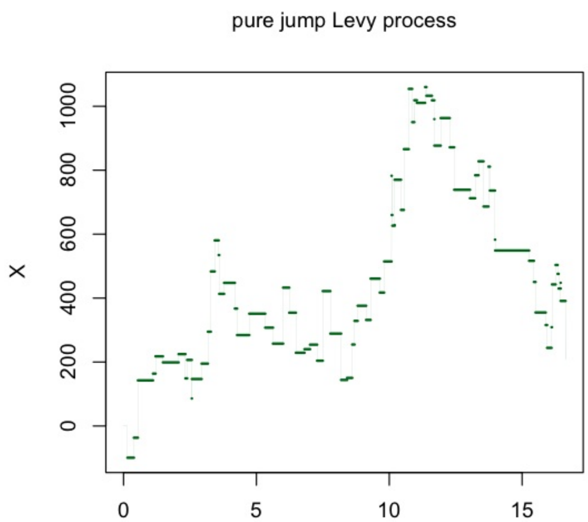
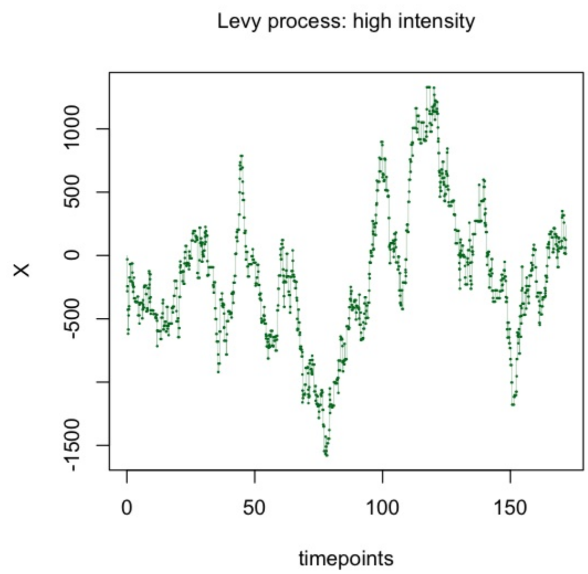
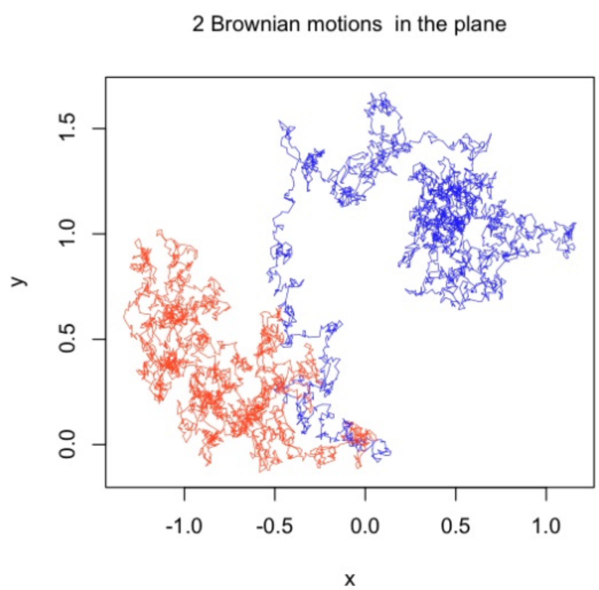
Markov processes







C. Geiss

December 10, 2019

Contents

1 Introduction

Why should one study Markov processes?

- Markov processes are quite general:

  A Brownian motion is a Lévy process.
  Lévy processes are Feller processes.
  Feller processes are Hunt processes, and the class of Markov processes comprises all of them.

- Solutions to certain SDEs are Markov processes.

- There exist many useful relations between Markov processes and

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

2 Definition of a Markov process

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space and (E, r) a complete separable metric space. By $(E, \mathcal{E})$ we denote a measurable space and $\mathbf{T} \subseteq \mathbb{R} \cup \{\infty\} \cup \{-\infty\}$.
We call $X = \{X_t; t \in \mathbf{T}\}$ a stochastic process if

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

The map $t \mapsto X_t(\omega)$ we call 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 $\mathbb{F}$ $\Longleftrightarrow_{df}$ X_t is $\mathcal{F}_t$ measurable for all $t \in \mathbf{T}$.

Obviously, X is always adapted to its natural filtration $\mathbb{F}^X = \{\mathcal{F}_t^X; t \in \mathbf{T}\}$ given by $\mathcal{F}_t^X = \sigma(X_s; s \leq t, s \in \mathbf{T})$.

Definition 2.1 (Markov process)**.** *The stochastic process X is a* Markov process *w.r.t.* $\mathbb{F} \iff_{df}$

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

(2) for all $t \in \mathbf{T} : \mathbb{P}(A \cap B|X_t) = \mathbb{P}(A|X_t)\mathbb{P}(B|X_t)$, a.s. whenever $A \in \mathcal{F}_t$ and $B \in \sigma(X_s; s \geq t)$. (for all $t \in \mathbf{T}$ 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) Recall that we define conditional probability using conditional expectation: $\mathbb{P}(C|X_t) := \mathbb{P}(C|\sigma(X_t)) = \mathbb{E}[\mathbb{1}_C|\sigma(X_t)]$.*

(2) If X is a Markov process w.r.t. $\mathbb{F}$, then X is a Markov process w.r.t. $\mathbb{G} = \{\mathcal{G}_t; s \in \mathbf{T}\}$, with $\mathcal{G}_t = \sigma(X_s; s \leq t, s \in \mathbf{T})$.

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

Theorem 2.3. *Let X be $\mathbb{F}$-adapted. TFAE:*

(i) X is a Markov process w.r.t. $\mathbb{F}$.

(ii) For each $t \in \mathbf{T}$ and each bounded $\sigma(X_s; s \geq t, s \in \mathbf{T})$-measurable Y one has

$$\mathbb{E}[Y|\mathcal{F}_t] = \mathbb{E}[Y|X_t]. \tag{1}$$

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

$$\mathbb{E}[f(X_s)|\mathcal{F}_t] = \mathbb{E}[f(X_s)|X_t] \tag{2}$$

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

Proof. (i) $\implies$ (ii):
Suppose (i) holds. The Monotone Class Theorem for functions (Theorem A.1) implies that it suffices to show (1) for $Y = \mathbb{1}_B$ where $B \in \sigma(X_s; s \geq t, s \in \mathbf{T})$. For $A \in \mathcal{F}_t$ we have

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

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

which implies (ii).
(ii) $\Longrightarrow$ (i):
Assume (ii) holds. If $A \in \mathcal{F}_t$ and $B \in \sigma(X_s; s \geq t, s \in \mathbf{T})$, then

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

which implies (i).
(ii) $\Longleftrightarrow$ (iii):
The implication (ii) $\Longrightarrow$ (iii) is trivial. Assume that (iii) holds. We want to use the Monotone Class Theorem for functions. Let

$$
\begin{aligned}
\mathcal{H} := \{Y; \quad & Y \text{ is bounded and } \sigma(X_s; s \geq t, s \in \mathbf{T}) - \text{measurable} \\
& \text{such that (1) holds.}\}
\end{aligned}
$$

Then $\mathcal{H}$ is a vector space containing the constants and is closed under bounded and monotone limits. We want that

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

It is enough to show that

$$
Y = \Pi_{i=1}^n f_i(X_{s_i}) \in \mathcal{H} \tag{3}
$$

for bounded $f_i : (E, \mathcal{E}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ and $t \leq s_1 < ... < s_n$ $(n \in \mathbb{N}^*)$.
(Notice that then especially $\mathbb{1}_A \in \mathcal{H}$ for any $A \in \mathcal{A}$ with

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

and $\sigma(\mathcal{A}) = \sigma(X_s; s \geq t, s \in \mathbf{T})$.
We show (3) by induction in n:

$\underline{n=1}$: This is assertion (iii).
$\underline{n>1}$:

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

By the factorization Lemma (Lemma A.2) there exists a $h : (E, \mathcal{E}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that $\mathbb{E}[f_n(X_{s_n})|X_{s_{n-1}}] = h(X_{s_{n-1}})$. By induction assumption:

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

By the tower property, since $\sigma(X_t) \subseteq \mathcal{F}_{s_{n-1}}$

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

□

Definition 2.4 (transition function)**.** *Let* $s, t \in \mathbf{T} \subseteq [0, \infty)$.

(1) The map

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

is called *Markov transition function* on $(E, \mathcal{E})$, provided that

(i) $A \mapsto P_{t,s}(x, A)$ is a probability measure on $(E, \mathcal{E})$ for each (t, s, x),
(ii) $x \mapsto P_{t,s}(x, A)$ is $\mathcal{E}$-measurable for each (t, s, A),
(iii) $P_{t,t}(x, A) = \delta_x(A)$
(iv) 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 $P_{t,s}(x,A)$ is homogeneous $\iff_{df}$ if there exists a map $\hat{P}_t(x,A)$ with $P_{t,s}(x,A) = \hat{P}_{s-t}(x,A)$ for all $0 \le t \le s, x \in E, A \in \mathcal{E}$.

(3) Let X be adapted to $\mathbb{F}$ and $P_{t,s}(x,A)$ with $0 \le t \le s$, $x \in E, A \in \mathcal{E}$ a Markov transition function. We say that X is a Markov process w.r.t. $\mathbb{F}$ having $P_{t,s}(x,A)$ as transition function if

$$\mathbb{E}[f(X_s)|\mathcal{F}_t] = \int_E f(y) P_{t,s}(X_t, dy) \tag{4}$$

for all $0 \le t \le s$ 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 2.5. *(1) There exist Markov processes which do not possess transition functions (see [4] Remark 1.11 page 446)*

(2) A Markov transition function for a Markov process is not necessarily unique.

Using the Markov property, one obtains the finite-dimensional distributions *of X:*
for $0 \le t_1 < t_2 < ... < t_n$ and bounded

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

it holds

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

3 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 3.1. For a measurable space $(E, \mathcal{E})$ and an arbitrary index set $\mathbf{T}$ define

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

where $X_t : \Omega \to E$ is the coordinate map $X_t(\omega) = \omega(t)$. For a finite subset $J = \{t_1, ..., t_n\} \subseteq \mathbf{T}$ we use the projections $\pi_J : \Omega \to E^J$

$$\begin{aligned} \pi_J \omega &= (\omega(t_1), ..., \omega(t_n)) \in E^J \\ \pi_J X &= (X_{t_1}, ..., X_{t_n}). \end{aligned}$$

(1) Let $\mathrm{Fin}(\mathbf{T}) := \{\mathrm{J} \subseteq \mathbf{T}; 0 < |\mathrm{J}| < \infty\}$. Then

$$\{\mathbf{P}_J : \mathbf{P}_J \text{ is a probability measure on } (E^J, \mathcal{E}^J), J \in \mathrm{Fin}(\mathbf{T})\}$$

is called the set of *finite-dimensional distributions* of X.

(2) The set of probability measures $\{\mathbf{P}_J : J \in \mathrm{Fin}(\mathbf{T})\}$ is called *Kolmogorov consistent* (or compatible or projective) provided that

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

for all $J \subseteq K,\ J, K \in \mathrm{Fin}(\mathbf{T})$.
(Here it is implicitly assumed that

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

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

Theorem 3.2 (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 set. Suppose that for each $J \in \mathrm{Fin}(\mathbf{T})$ there exists a probability measure P_J on $(E^J, \mathcal{E}^J)$ and that*

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

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

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

Proof: see, for example, Theorem 2.2 in Chapter 2 of [8].

Corollary 3.3 (Existence of Markov processes)**.** *Let $E = \mathbb{R}^d$, $\mathcal{E} = \mathcal{B}(\mathbb{R}^d)$ and $\mathbf{T} \subseteq [0, \infty)$. Assume μ is a probability measure on $(E, \mathcal{E})$, and*

$$\{P_{t,s}(x, A); t, s \in \mathbf{T},\ x \in E,\ A \in \mathcal{E}\}$$

is a family of Markov transition functions like in Definition 2.4. If $J = \{t_1, ..., t_n\} \subseteq \mathbf{T}$ *then by* $\{s_1, ..., s_n\} = J$ *with* $s_1 < ... < s_n$ *(i.e. the* $t_k's$ *are re-arranged according to their size) and define*

$$\mathbf{P}_J(A_1 \times ... \times A_n) := \int_E ... \int_E \mathbb{1}_{A_1\times...\times A_n}(x_1, .., x_n)\mu(dx_0)P_{0,s_1}(x_0, dx_1) ...P_{s_{n-1},s_n}(x_{n-1}, dx_n). \tag{5}$$

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.

Remark 3.4. *Using the monotone class theorem (Theorem A.1) one can show that (5) implies that for any bounded* $f : (E^n, \mathcal{E}^n) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ *it holds*

$$\mathbb{E}f(X_{s_1}, ..., X_{s_n}) = \int_E ... \int_E f(x_1, .., x_n)\mu(dx_0)P_{0,s_1}(x_0, dx_1) ...P_{s_{n-1},s_n}(x_{n-1}, dx_n). \tag{6}$$

Proof of the Corollary. 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: consider $K \subseteq J$,

$$K = \{s_{i_1} < ... < s_{i_k}\} \subseteq \{s_1 < ... < s_n\}, \quad k < n,$$

and

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

We have $\pi_K^{-1}(B_1 \times ... \times B_k) = A_1 \times ... \times A_n$ 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

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

$$= \mathbf{P}_{\{s_1,...,s_{n-2},s_n\}}(B_1 \times ... \times B_{n-2} \times B_n)$$

since, by Chapman-Kolmogorov, we have

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

According to Definition 2.1 we need to show that

$$\mathbb{P}(A \cap B|X_t) = \mathbb{P}(A|X_t)\mathbb{P}(B|X_t) \tag{7}$$

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

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

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

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

Indeed, by (5),

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

Using the tower property we get

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

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 (6) for $f(x_1) = \mathbb{1}_B(x_1)P_{t,s}(x_1, B_3)$. Again (6), 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

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

□

4 Strong Markov processes

4.1 Stopping times and optional times

For $(\Omega, \mathcal{F})$ we fix a filtration $\mathbb{F} = \{\mathcal{F}_t; t \in \mathbf{T}\}$, where $\mathbf{T} = [0, \infty) \cup \{\infty\}$ and $\mathcal{F}_\infty = \mathcal{F}$.

Definition 4.1. A map $\tau : \Omega \to \mathbf{T}$ is called a *stopping time w.r.t.* $\mathbb{F}$ provided that

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

Remark 4.2. *Note that* $\{\tau = \infty\} = \{\tau < \infty\}^c \in \mathcal{F}$ *since*

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

Then $\{\tau \le \infty\} = \{\tau < \infty\} \cup \{\tau = \infty\} \in \mathcal{F}_\infty$ *and hence*

$$\{\tau \le t\} \in \mathcal{F}_t \quad \textit{for all} \quad t \in \mathbf{T}.$$

We define

$$
\begin{aligned}
\mathcal{F}_{t+} &:= \bigcap_{s>t} \mathcal{F}_s, \quad t \in [0, \infty), \quad \mathcal{F}_{\infty+} := \mathcal{F}, \\
\mathcal{F}_{t-} &:= \sigma\left(\bigcup_{0\le s<t} \mathcal{F}_s\right), \quad t \in (0, \infty), \\
\mathcal{F}_{0-} &:= \mathcal{F}_0, \quad \mathcal{F}_{\infty-} := \mathcal{F}.
\end{aligned}
$$

Clearly,

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

Definition 4.3. 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 4.4. *If τ and σ are stopping times w.r.t. $\mathbb{F}$, then*

(1) $\tau + \sigma$,

(2) $\tau \wedge \sigma$, *(min)*

(3) $\tau \vee \sigma$, *(max)*

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

Definition 4.5. *Let τ be a stopping time w.r.t. $\mathbb{F}$. We define*

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

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

Note: $\mathcal{F}_\tau$ and $\mathcal{F}_{\tau+}$ are σ algebras.

Lemma 4.6. *Let $\sigma, \tau, \tau_1, \tau_2, ...$ be $\mathbb{F}$ - stopping times. Then it holds*

(i) τ *is* $\mathcal{F}_\tau$*-measurable,*

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

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

(iv) $\sup_n \tau_n$ *is an* $\mathbb{F}$ *- stopping time.*

Definition 4.7. *The map $\tau : \Omega \to \mathbf{T}$ is called optional time* $\Longleftrightarrow_{df}$

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

Note: For an optional time it holds

$$\tau : (\Omega, \mathcal{F}) \to (\mathbb{R} \cup \{\infty\}, \sigma(\mathcal{B}(\mathbb{R}) \cup \{\{\infty\}\}))$$

i.e. τ is an extended random variable.

Lemma 4.8.

(i) For $t_0 \in \mathbf{T}$ *the map* $\tau(\omega) = t_0 \quad \forall \omega \in \Omega$ *is a stopping time.*

(ii) Every stopping time is an optional time.

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

(iv) τ is an $\{\mathcal{F}_t; t \in \mathbf{T}\}$ optional time $\iff$ τ is an $\{\mathcal{F}_{t+}; t \in \mathbf{T}\}$ stopping time.

Proof. (i): Consider

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

(ii): Let τ be a stopping time. Then

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

(iii): We have that $\{\tau \leq t\} = \bigcap_{n=1}^{\infty} \underbrace{\left\{\tau < t + \frac{1}{n}\right\}}_{\in \mathcal{F}_{t+\frac{1}{n}}}$. 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 $\{\tau \leq t\} \in \mathcal{F}_{t+\frac{1}{M}} \quad \forall M \in \mathbb{N}^*$ and hence $\{\tau \leq t\} \in \mathcal{F}_{t+} = \mathcal{F}_t$ since $\{\mathcal{F}_t; t \in \mathbf{T}\}$ is right-continuous.
(iv) '$\implies$': 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 (exercise), we conclude from (iii) 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

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

$\square$

Lemma 4.9. *If τ is an optional time w.r.t. $\mathbb{F}$, then*

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

is a σ -algebra. It holds

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

4.2 Strong Markov property

Definition 4.10 (progressively measurable)**.** *Let E be a complete, separable metric space and $\mathcal{E} = \mathcal{B}(E)$. 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*

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

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

$$t \mapsto X_t(\omega)$$

is a right-continuous (left-continuous) function.

Lemma 4.11.

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

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

(iii) If τ is an $\mathbb{F}$-stopping time, and X is $\mathbb{F}$- progressively measurable, then X_τ (defined on $\{\tau < \infty\}$) is $\mathcal{F}_\tau$-measurable,

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

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

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

(v) If τ is an $\mathbb{F}$-optional time, and X is $\mathbb{F}$- progressively measurable, then X_τ (defined on $\{\tau < \infty\}$) is $\mathcal{F}_{\tau+}$-measurable.

Proof. (i), (ii) and (v) are exercises.
(iii): For $s \in [0,\infty)$ it holds

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

Hence $\tau \wedge t$ is $\mathcal{F}_t$-measurable. We have $h : \omega \mapsto (\tau(\omega) \wedge t, \omega)$:

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

Since X is $\mathbb{F}$- progressively measurable, we have

$$X : ([0,t] \times \Omega, \mathcal{B}([0,t]) \otimes \mathcal{F}_t) \to (E, \mathcal{E}). \tag{8}$$

Hence

$$X \circ h : (\Omega, \mathcal{F}_t) \to (E, \mathcal{E}). \tag{9}$$

It holds that X_τ is $\mathcal{F}_\tau$-measurable $\iff$

$$\{X_\tau \in B\} \cap \{\tau \le t\} \in \mathcal{F}_t \quad \forall t \in [0,\infty).$$

Indeed, this is true:

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

which is in $\mathcal{F}_t$ because of (9), and since τ is a stopping time.
(iv):
It holds

$$\begin{aligned} &H : (s,\omega) \mapsto (\tau(\omega) \wedge s, \omega) : \\ &([0,t] \times \Omega, \mathcal{B}([0,t]) \otimes \mathcal{F}_t) \to ([0,t] \times \Omega, \mathcal{B}([0,t]) \otimes \mathcal{F}_t), \quad t \ge 0, \end{aligned} \tag{10}$$

since

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

Because of (8) we have for the composition

$$\begin{aligned} &X \circ H : ([0,t] \times \Omega, \mathcal{B}([0,t]) \otimes \mathcal{F}_t) \to (E, \mathcal{E}), \\ &(X \circ H)(s,\omega) = X_{\tau(\omega) \wedge s}(\omega) = X^\tau_s(\omega). \end{aligned}$$

□

Definition 4.12 (strong Markov)**.** Assume X is an $\mathbb{F}$-progressively measurable homogeneous Markov process. Let $\{P_t(x,A)\}$ be its transition function. Then X is strong Markov if

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

for all $t \ge 0$, $A \in \mathcal{E}$ and all $\mathbb{F}$-optional times τ for which it holds $\tau < \infty$ a.s.

One can formulate the strong Markov property without transition functions:

Proposition 4.13. *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 4.12:*

(1) X is called strong Markov provided that for all $A \in \mathcal{E}$

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

for all $\mathbb{F}$-optional times τ such that $\tau < \infty$ a.s.

(2) $\forall t_1, ..., t_n \in \mathbf{T},\ A_1, ..., A_n \in \mathcal{E}$

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

for all $\mathbb{F}$-optional times τ such that $\tau < \infty$ a.s.

4.3 Lévy processes are strong Markov

Definition 4.14. A process X is called Lévy process if

(i) The paths of X are a.s. càdlàg (i.e. they are right-continuous and have left limits for $t > 0$.),

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

(iii) $\forall\, 0 \le s \le t: \quad X_t - X_s \stackrel{d}{=} X_{t-s}$,

(iv) $\forall\, 0 \le s \le t: \quad X_t - X_s$ is independent of $\mathcal{F}_s^X$.

The strong Markov property for a Lévy process is formulated as follows.

Theorem 4.15 (strong Markov property for a Lévy process)**.** *Let X be a Lévy process. Assume that τ is an $\mathbb{F}^X$-optional time such that $\tau < \infty$ almost surely. Define the process $\tilde{X} = \{\tilde{X}_t; t \ge 0\}$ by*

$$\tilde{X}_t = \mathbb{1}_{\{\tau<\infty\}}(X_{t+\tau} - X_\tau), \quad t \ge 0.$$

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

Remark 4.16. *To show that Theorem 4.15 implies that X is strong Markov according to Defintion 4.12 we proceed as follows. Assume that τ is an $\mathbb{F}^X$-optional time such that $\tau < \infty$ a.s. Since by Lemma 4.11 (v) we have that $X_\tau \mathbb{1}_{\{\tau<\infty\}}$ is $\mathcal{F}^X_{\tau+}$ measurable, and from the above Theorem we have that $\mathbb{1}_{\{\tau<\infty\}}(X_{t+\tau} - X_\tau)$ is independent from $\mathcal{F}^X_{\tau+}$, we get from the Factorization Lemma (Lemma A.2) that for any $A \in \mathcal{E}$ it holds*

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

The assertion from the theorem that $\mathbb{1}_{\{\tau<\infty\}}(X_{t+\tau} - X_\tau) \overset{d}{=} X_t$ allows us to write

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

Consequently, we have shown that on $\{\tau < \infty\}$,

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

Proof of Theorem 4.15. The finite dimensional distributions determine the law of a stochastic process. Hence it is sufficient to show for arbitrary $0 = t_0 < t_1 < ... < t_m$ ($m \in \mathbb{N}^*$) 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

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

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

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

$$= \sum_{k=1}^{\infty} \mathbb{E}\mathbb{1}_{G\cap\{\tau^{(n)}=\frac{k}{2^n}\}} \mathbb{E}\Big[\exp\Big\{ i\sum_{l=1}^{m} \theta_l (X_{\frac{k}{2^n}+t_l} - X_{\frac{k}{2^n}+t_{l-1}}) \Big\} \Big| \mathcal{F}_{\frac{k}{2^n}} \Big]$$

$$= \sum_{k=1}^{\infty} \mathbb{E}\mathbb{1}_{G\cap\{\tau^{(n)}=\frac{k}{2^n}\}} \mathbb{E}\exp\Big\{ i\sum_{l=1}^{m} \theta_l (X_{\frac{k}{2^n}+t_l} - X_{\frac{k}{2^n}+t_{l-1}}) \Big\}, \qquad (11)$$

since $G \cap \{\tau^{(n)} = \frac{k}{2^n}\} \in \mathcal{F}_{\frac{k}{2^n}}$ and property (iv) of Definition 4.14.
For $\omega \in \{\tau < \infty\}$ we have $\tau^{(n)}(\omega) \downarrow \tau(\omega)$. Since X is right-continuous:

$$X_{\tau^{(n)}(\omega)+s} \to X_{\tau(\omega)+s}, \quad n \to \infty, \forall s \geq 0.$$

By dominated convergence and property (iii) of Definition 4.14:

$$\mathbb{E}\exp\Big\{ i\sum_{l=1}^{m} \theta_l (X_{\tau+t_l} - X_{\tau+t_{l-1}}) \Big\} \mathbb{1}_{G\cap\{\tau<\infty\}}$$

$$= \lim_{n\to\infty} \mathbb{E}\exp\Big\{ i\sum_{l=1}^{m} \theta_l (X_{\tau^{(n)}+t_l} - X_{\tau^{(n)}+t_{l-1}}) \Big\} \mathbb{1}_G$$

$$= \lim_{n\to\infty} \mathbb{P}(G)\mathbb{E}\exp\Big\{ i\sum_{l=1}^{m} \theta_l (X_{t_l} - X_{t_{l-1}}) \Big\}$$

$$= \mathbb{P}(G)\mathbb{E}\exp\Big\{ i\sum_{l=1}^{m} \theta_l (X_{t_l} - X_{t_{l-1}}) \Big\},$$

where we used (11).

□

4.4 Right-continuous filtrations

We denote as above by $(\Omega, \mathcal{F})$ a measurable space and use $\mathbf{T} = [0, \infty) \cup \{\infty\}, \mathbb{F} = \{\mathcal{F}_t; t \in \mathbf{T}\}, \mathcal{F}_\infty = \mathcal{F}$.

Definition 4.17. The system $\mathcal{D} \subseteq 2^\Omega$ is called *Dynkin system* if $\iff_{df}$

(i) $\Omega \in \mathcal{D}$,

(ii) $A, B \in \mathcal{D}$ and $B \subseteq A \implies A \setminus B \in \mathcal{D}$,

(iii) $(A_n)_{n=1}^\infty \subseteq \mathcal{D}, \quad A_1 \subseteq A_2 \subseteq ... \implies \bigcup_{n=1}^\infty A_n \in \mathcal{D}$.

Theorem 4.18 (Dynkin system theorem)**.** *Let $\mathcal{C} \subseteq 2^\Omega$ be a π-system. If $\mathcal{D}$ is a Dynkin system and $\mathcal{C} \subseteq \mathcal{D}$, then*

$$\sigma(\mathcal{C}) \subseteq \mathcal{D}.$$

Definition 4.19 (augmented natural filtration)**.** Let X be a process on $(\Omega, \mathcal{F}, \mathbb{P})$. We set

$$\mathcal{N}^{\mathbb{P}} := \{A \subseteq \Omega : \exists B \in \mathcal{F} \text{ with } A \subseteq B \text{ and } \mathbb{P}(B) = 0\},$$

the set of '$\mathbb{P}$-null-sets'. If $\mathcal{F}_t^X = \sigma(X_u : u \leq t)$, then the filtration $\{\mathcal{F}_t^{\mathbb{P}}; t \in \mathbf{T}\}$ given by

$$\mathcal{F}_t^{\mathbb{P}} := \sigma(\mathcal{F}_t^X \cup \mathcal{N}^{\mathbb{P}})$$

is called the *augmented natural filtration* of X.

Theorem 4.20 (the augmented natural filtration of a strong Markov process is right-continuous)**.** *Assume $(E, \mathcal{E}) = (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ and let X be a strong Markov process with initial distribution μ (which means $\mathbb{P}(X_0 \in B) = \mu(B) \quad \forall B \in \mathcal{B}(\mathbb{R}^d)$). Then the augmented natural filtration*

$$\{\mathcal{F}_t^{\mathbb{P}}; t \in \mathbf{T}\}$$

is right-continuous.

Proof. **Step 1**
We show that $\forall s \geq 0$ and $G \in \mathcal{F}_\infty^X : \mathbb{P}(G|\mathcal{F}_{s+}^X) = \mathbb{P}(G|\mathcal{F}_s^X) \quad \mathbb{P}$-a.s:
Fix $s \in [0, \infty)$. Then $\sigma \equiv s$ is a stopping time w.r.t. $\mathbb{F}^X := \{\mathcal{F}_t^X; t \in \mathbf{T}\}$ by Lemma 4.8 (i) and (ii) we get that σ is an $\mathbb{F}^X$ optional time. For arbitrary $0 \leq t_0 < t_1 < ... < t_n \leq s < t_{n+1} < ... < t_m$ and $A_0, A_1, ..., A_m \in \mathcal{B}(\mathbb{R}^d)$ we have from Proposition 4.13 about the strong Markov property that

$$\begin{aligned}
& \mathbb{P}(X_{t_0} \in A_0, ..., X_{t_m} \in A_m|\mathcal{F}_{s+}^X) \\
= \; & \mathbb{E}[\mathbb{1}_{\{X_{t_0} \in A_0, ..., X_{t_n} \in A_n\}} \mathbb{1}_{\{X_{t_{n+1}} \in A_{n+1}, ..., X_{t_m} \in A_m\}}|\mathcal{F}_{s+}^X] \\
= \; & \mathbb{1}_{\{X_{t_0} \in A_0, ..., X_{t_n} \in A_n\}} \mathbb{P}(X_{t_{n+1}} \in A_{n+1}, ..., X_{t_m} \in A_m|\mathcal{F}_{s+}^X) \\
= \; & \mathbb{1}_{\{X_{t_0} \in A_0, ..., X_{t_n} \in A_n\}} \mathbb{P}(X_{t_{n+1}} \in A_{n+1}, ..., X_{t_m} \in A_m|X_s) \quad a.s.
\end{aligned}$$

Hence the RHS is a.s. $\mathcal{F}_s^X$-measurable. Define

$$\mathcal{D} := \{G \in \mathcal{F}_\infty^X : \mathbb{P}(G|\mathcal{F}_{s+}^X) \text{ has an } \mathcal{F}_s^X \text{ measurable version } \}.$$

Then $\Omega \in \mathcal{D}$. If $G_1, G_2 \in \mathcal{D}$ and $G_1 \subseteq G_2$, then

$$\mathbb{P}(G_2 \setminus G_1 | \mathcal{F}^X_{s+}) = \mathbb{P}(G_2|\mathcal{F}^X_{s+}) - \mathbb{P}(G_1|\mathcal{F}^X_{s+})$$

has an $\mathcal{F}^X_s$-measurable version. Finally, for $G_1, G_2, ... \in \mathcal{D}$ with $G_1 \subseteq G_2 \subseteq ...$ we get by monotone convergence applied to $\mathbb{E}[\mathbb{1}_{G_k}|\mathcal{F}^X_{s+}]$ that $\bigcup_{k=1}^\infty G_k \in \mathcal{D}$. We know that

$$\mathcal{C} := \{\{X_{t_0} \in A_0, ..., X_{t_m} \in A_m\} : 0 \leq t_0 < t_1 < ... < t_n \leq s < t_{n+1} < ... < t_m, A_k \in \mathcal{B}(\mathbb{R}^d)\}$$

is a π -system which generates $\mathcal{F}^X_\infty$. By the Dynkin system theorem we get that for any $G \in \mathcal{F}^X_\infty$

$$\mathbb{P}(G|\mathcal{F}^X_{s+})$$

has an $\mathcal{F}^X_s$-measurable version.

Step 2 We show $\mathcal{F}^X_{s+} \subseteq \mathcal{F}^\mathbb{P}_s$:
If $G \in \mathcal{F}^X_{s+} \subseteq \mathcal{F}^X_\infty$ then $\mathbb{P}(G|\mathcal{F}^X_{s+}) = \mathbb{1}_G$ a.s. By Step 1 there exists an $\mathcal{F}^X_s$ - measurable random variable $Y := \mathbb{P}(G|\mathcal{F}^X_s)$. Then $H := \{Y = 1\} \in \mathcal{F}^X_s$ and

$$H \Delta G := (H \setminus G) \cup (G \setminus H) \subseteq \{\mathbb{1}_G \neq Y\} \in \mathcal{N}^\mathbb{P}.$$

From the exercises we know that for any $t \geq 0$ it holds

$$\mathcal{F}^\mathbb{P}_t = \{G \subseteq \Omega : \exists H \in \mathcal{F}^X_t : H \Delta G \in \mathcal{N}^\mathbb{P}\}. \tag{12}$$

Hence $G \in \mathcal{F}^\mathbb{P}_s$, which means $\mathcal{F}^X_{s+} \subseteq \mathcal{F}^\mathbb{P}_s$.

Step 3 We show $\mathcal{F}^\mathbb{P}_{s+} \subseteq \mathcal{F}^\mathbb{P}_s$:
If $G \in \mathcal{F}^\mathbb{P}_{s+}$ then $\forall n \geq 1 \quad G \in \mathcal{F}^\mathbb{P}_{s+\frac{1}{n}}$. We use again (12) and conclude that there exists a set $H_n \in \mathcal{F}^X_{s+\frac{1}{n}}$ with $G \Delta H_n \in \mathcal{N}^\mathbb{P}$. Put

$$H = \bigcap_{m=1}^\infty \bigcup_{n=m}^\infty H_n.$$

Since $\bigcup_{n=m}^\infty H_n \supseteq \bigcup_{n=m+1}^\infty H_n$ we have $H = \bigcap_{m=M}^\infty \underbrace{\bigcup_{n=m}^\infty H_n}_{\in \mathcal{F}^X_{s+\frac{1}{m}}}$ $\forall M \in \mathbb{N}$. We get $H \in \mathcal{F}^X_{s+\frac{1}{M}}$ $\forall M \in \mathbb{N}$ and therefore $H \in \mathcal{F}^X_{s+} \subseteq \mathcal{F}^\mathbb{P}_s$. We show $G \in \mathcal{F}^\mathbb{P}_s$ by representing $G = (G \cup H) \setminus (H \setminus G) = ((H \Delta G) \cup H) \setminus (H \setminus G)$ where we have $H \in \mathcal{F}_s^{\mathbb{P}}$, and $H \Delta G \in \mathcal{N}^{\mathbb{P}}$ will be shown below (which especially implies also $H \setminus G \in \mathcal{N}^{\mathbb{P}}$). Indeed, we notice that $H \Delta G = (H \setminus G) \cup (G \setminus H)$,

$$H \setminus G \subseteq \left(\bigcup_{n=1}^{\infty} H_n \right) \setminus G = \bigcup_{n=1}^{\infty} (H_n \setminus G) \in \mathcal{N}^{\mathbb{P}},$$

and

$$\begin{aligned} G \setminus H = G \cap H^c &= G \cap \left(\bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} H_n \right)^c \\ &= G \cap \left(\bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} H_n^c \right) \\ &= \bigcup_{m=1}^{\infty} \left(G \cap \bigcap_{n=m}^{\infty} H_n^c \right) \\ &\subseteq \bigcup_{m=1}^{\infty} \left(\underbrace{G \cap H_m^c}_{G \setminus H_m \subseteq G \Delta H_m \in \mathcal{N}^{\mathbb{P}}} \right) \in \mathcal{N}^{\mathbb{P}}. \end{aligned}$$

So $H \Delta G \in \mathcal{N}^{\mathbb{P}}$ and hence $G \in \mathcal{F}_s^{\mathbb{P}}$.

$\square$

5 The semigroup/infinitesimal generator approach

5.1 Contraction semigroups

Definition 5.1 (semigroup)**.** (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 *semigroup* if

- $T(0) = Id$,
- $T(s+t) = T(s)T(t), \quad \forall s, t \geq 0.$

(2) A semigroup $\{T(t); t \geq 0\}$ is called *strongly continuous* (or C_0 semigroup) if

$$\lim_{t \downarrow 0} T(t)f = f, \quad \forall f \in \mathcal{B}.$$

(3) The semigroup $\{T(t); t \geq 0\}$ is a *contraction semigroup* if

$$\|T(t)\| = \sup_{\|f\|=1} \|T(t)f\| \leq 1, \quad \forall t \geq 0.$$

As a simple example consider $\mathcal{B} = \mathbb{R}^d$, let A be a $d \times d$ matrix and

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

with A^0 as identity matrix. One can show that $e^{(s+t)A} = e^{sA}e^{tA}, \quad \forall s, t \geq 0$, $\{e^{tA}; t \geq 0\}$ is strongly continuous, and $\|e^{tA}\| \leq e^{t\|A\|}, \quad t \geq 0$.

Definition 5.2. Let E be a separable metric space. By $\mathcal{B}_E$ we denote the space of bounded measurable functions

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

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

Lemma 5.3. *Let E be a complete separable metric space and X a homogeneous Markov process with transition function $\{P_t(x,A)\}$. The space $\mathcal{B}_E$ defined in Definition 5.2 is a Banach space, and $\{T(t); t \geq 0\}$ with*

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

is a contraction semigroup.

Proof. **Step 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.

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

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

we approximate f by $f_n = \sum_{k=1}^{N_n} a_k^n \mathbb{1}_{A_k^n}$, $A_k^n \in \mathcal{B}(E), a_k^n \in \mathbb{R}$ such that $|f_n| \uparrow |f|$. Then

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

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 $T(t)f$.

$$\|T(t)f\| \;=\; \sup_{x\in E} |T(t)f(x)|$$

$$
\begin{aligned}
&\leq \sup_{x\in E}\int_E |f(y)|P_t(x,dy) \\
&\leq \sup_{x\in E}\|f\|P_t(x,E) = \|f\|. \qquad (13)
\end{aligned}
$$

Hence $T(t)f \in \mathcal{B}_E$.
Step 3 $\{T(t); t \geq 0\}$ is a semigroup:
We have $T(0)f(x) = \int_E f(y)P_0(x,dy) = \int_E f(y)\delta_x(dy) = f(x)$. This implies that $T(0) = Id$. From Chapman Kolmogorov's equation we derive

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

Step 4 We have already seen in (13) that $\{T(t); t \geq 0\}$ is a contraction. □

5.2 Infinitesimal generator

Definition 5.4 (infinitesimal generator)**.** Let $\{T(t); t \geq 0\}$ be a contraction semigroup on $\mathcal{B}_E$. Define

$$
Af := \lim_{t\downarrow 0}\frac{T(t)f - f}{t}
$$

for each $f \in \mathcal{B}_E$ for which it holds: there exists a $g \in \mathcal{B}_E$ such that:

$$
\text{There exists a } g \in \mathcal{B}_E \text{ such that } \left\|\frac{T(t)f-f}{t} - g\right\| \to 0, \quad \text{for } t\downarrow 0. \qquad (14)
$$

Let $D(A) := \{f \in \mathcal{B}_E : (14) \text{ holds}\}$. Then

$$
A : D(A) \to \mathcal{B}_E
$$

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

Example 5.5. If W is the Brownian motion (one-dimensional) then $A = \frac{1}{2}\frac{d^2}{dx^2}$ and $C_c^2(\mathbb{R}) \subseteq D(A)$, where

$C_c^2(\mathbb{R}) := \{f : \mathbb{R} \to \mathbb{R} :$ twice continuously differentiable, compact support $\}$

We have $P_t(x, A) = \mathbb{P}(W_t \in A | W_0 = x)$ and

$$\begin{aligned} T(t)f(x) &= \mathbb{E}[f(W_t)|W_0 = x] \\ &= \mathbb{E}f(\widetilde{W}_t + x), \end{aligned}$$

where $\widetilde{W}$ is a standard Brownian motion starting in 0. By Itô's formula,

$$f(\widetilde{W}_t + x) = f(x) + \int_0^t f'(\widetilde{W}_s + x) d\widetilde{W}_s + \frac{1}{2}\int_0^t f''(\widetilde{W}_s + x) ds.$$

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

$$\mathbb{E}\int_0^t f'(\widetilde{W}_s + x) d\widetilde{W}_s = 0.$$

This implies

$$\mathbb{E}f(\widetilde{W}_t + x) = f(x) + \frac{1}{2}\mathbb{E}\int_0^t f''(\widetilde{W}_s + x) ds.$$

By Fubini's Theorem we get $\mathbb{E}\int_0^t f''(\widetilde{W}_s + x) ds = \int_0^t \mathbb{E}f''(\widetilde{W}_s + x) ds$. We notice that g given by $g(s) := \mathbb{E}f''(\widetilde{W}_s + x)$ is a continuous function. By the mean value theorem we may write

$$\int_0^t \mathbb{E}f''(\widetilde{W}_s + x) ds = \int_0^t g(s) ds = g(\xi)t, \quad \text{for some } \xi \in [0, t].$$

Hence

$$\begin{aligned} \frac{T(t)f(x) - f(x)}{t} &= \frac{\mathbb{E}f(\widetilde{W}_t + x) - f(x)}{t} = \frac{\frac{1}{2}\mathbb{E}\int_0^t f''(\widetilde{W}_s + x) ds}{t} \\ &= \frac{1}{2}\mathbb{E}f''(\widetilde{W}_\xi + x). \end{aligned}$$

This implies that for any given $\varepsilon > 0$ we can find by uniform continuity of f'' a $\delta > 0$ and get Chebyshev's inequality that

$$\begin{aligned}
&\left|\frac{T(t)f(x)-f(x)}{t}-\frac{1}{2}f''(x)\right| \\
=\;& \frac{1}{2}\left|\mathbb{E}f''(\widetilde{W}_\xi+x)-f''(x)\right| \\
\le\;& \frac{1}{2}\left|\mathbb{E}f''(\widetilde{W}_\xi+x)\mathbb{1}_{\{|\widetilde{W}_\xi|\le\delta\}}-f''(x)\right|+\frac{1}{2}\left|\mathbb{E}f''(\widetilde{W}_\xi+x)\mathbb{1}_{\{|\widetilde{W}_\xi|>\delta\}}\right| \\
\le\;& \frac{1}{2}\sup_{|y-x|\le\delta}|f''(y)-f''(x)|+\frac{1}{2}\sup_{x\in\mathbb{R}}|f''(x)|\mathbb{P}(|\widetilde{W}_\xi|>\delta) \\
\le\;& \frac{1}{2}\varepsilon+\frac{1}{2}\sup_{x\in\mathbb{R}}|f''(x)|\frac{\mathbb{E}|\widetilde{W}_\xi|^2}{\delta^2}\le\varepsilon
\end{aligned}$$

for $0 \le \xi \le t$ small.

Theorem 5.6. *Let $\{T(t); t \ge 0\}$ be a contraction semigroup and A its infinitesimal generator with domain $D(A)$. Then*

(i) If $f \in \mathcal{B}_E$ such that $\lim_{t\downarrow 0} T(t)f = f$,
then for $t \ge 0$ it holds $\int_0^t T(s)fds \in D(A)$ and

$$T(t)f - f = A\int_0^t T(s)fds.$$

(ii) If $f \in D(A)$ and $t \ge 0$, then $T(t)f \in D(A)$ and

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

(iii) If $f \in D(A)$ and $t \ge 0$ then $\int_0^t T(s)fds \in D(A)$ and

$$T(t)f - f = A\int_0^t T(s)fds = \int_0^t AT(s)fds = \int_0^t T(s)Afds.$$

Proof. (i) 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 $T(u): \mathcal{B}_E \to \mathcal{B}_E$:

$$\|T(u)f_n - T(u)f\| = \|T(u)(f_n - f)\| \le \|f_n - f\|$$

Hence the Riemann integral

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

exists for all $t, s \geq 0$. 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

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

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

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

(ii) 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 $AT(t)f = T(t)Af$.

(iii) 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 (i), we get $\int_0^t T(u)f du \in D(A)$. From (ii) we get by integrating

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

On the other hand, in the proof of (i) 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:

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

□

5.3 Martingales and Dynkin's formula

Definition 5.7 (martingale)**.** *An $\mathbb{F}$-adapted stochastic process $X = \{X_t; t \in \mathbf{T}\}$ such that $\mathbb{E}|X_t| < \infty$ $\forall t \in \mathbf{T}$ is called $\mathbb{F}$-martingale (submartingale, supermartingale) if for all $t, t+h \in \mathbf{T}$ with $h \geq 0$ it holds*

$$\mathbb{E}[X_{t+h}|\mathcal{F}_t] = (\geq, \leq) X_t \quad a.s.$$

Theorem 5.8 (Dynkin's formula)**.** *Let X be a homogeneous Markov process with cádlág paths for all $\omega \in \Omega$ and transition function $\{P_t(x, A)\}$. Let $\{T(t); t \geq 0\}$ denote its semigroup $T(t)f(x) = \int_E f(y)P_t(x, dy)$ $(f \in \mathcal{B}_E)$ 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*

$$M_t := g(X_t) - g(X_0) - \int_0^t Ag(X_s)ds. \tag{15}$$

(The integral $\int_0^t Ag(X_s)ds$ is understood as a Lebesgue-integral for each ω:

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

where λ denotes the Lebesgue measure.)

Proof. Since by Definition 5.4 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 cádlág paths and is adapted, it is (see Lemma 4.11) progressively measurable:

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

Hence for the composition we have

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

Moreover, Ag is bounded as it is from $\mathcal{B}_E$. So the integral w.r.t. the Lebesgue measure λ is well-defined:

$$\int_0^t Ag(X_s(\omega))\lambda(ds), \quad \omega \in \Omega.$$

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 (15)

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

The Markov property from Definition 2.4(3) (equation (4)) 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 $\mathbb{E}\left[\int_t^{t+h} Ag(X_s)ds \middle| \mathcal{F}_t^X\right] = \int_t^{t+h} \mathbb{E}[Ag(X_s)|\mathcal{F}_t^X]ds$. 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

$$\begin{aligned}\int_\Omega \int_t^{t+h} Ag(X_s)ds\mathbb{1}_G d\mathbb{P} &= \int_t^{t+h}\int_\Omega Ag(X_s)\mathbb{1}_G d\mathbb{P}ds \\ &= \int_t^{t+h}\int_\Omega \mathbb{E}[Ag(X_s)|\mathcal{F}_t^X]\mathbb{1}_G d\mathbb{P}ds.\end{aligned}$$

The Markov property implies that $\mathbb{E}[Ag(X_{t+h})|\mathcal{F}_t^X] = \int_E Ag(y)P_h(X_t, dy)$. Therefore we have

$$\begin{aligned}&\mathbb{E}\left[\left(g(X_{t+h}) - \int_t^{t+h} Ag(X_s)ds\right)\middle|\mathcal{F}_t^X\right] - \int_0^t Ag(X_s)ds \\ &= \int_E g(y)P_h(X_t, dy) - \int_t^{t+h}\int_E Ag(y)P_{s-t}(X_t, dy)ds \\ &\qquad - \int_0^t Ag(X_s)ds.\end{aligned}$$

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 5.6 (iii).

□

6 Weak solutions of SDEs and martingale problems

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

Definition 6.1. 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

$$dX_t = \sigma(X_t)dB_t + b(X_t)dt, \quad X_0 = x, \quad t \geq 0 \tag{16}$$

is a triple $(X_t, B_t)_{t\geq 0}$, $(\Omega, \mathcal{F}, \mathbb{P})$, $(\mathcal{F}_t)_{t\geq 0}$ such that

(i) $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$ 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,

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

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

(iv) $X_t^{(i)} = x^{(i)} + \sum_{j=1}^m \int_0^t \sigma_{ij}(X_u)dB_u^{(j)} + \int_0^t b_i(X_u)du$,
$t \geq 0$, $1 \leq i \leq d$, a.s.

Let $a_{ij}(x) = \sum_{k=1}^m \sigma_{ik}(x)\sigma_{jk}(x)$ (or using the matrices: $a(x) = \sigma(x)\sigma^T(x)$). Consider now the differential operator

$$Af(x) = \frac{1}{2}\sum_{ij} a_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} f(x) + \sum_i b_i(x)\frac{\partial}{\partial x_i} f(x).$$

with domain $D(A) = C_c^2(\mathbb{R}^d)$, the twice continuously differentiable functions with compact support in $\mathbb{R}^d$. Then it follows from Itô's formula that

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

is a martingale.

By $\Omega := C_{\mathbb{R}^d}[0,\infty)$ we denote the space of continuous functions $\omega : [0,\infty) \to \mathbb{R}^d$. One can introduce a metric on this space setting

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

Then $C_{\mathbb{R}^d}[0,\infty)$ with this metric is a complete separable metric space ([8, Problem 2.4.1]). We set

$$\mathcal{F}_t := \sigma\{\pi_s, s \in [0,t]\}$$

where

$$\pi_s : C_{\mathbb{R}^d}[0,\infty) \to \mathbb{R}^d : \omega \mapsto \omega(s)$$

is the coordinate mapping. For $0 \le t \le u$ we have

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

([8, Problem 2.4.2]). We define local martingales to introduce the concept of a martingale problem.

Definition 6.2 (local martingale)**.** *A continuous $(\mathcal{F}_t)_{t\ge0}$ adapted process $M = (M_t)_{t\ge0}$ with $M_0 = 0$ is a* **local martingale** *if there exists a sequence of stopping times $\tau_1 \le \tau_2 \le \tau_3 ... \uparrow \infty$ a.s. 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 6.3. *The process which solves*

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

is a martingale if $0 \le \alpha \le 1$ and it is a local martingale but not a martingale for $\alpha > 1$.

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

Definition 6.4 ($C_{\mathbb{R}^d}[0,\infty)$- 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 A is probability measure $\mathbb{P}$ on
$(C_{\mathbb{R}^d}[0,\infty), \mathcal{B}(C_{\mathbb{R}^d}[0,\infty)))$ satisfying

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

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.

Theorem 6.5. *X (or more exactly, the distribution of X given by a probability measure $\mathbb{P}$ on $(C_{\mathbb{R}^d}[0,\infty), \mathcal{B}(C_{\mathbb{R}^d}[0,\infty)))$) is a solution of the $C_{\mathbb{R}^d}[0,\infty)$-martingale problem for A $\iff$ X is a weak solution of (16).*

Proof. We have seen above that $\Longleftarrow$ follows from Itô's formula.
We will show $\Longrightarrow$ only for the case $d = m$. See [8, 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 A.
One can conclude that then for any $f(x) = x_i$ $(i = 1, ..., d)$ the process $\{M_t^i := M_t^f; t \geq 0\}$ is a continuous, local martingale. This can be seen as follows: If we define the stopping times for $n > \max\{|x^{(1)}|, ..., |x^{(d)}|\}$

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

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.
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

$$\mathbb{P}(\{\omega : \int_0^t |b_i(X_s(\omega))|ds < \infty; 0 \leq t < \infty\}) = 1.$$

Also for $f(x) = x_i x_j$ the process $M_t^{(ij)} := M_t^f$ is a continuous, local martingale.

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

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

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

Indeed,

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

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 6.6. *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 [8, Theorem 3.3.16]) implies that B is a Brownian motion. □

Definition 6.7. (1) Given an initial distribution μ on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, we say that *uniqueness holds for the $C_{\mathbb{R}^d}[0, \infty)$-martingale problem* for (A, μ) if any two solutions of the $C_{\mathbb{R}^d}[0, \infty)$-martingale problem for A with initial distribution μ have the same finite dimensional distributions.

(2) *Weak uniqueness* holds for (16) with initial distribution μ if any two weak solutions of (16) with initial distribution μ have the same finite dimensional distributions.

Note that Theorem 6.5 does not assume uniqueness. Consequently, existence and uniqueness for the two problems are equivalent.

Corollary 6.8. *Let μ be a probability measure on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. The following assertions are equivalent:*

(1) Uniqueness holds for the martingale problem for (A, μ).

(2) Weak uniqueness holds for (16) with initial distribution μ.

Remark 6.9. *There exist sufficient conditions on σ and b such that the martingale problem with $a = \sigma\sigma^T$ has a unique weak solution. For example, it is enough to require that σ and b are continuous and bounded.*

7 Feller processes

7.1 Feller semigroups, Feller transition functions and Feller processes

Definition 7.1.

(1) $C_0(\mathbb{R}^d) := \{f : \mathbb{R}^d \to \mathbb{R} : f \text{ continuous}, \lim_{|x|\to\infty} |f(x)| = 0\}$.

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

  (a) $T(t) : C_0(\mathbb{R}^d) \to C_0(\mathbb{R}^d)$ is positive $\forall t \geq 0$ (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 semigroup.

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

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

Proposition 7.2. *Let $\{T(t); t \geq 0\}$ be a conservative Feller semigroup on $C_0(\mathbb{R}^d)$. Then there exists a (homogeneous) transition function $\{P_t(x, A)\}$ such that*

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

Proof. Recall the Riesz representation theorem (see, for example, [6, 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

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

Definition 7.3. A Borel measure on $(E, \mathcal{B}(E))$ (if E is a locally compact Hausdorff space) is a *Radon measure* $\Longleftrightarrow_{df}$

(1) $\mu(K) < \infty, \quad \forall K$ compact,

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

(3) $\forall B$ open: $\mu(B) = \sup\{\mu(K) : K \subseteq B, K \text{ compact }\}$,

Remark: Any probability measure on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ is a Radon measure.

By 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.

Step 1 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 $P_t(x, \mathbb{R}^d) = 1$. This left as an exercise.

Step 2 For $A \in \mathcal{B}(\mathbb{R}^d)$ we have to show that

$$x \mapsto P_t(x, A) : (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)) \to (\mathbb{R}, \mathcal{B}(\mathbb{R})). \tag{17}$$

Using the monotone class theorem for

$$\begin{aligned} H \;=\; & \{f : \mathbb{R}^d \to \mathbb{R} : \mathcal{B}(\mathbb{R}^d) \text{ measurable and bounded,} \\ & T(t)f \text{ is } \mathcal{B}(\mathbb{R}^d) \text{ measurable }\} \end{aligned}$$

we see that it is enough to show that

$$\forall A \in \mathcal{A} := \{[a_1, b_1] \times ... \times [a_n, b_n]; a_k \leq b_k\} \cup \emptyset : \mathbb{1}_A \in H.$$

We will approximate such $\mathbb{1}_A$ by $f_n \in C_c(\mathbb{R}^d)$: 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 $T(t)f : C_0(\mathbb{R}^d) \to C_0(\mathbb{R}^d)$ and $C_c(\mathbb{R}^d) \subseteq C_0(\mathbb{R}^d)$, we get

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

It holds $T(t)f_n(x) = \int_{\mathbb{R}^d} f_n(y)P_t(x, dy) \to P_t(x, A)$ for $n \to \infty$. Hence $P_t(\cdot, A) : (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, which means $\mathbb{1}_A \in H$.

Step 3 The Chapman-Kolmogorov equation for $\{P_t(x, A)\}$ we conclude from $T(t+s) = T(t)T(s)\ \forall s, t \geq 0$ (This can be again done by approximating $\mathbb{1}_A$, $A \in \mathcal{A}$ and using dominated convergence and the Monotone Class Theorem).

Step 4 $T(0) = Id$ gives $P_0(x, A) = \delta_x(A)$ (again by approximating).

□

Definition 7.4. A transition function associated to a Feller semigroup is called a *Feller transition function.*

Proposition 7.5. *A transition function $\{P_t(x, A)\}$ is Feller* $\iff$

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

(ii) $\forall f \in C_0(\mathbb{R}^d)$, $x \in \mathbb{R}^d :\ \lim_{t\downarrow 0} \int_{\mathbb{R}^d} f(y)P_t(x, dy) = f(x)$.

Proof. $\Longleftarrow$ We will show that (i) and (ii) 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 semigroup. By know by Lemma 5.3 that $\{T(t);\ t \geq 0\}$ is a contraction semigroup. By (i) we have that $T(t) : C_0(\mathbb{R}^d) \to C_0(\mathbb{R}^d)$. Any $T(t)$ is positive. So we only have to show that $\forall f \in C_0(\mathbb{R}^d)$

$$\|T(t)f - f\| \to 0, \quad t \downarrow 0$$

which is the strong continuity.
Since by (i) $T(t)f \in C_0(\mathbb{R}^d)$ we conclude by (ii) that for all $x \in \mathbb{R}^d$: $\lim_{s\downarrow 0} T(t + s)f(x) = T(t)f(x)$. Hence we have

$$t \mapsto T(t)f(x) \text{ is right-continuous,}$$

$$x \mapsto T(t)f(x) \text{ is continuous.}$$

This implies (similarly to the proof 'right-continuous + adapted $\Longrightarrow$ progressively measurable')

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

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 semigroup, it holds $\|T(\frac{u}{p})f\| \leq \|f\|$. Hence we can use dominated convergence in the following expression, and it follows from (ii) that

$$p\mathcal{R}_p f(x) = \int_0^\infty e^{-pt} T(t)f(x)dt = \int_0^\infty e^{-u} T\left(\frac{u}{p}\right) f(x)du \to f(x). \tag{18}$$

for $p \to \infty$. Moreover, one can easily show that $\mathcal{R}_p f \in C_0(\mathbb{R}^d)$.
For the resolvent: $f \mapsto \mathcal{R}_p f$ it holds

$$\begin{aligned}
(q-p)\mathcal{R}_p\mathcal{R}_q f &= (q-p)\mathcal{R}_p \int_0^\infty e^{-qt} T(t) f dt \\
&= (q-p) \int_0^\infty e^{-ps} T(s) \int_0^\infty e^{-qt} T(t) f dt ds \\
&= (q-p) \int_0^\infty e^{-(p-q)s} \int_0^\infty e^{-q(t+s)} T(t+s) f dt ds \\
&= (q-p) \int_0^\infty e^{-(p-q)s} \int_s^\infty e^{-qu} T(u) f du ds \\
&= (q-p) \int_0^\infty e^{-qu} T(u) f \int_0^u e^{-(p-q)s} ds du \\
&= (q-p) \int_0^\infty e^{-qu} T(u) f \frac{1}{q-p} (e^{-(p-q)u} - 1) du
\end{aligned}$$

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

Let $D_p := \{\mathcal{R}_p f; f \in C_0(\mathbb{R}^d)\}$. Then $D_p = D_q =: D$. Indeed, if $g \in D_p$ then there exists $f \in C_0(\mathbb{R}^d): \; g = \mathcal{R}_p f$. Since

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

we conclude $g \in D_q$ and hence $D_p \subseteq D_q$ and, by symmetry, $D_q \subseteq D_p$.
By (18)

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

We show that $D \subseteq C_0(\mathbb{R}^d)$ is dense. We follow [6, Section 7.3] and notice that $C_0(\mathbb{R}^d)$ is the closure of $C_c(\mathbb{R}^d)$ with respect to $\|f\| := \sup_{x\in\mathbb{R}^d} |f(x)|$. A positive linear functional L on $C_c(\mathbb{R}^d)$ can be represented uniquely by a Radon measure on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$:

$$L(f) = \int_{\mathbb{R}^d} f(x)\mu(dx).$$

Since $\mu(\mathbb{R}^d) = \sup\{\int_{\mathbb{R}^d} f(x)\mu(dx) : f \in C_c(\mathbb{R}^d), 0 \leq f \leq 1\}$, we see that we can extend L to a positive linear functional on $C_0(\mathbb{R}^d) \iff \mu(\mathbb{R}^d) < \infty$.
In fact, any positive linear functional on $C_0(\mathbb{R}^d)$ has the representation $L(f) = \int_{\mathbb{R}^d} f(x)\mu(dx)$ with a finite Radon measure μ ([6, Proposition 7.16]).

Since D is a linear space in view of Hahn-Banach we should have a linear functional L on $C_0(\mathbb{R}^d)$ given by $L(f) = \int_{\mathbb{R}^d} f(x)\mu(dx)$ (here μ is a signed measure) which is 0 on D and positive for an $f \in C_0(\mathbb{R}^d)$ which is outside the closure of D. But by dominated convergence we have

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

which implies that D is dense. We have

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

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

This implies

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

So we have shown that $\{T(t); t \geq 0\}$ is strongly continuous on D. Since $D \subseteq C_0(\mathbb{R}^d)$ is dense, we can also show strong continuity on $C_0(\mathbb{R}^d)$. The direction $\Longrightarrow$ is obviously trivial. □

Definition 7.6. A Markov process having a Feller transition function is called a *Feller process.*

7.2 Càdlàg modifications of Feller processes

In Definition 4.14 we defined a Lévy process as a stochastic process with a.s. càdlàg paths. In Theorem 4.15 we have shown that a Lévy process (with càdlàg paths) is a strong Markov process. By the Daniell-Kolmogorov Theorem (Theorem 3.2) we know that Markov processes exist. 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 7.7 (Lévy process in law)**.** *A stochastic process* $X = \{X_t; t \geq 0\}$ *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.* $\forall t \geq 0, \forall \varepsilon > 0$

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

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

(3) $\forall\, 0 \leq s \leq t: \; X_t - X_s \overset{d}{=} X_{t-s}$,

(4) $\forall\, 0 \leq s \leq t: \; X_t - X_s$ *is independent of* $\mathcal{F}_s^X$.

Theorem 7.8. *Let* X *be an* $\{\mathcal{F}_t; t \geq 0\}$*-submartingale. Then it holds* *(i)* *For any countable dense subset* $D \subseteq [0,\infty)$, $\exists \Omega^* \in \mathcal{F}$ *with* $\mathbb{P}(\Omega^*) = 1$, *such that for every* $\omega \in \Omega^*$:

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

exists $\forall t \geq 0$ *(*$t > 0$*, respectively).*

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

(iii) *Assume that* $\{\mathcal{F}_t; t \geq 0\}$ *satisfies the usual conditions. Then it holds:* X *has a càdlàg modification* $\iff$ $t \mapsto \mathbb{E}X_t$ *is right-continuous.*

Proof: See [8, Proposition 1.3.14 and Theorem 1.3.13]

Lemma 7.9. *Let* X *be a Feller process. For any* $p > 0$ *and any* $f \in C_0(\mathbb{R}^d; [0,\infty)) := \{f \in C_0(\mathbb{R}^d) : f \geq 0\}$ *the process* $\{e^{-pt}R_pf(X_t); t \geq 0\}$ *is a supermartingale w.r.t. the natural filtration* $\{\mathcal{F}_t^X; t \geq 0\}$ *and for any probability measure* $\mathbb{P}_\nu$:

$$\mathbb{P}_\nu(X_0 \in B) = \nu(B), \quad B \in \mathcal{B}(\mathbb{R}^d),$$

where ν *denotes the initial distribution.*

Proof. Recall that for $p > 0$ we defined in the proof of Proposition 7.5 the resolvent

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

Step 1 We show that $\mathcal{R}_p : C_0(\mathbb{R}^d) \to C_0(\mathbb{R}^d)$:

Since

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

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

$$\lim_{x_n\to x} \mathcal{R}_pf(x_n) \quad = \quad \lim_{x_n\to x} \int_0^\infty e^{-pt}T(t)f(x_n)dt$$

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

In the same way: $\lim_{|x_n|\to\infty} \mathcal{R}_p f(x_n) = 0$.
Step 2 We show that $\forall x \in \mathbb{R}^d$: $e^{-ph}T(h)\mathcal{R}_p f(x) \leq \mathcal{R}_p f(x)$ provided that $f \in C_0(\mathbb{R}^d; [0,\infty))$ and $h > 0$:

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

Step 3 $\{e^{-pt}R_p f(X_t); t \geq 0\}$ is a supermartingale:
Let $0 \leq s \leq t$. Since X is a Feller process, it has a transition function, and by Definition 2.4 (3) 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 2 we conclude

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

□

Lemma 7.10. *Let Y_1 and Y_2 be random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ with values in $\mathbb{R}^d$. Then*

$$Y_1 = Y_2 \quad a.s. \quad \Longleftrightarrow \quad \begin{array}{l} \mathbb{E}f(Y_1)g(Y_2) = \mathbb{E}f(Y_1)g(Y_1) \\ \forall f, g : \mathbb{R}^d \to \mathbb{R} \text{ bounded and continuous} \end{array}$$

Proof. The direction $\implies$ is clear.
We will use the Monotone Class Theorem (Theorem A.1) to show $\impliedby$. 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 $\mathbb{1}_{[a_1,b_1]\times\ldots\times[a_{2d},b_{2d}]}$ 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 $h := \mathbb{1}_{\{(x,y)\in\mathbb{R}^d\times\mathbb{R}^d : x\neq y\}}$ and infer

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

□

Theorem 7.11. *If X is a Feller process, then it has a càdlàg modification.*

Proof. **Step 1.** We need instead of the $\mathbb{R}^d$ a compact space. We use the *one-point compactification* (Alexandroff extension):
Let ∂ be a point not in $\mathbb{R}^d$. We define a topology $\mathcal{O}'$ on $(\mathbb{R}^d)^\partial := \mathbb{R}^d \cup \{\partial\}$ as follows: Denote by $\mathcal{O}$ the open sets of $\mathbb{R}^d$. We define

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

Then $((\mathbb{R}^d)^\partial, \mathcal{O}')$ is a compact Hausdorff space.

Remark. This construction can be done for any locally compact Hausdorff space.

Step 2. Let $(f_n)_{n=1}^\infty \subseteq C_0(\mathbb{R}^d; [0,\infty))$ be a sequence which separates the points: For any $x, y \in (\mathbb{R}^d)^\partial$ with $x \neq y$ $\exists k \in \mathbb{N}: f_k(x) \neq f_k(y)$. (Such a sequence exists: exercise).
We want to show that then also

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

is a countable set (this is clear) which separates points: It holds for any $p > 0$

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

This implies

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

by dominated convergence since $\|T\left(\frac{u}{p}\right)f - f\| \leq 2\|f\|$, and the strong continuity of the semigroup 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_k with $f_k(x) \neq f_k(y)$ and can find a $p \in \mathbb{N}$ such that $\mathcal{R}_p f_k(x) \neq \mathcal{R}_p f_k(y)$.

Step 3. We fix a set $D \subseteq [0,\infty)$ which is countable and dense. We show that $\exists \Omega^* \in \mathcal{F}$ with $\mathbb{P}(\Omega^*) = 1$:

$$\forall\omega \in \Omega^* \forall n, p \in \mathbb{N}^* : [0,\infty) \ni t \mapsto \mathcal{R}_p f_n(X_t(\omega)) \tag{19}$$

has right and left (for $t > 0$) limits along D.

For this we conclude from Lemma 7.9 that

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

By Theorem 7.8 (i) 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 $\forall\omega \in \Omega^*_{n,p} \forall t \geq 0 (t > 0)$

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

Since $s \mapsto e^{ps}$ is continuous we get assertion (19) by setting $\Omega^* := \bigcap_{n=1}^\infty \bigcap_{p=1}^\infty \Omega^*_{n,p}$.

Step 4. We show: $\forall \omega \in \Omega^* : \ t \to X_t(\omega)$ has right limits along D. If $\nexists \lim_{s \downarrow t, s \in D} X_s(\omega)$ then $\exists x, y \in (\mathbb{R}^d)^\partial$ 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 $\exists p, k : \ \mathcal{R}_p f_k(x) \neq \mathcal{R}_p f_k(y)$ which is a contradiction to the fact that $s \mapsto \mathcal{R}_p f_k(X_s(\omega))$ has right limits along D.

Step 5. Construction of a right-continuous modification:
For $\omega \in \Omega^*$ set $\forall t \geq 0$:

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

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

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

Since for $f, g \in C((\mathbb{R}^d)^\partial)$ 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 7.10 we conclude $\tilde{X}_t = X_t$ a.s.

We check that $t \to \tilde{X}_t$ is right-continuous $\forall \omega \in \Omega$: For $\omega \in \Omega^*$ consider for $\delta > 0$

$$|\tilde{X}_t(\omega) - \tilde{X}_{t+\delta}(\omega)| \leq |\tilde{X}_t(\omega) - X_s(\omega)| + |X_s(\omega) - \tilde{X}_{t+\delta}(\omega)|$$

where $|\tilde{X}_t(\omega) - X_s(\omega)| < \varepsilon$ for all $s \in (t, t+\delta_1(t)) \cap D$ and $|X_s(\omega) - \tilde{X}_{t+\delta}(\omega)| < \varepsilon$ for all $\delta < \delta_1(t)$ and $s \in (t+\delta, t+\delta+\delta_2(t+\delta)) \cap D$. Hence $t \to \tilde{X}_t$ is right-continuous.

Step 6. càdlàg modifications:

We use [8, 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 $\omega \in \Omega$.

□

Remark 7.12. *Since we used the one point compactification of $\mathbb{R}^d$, we are not able to distinguish, for example, if a sequence $(X_{s_n})_{n\geq 1}$ converges to $-\infty$ or $+\infty$ if $d = 1$.*
However, for a Lévy process it can be shown (see [7, Theorem II.2.68]) that for every $c > 0$

$$\mathbb{P}(\sup\{|X_s| : s \in [0, c] \cap D\} < \infty) = 1.$$

Consequently, $\lim_n |X_{s_n}| = \partial$ has probability 0 and $\tilde{X}$ is a càdlàg version with values in $\mathbb{R}^d$.

A Appendix

Theorem A.1 (Monotone Class Theorem for functions)**.** *Let $\mathcal{A} \subseteq 2^\Omega$ be a π-system that contains Ω. Assume that for $\mathcal{H} \subseteq \{f; f : \Omega \to \mathbb{R}\}$ it holds*

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

(ii) linear combinations of elements of $\mathcal{H}$ *are again in* $\mathcal{H}$,

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

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

Proof. see [7].

Lemma A.2 (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:*

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

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

Proof. see [2, p. 62]

Theorem A.3. *Suppose $M_t^1, ..., M_t^d$ are continuous, local martingales on $(\Omega, \mathcal{F}, \mathbb{P})$ w.r.t. $\mathcal{F}_t$. If for $1 \leq i, j \leq d$ the processes $\langle M^i, M^j \rangle_t$ is an absolutely continuous function in t $\mathbb{P}$- a.s. then there exists an extension $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ of $(\Omega, \mathcal{F}, \mathbb{P})$ carrying a d-dimensional $\tilde{\mathcal{F}}$ Brownian motion B and measurable, adapted 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.*

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

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

Proof: [8, Theorem 3.4.2, page 170]
An Itô process has the form

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

μ and σ are progressively measurable and satisfy

$$\int_0^t \mu(s)ds < \infty, \quad \int_0^t \sigma(s)^2 ds < \infty \, a.s.$$

Lemma 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

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

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

Index

References

[1] D. Applebaum, *Lévy Processes and Stochastic Calculus* (Second Edition), Cambridge 2007.

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

[3] R.M. Blumenthal, R.K. Getoor, *Markov Processes and Potential Theory.* Academic Press, New York, 1968.

[4] Erhan Cinlar, *Probability and Statistics* , Springer, 2011.

[5] S.N. Ethier, T.G. Kurtz, *Markov Processes*, Wiley, 2005.

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

[7] Sheng-wu He, Jia-gang Wang, Jia-An Yan, *Semimartingale Theory and Stochastic Calculus*, Taylor & Francis, 1992.

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

[9] F. Klebaner, *Introduction to Stochastic Calculus*

[10] D. Revuz, M. Yor, *Continuous Martingales and Brownian Motion*, 1998

[11] L.C.G. Rogers, D. Williams, *Diffusions, Markov Processes, and Martingales I*, Cambridge University Press 1994.

[12] L.C.G. Rogers, D. Williams, *Diffusions, Markov Processes, and Martingales II*, Cambridge University Press 2000.

[13] T. Rolski, H. Schmidli, V. Schmidt and J.Teugels, *Stochastic Processes for Insurance and Finance*, 1998.

[14] Ken-iti Sato, *Lévy processes and infinitely divisible distributions*, Cambridge, 1999.

[15] http://almostsure.wordpress.com/tag/levy-process/ ☺