}(\mathbb{R}^d))$ is a Lévy process in law if

(1) X is continuous in probability, i.e. for all $t\ge 0$ and $\varepsilon>0$ one has

$$\lim_{s\downarrow t}\mathbb{P}(|X_s-X_t|>\varepsilon)=0,$$

(2) $\mathbb{P}(X_0=0)=1$,

(3) for all $0\le s\le t$ one has $X_t-X_s\overset{d}{=}X_{t-s}$,

(4) for all $0 \le s \le t$ one has $X_t - X_s$ is independent of $\mathcal{F}_s^X$.

Theorem 9.9. *A Lévy process in law is a* Feller *process.*

We shall prove this as an exercise.

Theorem 9.10. *Let X be an $\{\mathcal{F}_t; t \ge 0\}$-submartingale. Then the following holds:*

(1) *For any countable dense subset $D \subseteq [0, \infty)$ there is a $\Omega^* \in \mathcal{F}$ with $\mathbb{P}(\Omega^*) = 1$ such that for every $\omega \in \Omega^*$ one has*

$$X_{t+}(\omega) := \lim_{s \downarrow t, s \in D} X_s(\omega) \quad \text{and} \quad X_{t-}(\omega) := \lim_{s \uparrow t, s \in D} X_s(\omega)$$

exists for all $t \ge 0$ ($t > 0$, respectively).

(2) *$\{X_{t+}; t \ge 0\}$ is an $\{\mathcal{F}_{t+}; t \ge 0\}$ submartingale with a.s. càdlàg paths.*

(3) *Assume that $\{\mathcal{F}_t; t \ge 0\}$ satisfies the usual conditions. Then X has a càdlàg modification if and only if $t \mapsto \mathbb{E}X_t$ is right-continuous.*

The proof can be found in [5, Proposition 1.3.14 and Theorem 1.3.13].

Lemma 9.11. *Let X be a* Feller *process. For any $p > 0$ and any*

$$f \in C_0(\mathbb{R}^d; [0, \infty)) := \{f \in C_0(\mathbb{R}^d) : f \ge 0\}$$

the process

$$\{e^{-pt} R_p f(X_t); t \ge 0\}$$

is a supermartingale w.r.t. the natural filtration $\{\mathcal{F}_t^X; t \ge 0\}$ and for any initial distribution $\mathbb{P}_\nu(X_0 \in B) = \nu(B)$ for $B \in \mathcal{B}(\mathbb{R}^d)$.

Proof. Recall that for $p > 0$ we defined in the proof of Theorem 9.7 the resolvent

$$f \mapsto \mathcal{R}_p f := \int_0^\infty e^{-pt} T(t) f dt, \quad f \in C_0(\mathbb{R}^d).$$

(a) We show that $\mathcal{R}_p : C_0(\mathbb{R}^d) \to C_0(\mathbb{R}^d)$: Since

$$\|\mathcal{R}_p f\| = \left\| \int_0^\infty e^{-pt} T(t) f dt \right\| \le \int_0^\infty e^{-pt} \|T(t) f\| dt$$

and $\|T(t)f\| \leq \|f\|$, we may use dominated convergence, and since $T(t)f \in C_0(\mathbb{R}^d)$ it holds

$$\begin{aligned} \lim_{x_n \to x} \mathcal{R}_p f(x_n) &= \lim_{x_n \to x} \int_0^\infty e^{-pt} T(t) f(x_n) dt \\ &= \int_0^\infty e^{-pt} \lim_{x_n \to x} T(t) f(x_n) dt \\ &= \mathcal{R}_p f(x). \end{aligned}$$

In the same way we verify that $\lim_{|x_n| \to \infty} \mathcal{R}_p f(x_n) = 0$.

(b) For $x \in \mathbb{R}^d$, $f \in C_0(\mathbb{R}^d; [0, \infty))$, and $h > 0$ one has

$$\begin{aligned} e^{-ph} T(h) \mathcal{R}_p f(x) &= e^{-ph} T(h) \int_0^\infty e^{-pt} T(t) f(x) dt \\ &= \int_0^\infty e^{-p(t+h)} T(t+h) f(x) dt \\ &= \int_h^\infty e^{-pu} T(u) f(x) du \\ &\leq \int_0^\infty e^{-pu} T(u) f(x) du \\ &= \mathcal{R}_p f(x). \end{aligned}$$

(c) The process $\{e^{-pt} R_p f(X_t); t \geq 0\}$ is a supermartingale: Let $0 \leq s \leq t$. Since X is a FELLER process, it has a transition function, and by Definition 3.1 we may write

$$\begin{aligned} \mathbb{E}_{\mathbb{P}_\nu}[e^{-pt} \mathcal{R}_p f(X_t) | \mathcal{F}_s^X] &= e^{-pt} \int_{\mathbb{R}^d} \mathcal{R}_p f(y) P_{t-s}(X_s, dy) \\ &= e^{-pt} T(t-s) \mathcal{R}_p f(X_s). \end{aligned}$$

From step **(b)** we conclude

$$e^{-pt} T(t-s) \mathcal{R}_p f(X_s) \leq e^{-ps} \mathcal{R}_p f(X_s).$$

□

Lemma 9.12. *Let Y_1 and Y_2 be random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ with values in $\mathbb{R}^d$. Then the following holds:*

$$Y_1 = Y_2 \quad a.s. \quad \Longleftrightarrow \quad \begin{array}{l} \mathbb{E} f_1(Y_1) f_2(Y_2) = \mathbb{E} f_1(Y_1) f_2(Y_1) \\ \textit{for all } f_1, f_2 \in C_0(\mathbb{R}^d) \end{array}$$

Proof. The direction $\implies$ is evident. We will use the Monotone Class Theorem A.2 to verify $\impliedby$. Let

$$H := \{h : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R} : \quad h \text{ bounded and measurable, } \mathbb{E}h(Y_1, Y_2) = \mathbb{E}h(Y_1, Y_1)\}$$

As before we can approximate $\mathbb{1}_{[a_1,b_1]\times\ldots\times[a_{2d},b_{2d}]}$ for $-\infty < a_i \leq b_i < \infty$ by continuous functions with values in $[0, 1]$. Since by the Monotone Class Theorem the equality

$$\mathbb{E}h(Y_1, Y_2) = \mathbb{E}h(Y_1, Y_1)$$

holds for all $h : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ which are bounded and measurable, we choose $h(x, y) := \mathbb{1}_{\{(x,y)\in\mathbb{R}^d\times\mathbb{R}^d : x\neq y\}}$ and infer

$$\mathbb{P}(Y_1 \neq Y_2) = \mathbb{P}(Y_1 \neq Y_1) = 0. \quad \square$$

Theorem 9.13. *If X is a* Feller *process such that there is a dense set $D \subseteq [0, \infty)$ such that*

$$\mathbb{P}\left(\sup_{t\in[0,T]\cap D} |X_t| < \infty\right) = 1 \quad \textit{for all} \quad T > 0,$$

then it has a càdlàg modification.

Sketch of the proof. **(a)** *One-point compactification* (Alexandroff extension) of $\mathbb{R}^d$: Let ∂ be a point not in $\mathbb{R}^d$ and denote by $\mathcal{O}$ the open sets of $\mathbb{R}^d$. We define a topology $\mathcal{O}'$ on $(\mathbb{R}^d)^\partial := \mathbb{R}^d \cup \{\partial\}$ as

$$\mathcal{O}' := \{A \subset (\mathbb{R}^d)^\partial : \quad \text{either } A \in \mathcal{O} \text{ or } \partial \in A \textbf{ and } A^c \text{ is a compact subset of } \mathbb{R}^d\}.$$

Then $((\mathbb{R}^d)^\partial, \mathcal{O}')$ is a compact Hausdorff space. Any function $f \in C_0(\mathbb{R}^d)$ will be extended to $f \in C_0((\mathbb{R}^d)^\partial)$ by $f(\partial) := 0$.

(b) Let $(f_n)_{n=1}^\infty \subseteq C_0(\mathbb{R}^d; [0, \infty))$ be a sequence which separates the points, i.e. for any $x, y \in (\mathbb{R}^d)^\partial$ with $x \neq y$ there exists $n \in \mathbb{N}$ such that $f_n(x) \neq f_n(y)$, where we set $f_n(\partial) := 0$. Such a sequence exists, which we will not prove here. We want to show that then also

$$\mathcal{S} := \{\mathcal{R}_p f_n : p, n \in \mathbb{N}\}$$

is a countable set (which is clear) and separates the points: in fact, it holds for any $p > 0$ that

$$p\mathcal{R}_p f(x) = p\int_0^\infty e^{-pt} T(t) f(x) dt = \int_0^\infty e^{-u} T\left(\frac{u}{p}\right) f(x) du.$$

This implies

$$\begin{aligned} \sup_{x\in(\mathbb{R}^d)^\partial} |p\mathcal{R}_p f(x) - f(x)| &= \sup_{x\in(\mathbb{R}^d)^\partial} \left| \int_0^\infty e^{-u} \left(T\left(\frac{u}{p}\right) f \right)(x) - f(x)) du \right| \\ &\le \int_0^\infty e^{-u} \left\| T\left(\frac{u}{p}\right) f - f \right\| du \to 0, \quad p \to \infty, \end{aligned}$$

by dominated convergence since $\|T\left(\frac{u}{p}\right) f - f\| \le 2\|f\|$ and the strong continuity of the semi-group implies $\|T\left(\frac{u}{p}\right) f - f\| \to 0$ for $p \to \infty$. Then, if $x \neq y$ there exists a function f_n with $f_n(x) \neq f_n(y)$ and can find a $p \in \mathbb{N}$ such that $\mathcal{R}_p f_n(x) \neq \mathcal{R}_p f_n(y)$.

(c) We fix a set $D \subseteq [0, \infty)$ which is countable and dense. We show that there exists $\Omega^* \in \mathcal{F}$ with $\mathbb{P}(\Omega^*) = 1$ and such that for all $\omega \in \Omega^*$ and for all $n, p \in \mathbb{N}$ one has

$$[0, \infty) \ni t \mapsto \mathcal{R}_p f_n(X_t(\omega)) \tag{9.3}$$

has right and left (for $t > 0$) limits along D. From Lemma 9.11 we know that

$$\{e^{-pt}\mathcal{R}_p f_n(X_t); t \ge 0\} \text{ is an } \{\mathcal{F}_t^X; t \ge 0\} \text{ supermartingale.}$$

By Theorem 9.10 (1) we have for any $p, n \in \mathbb{N}$ a set $\Omega_{n,p}^* \in \mathcal{F}$ with $\mathbb{P}(\Omega_{n,p}^*) = 1$ such that for all $\omega \in \Omega_{n,p}^*$ and for all $t \ge 0$ ($t > 0$, respectively) the limits

$$\lim_{s\downarrow t, s\in D} e^{-ps}\mathcal{R}_p f_n(X_s(\omega)) \quad \left(\lim_{s\uparrow t, s\in D} e^{-ps}\mathcal{R}_p f_n(X_s(\omega)) \right)$$

exist. Since $s \mapsto e^{ps}$ is continuous we get assertion (9.3) by setting

$$\Omega^* := \bigcap_{n=1}^\infty \bigcap_{p=1}^\infty \Omega_{n,p}^*.$$

(d) We show that for all $\omega \in \Omega^*$ the map $t \to X_t(\omega)$ has right limits along D: If the limit $\lim_{s \downarrow t, s \in D} X_s(\omega)$ does not exist, then there are $x, y \in (\mathbb{R}^d)^\partial$ and sequences $(s_n)_n, (\bar{s}_m)_m \subseteq D$ with $s_n \downarrow t$, $\bar{s}_m \downarrow t$, such that

$$\lim_{n \to \infty} X_{s_n}(\omega) = x \quad \text{and} \quad \lim_{m \to \infty} X_{\bar{s}_m}(\omega) = y.$$

But there are $p, k \in \mathbb{N}$ such that $\mathcal{R}_p f_k(x) \neq \mathcal{R}_p f_k(y)$ which is a contradiction to the fact that $s \mapsto \mathcal{R}_p f_k(X_s(\omega))$ has right limits along D.

(e) Construction of a right-continuous modification: For $\omega \in \Omega^*$ we set for all $t \geq 0$

$$\tilde{X}_t(\omega) := \lim_{s \downarrow t, s \in D} X_s(\omega),$$

and for $\omega \notin \Omega^*$ we set $\tilde{X}_t(\omega) := x$, where $x \in \mathbb{R}^d$ is arbitrary and fixed. Then we have that

$$\tilde{X}_t = X_t \quad a.s.$$

where we argue as follows: Since for $f, g \in C_0(\mathbb{R}^d)$ we have

$$\begin{aligned}
\mathbb{E} f(X_t) g(\tilde{X}_t) &= \lim_{s \downarrow t, s \in D} \mathbb{E} f(X_t) g(X_s) \\
&= \lim_{s \downarrow t, s \in D} \mathbb{E}\mathbb{E}[f(X_t) g(X_s) | \mathcal{F}_t^X] \\
&= \lim_{s \downarrow t, s \in D} \mathbb{E} f(X_t) \mathbb{E}[g(X_s) | \mathcal{F}_t^X] \\
&= \lim_{s \downarrow t, s \in D} \mathbb{E} f(X_t) T(s-t) g(X_t) \\
&= \mathbb{E} f(X_t) g(X_t),
\end{aligned}$$

where we used the Markov property for the second last equation while the last equation follows from the fact that $\|T(s-t)h - h\| \to t$ for $s \downarrow 0$. By Lemma 9.12 we conclude $\tilde{X}_t = X_t$ a.s.

It is an exercise to verify that $t \to \tilde{X}_t$ is right-continuous for all $\omega \in \Omega$.

(f) Càdlàg modifications: We use [5, Theorem 1.3.8(v)] which states that almost every path of a right-continuous submartingale has left limits for any $t \in (0, \infty)$. Since $\{-e^{-pt} \mathcal{R}_p f_n(\tilde{X}_t); t \geq 0\}$ is a right-continuous submartingale, we can proceed as above (using the fact that $\mathcal{S}$ separates the points) so show that $t \mapsto \tilde{X}(\omega)$ is càdlàg for almost all $\omega \in \Omega$. $\square$

Remark 9.14. For a Lévy process in law it can be shown (see [4, Theorem II.2.68]) that the assumption

$$\mathbb{P}(\sup\{|X_t| : t \in [0,T] \cap D\} < \infty) = 1$$

is satisfied for all $T > 0$.

A Appendix

Lemma A.1 (Factorization Lemma)**.** *Assume $\Omega \neq \emptyset$, $(E, \mathcal{E})$ be a measurable space, maps $g : \Omega \to E$ and $F : \Omega \to \mathbb{R}$, and $\sigma(g) = \{g^{-1}(B) : B \in \mathcal{E}\}$. Then the following assertions are equivalent:*

(1) *The map F is $(\Omega, \sigma(g)) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is measurable.*

(2) *There exists a measurable $h : (E, \mathcal{E}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that $F = h \circ g$.*

For the proof see [1, p. 62].

Theorem A.2 (Monotone Class Theorem for functions)**.** *Let $\mathcal{A} \subseteq 2^\Omega$ be a π-system that contains Ω and assume $\mathcal{H} \subseteq \{f; f : \Omega \to \mathbb{R}\}$ such that*

(1) *$\mathbb{1}_A \in \mathcal{H}$ for $A \in \mathcal{A}$,*

(2) *$\mathcal{H}$ is a linear space,*

(3) *If $(f_n)_{n=1}^\infty \subseteq \mathcal{H}$ such that $0 \leq f_n \uparrow f$ and f is bounded, then $f \in \mathcal{H}$.*

Then $\mathcal{H}$ contains all bounded functions that are $\sigma(\mathcal{A})$ measurable.

For the proof see [4].

Theorem A.3. *Suppose a stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F})_{t\geq 0})$ satisfying the usual assumptions and continuous, local martingales $(M_t^1)_{t\geq 0}, \ldots, (M_t^d)_{t\geq 0}$. If for $1 \leq i, j \leq d$ and all $\omega \in \Omega$ the processes $\langle M^i, M^j\rangle_t(\omega)$ are absolutely continuous in t, then there exists an extension $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}, (\tilde{\mathcal{F}})_{t\geq 0}))$ of $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F})_{t\geq 0})$ satisfying the usual conditions and an d-dimensional $(\tilde{\mathcal{F}})_{t\geq 0}$-Brownian motion $(B_t)_{t\geq 0}$ and progressively measurable processes $(X_t^{i,j})_{t\geq 0}$ $i, j = 1, ..., d$ with*

$$\tilde{\mathbb{P}}\left(\int_0^t (X_s^{i,j})^2 ds < \infty\right) = 1, \quad 1 \leq i, j \leq d; 0 \leq t < \infty,$$

such that $\tilde{\mathbb{P}}$-a.s.

$$M_t^i = \sum_{j=1}^d \int_0^t X_s^{i,j} dB_s^j, \quad 1 \leq i \leq d; 0 \leq t < \infty,$$

$$\langle M^i, M^j\rangle_t = \sum_{k=1}^d \int_0^t X_s^{i,k} X_s^{k,j} ds \quad 1 \leq i, j \leq d; 0 \leq t < \infty.$$

For the proof see [5, Theorem 3.4.2].

A continuous adapted process is an Itô process provided that

$$X(t) = x + \int_0^t \mu(s)ds + \int_0^t \sigma(s)dB(s), \quad t \geq 0,$$

where μ and σ are progressively measurable and satisfy

$$\int_0^t \mu(s)ds < \infty, \quad \int_0^t \sigma(s)^2 ds < \infty \, a.s. \quad \text{for all} \quad t \geq 0.$$

Theorem A.4 (Itô's formula)**.** *If $B(t) = (B_1(t), ..., B_d(t))$ is a d-dimensional $(\mathcal{F}_t)$ Brownian motion and*

$$X_i(t) = x_i + \int_0^t \mu_i(s)ds + \sum_{j=1}^d \int_0^t \sigma_{ij}(s)dB_j(s),$$

are Itô processes, then for any C^2 function $f : \mathbb{R}^d \to \mathbb{R}$ we have

$$\begin{aligned} f(X_1(t), ..X_d(t)) &= f(x_1, .., x_d) + \sum_{i=1}^d \int_0^t \frac{\partial}{\partial x_i} f(X_1(s), ..X_d(s))dX_i(s) \\ &+ \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d \int_0^t \frac{\partial^2}{\partial x_i \partial x_j} f(X_1(s), ..X_d(s))d\langle X_i, X_j \rangle_s, \end{aligned}$$

and $d\langle X_i, X_j \rangle_s = \sum_{k=1}^d \sigma_{ik}\sigma_{jk}ds$.

References

[1] H. Bauer, *Measure and integration theory*, Walter de Gruyter, 2001.

[2] E. Cinlar, *Probability and Statistics*, Springer, 2011.

[3] G. Folland, *Real Analysis*, 1999.

[4] S. He, J. Wang, J. Yan, *Semimartingale Theory and Stochastic Calculus*, Taylor & Francis, 1992.

[5] I. Karatzas, S. Shreve, *Brownian Motion and Stochastic Calculus*, Springer, 1991.

[6] https://almostsure.wordpress.com/2010/08/16/failure-of-the-martingale-property/#more-816