Oberseminar dynamische Systeme im Sommersemester 2020

Markov and Feller Processes

A Semigroup Theoretic Approach

Julian Hölz
(Matrikelnummer 77903)

betreut durch:
Prof. Dr. Fabian Wirth
Lehrstuhl für dynamische Systeme

29. Juli 2020

Abstract

We present the notions of Markov processes and investigate their relation to semigroups of operators, the so called transition semigroups. We then focus on the special case of Feller processes and their corresponding transition semigroups. We proof a result on the strong continuity of this semigroup and give a canonical construction of a Feller process. We proof Dynkin's formula and investigate the generator of a Feller semigroup. We conclude the paper by presenting a result for the special case of a diffusion processes, hereby linking the concept of Feller processes to partial differential equations.

Contents

1 Motivation

In deterministic dynamics the cocycle or the semigroup properties assure that the future evolution of the process does not depend on states from the past but only on the current state of the system. For stochastic dynamics the Markov property can be view as its analog. It is therefore natural that Markov processes are closely connected to the theory of semigroups, more precisely that of semigroups of operators on Banach spaces. A central role here is played by the notions of transition functions and transition semigroups.

The theory of Feller processes is basically a branch of that of Markov processes and any systematic treatment of them is usually preceded by some theory on Markov processes and their corresponding transition semigroups, which we refer to as Feller semigroup. For Feller semigroups we assume additional invariance and continuity properties, which in turn yield that Feller processes can be assumed to have càdlàg paths. The Feller properties turn out to be sufficient to obtain strong continuity of the Feller semigroup, which by Hille's and Yoshida's theorem yields the existence and uniqueness of a generator of said semigroup.

In Chapter 2 we introduce some important notion for future investigation. Chapter 3 investigates the notions of transition functions, transition semigroups on the space of bounded measurable real valued functions and Markov processes. We will first treat each topic separately by presenting some important properties. The main objective here is to find a framework in which we can relate these concepts to each other. Finally, in Chapter 4 we start by introducing the Feller semigroups and Feller processes and proceed by stating the results guaranteeing the existence of a generator of the semigroup. We continue by giving a construction of a canonical Feller process, proving Dynkin's formula and providing a result on diffusion processes.

2 Preliminaries

2.1 Notation

The tuple $(E, \mathcal{B})$ usually denotes a measurable spaces. By a *Polish space* we refer to a complete separable metrizable space. To every Polish space we can assign the *Borel-σ-algebra*, which is the σ-algebra generated by the open sets of E.

For a measure space $(\Omega, \mathcal{A}, \mu)$ and a measurable map $f : (\Omega, \mathcal{A}) \to (E, \mathcal{B})$ we denote by μf^{-1} the *pushforward measure* defined by

$$\mu f^{-1}(\Gamma) = \mu(\{f \in \Gamma\}), \quad \Gamma \in \mathcal{B}.$$

We continue by introducing some important function spaces.

The space of càdlàg functions $D(\mathbb{R}_+; E)$. Some stochastic processes can be chosen such that its paths satisfy certain regularity properties. One very important path regularity is that of having left limits and being continuous from the right. We will quickly introduce this space and remark some of its properties. For this let E be a Polish space.

The space of *càdlàg* (continue à droite, limite à gauche) functions on $\mathbb{R}_+$ mapping into E is given by

$$D(\mathbb{R}_+; E) := \{\omega \in E^{\mathbb{R}_+} \mid \forall t \geq 0 : \lim_{s \to t^-} \omega(s) =: \omega(t-) \in E,\ \lim_{s \to t^+} \omega(s) = \omega(t)\}.$$

In Billingsley [Bil68, Section 16] it is shown that $D(\mathbb{R}_+; E)$ can be made a Polish space. The associated topology is called the *Skorohod topology*. Moreover, the Borel-σ-algebra with respect to the Skorohod topology $\mathcal{D}$ on $D(\mathbb{R}_+; E)$ coincides with the trace-σ-algebra of $\mathcal{B}^{\mathbb{R}_+}$ in $D(\mathbb{R}_+; E)$, cf. Billingsley [Bil68, Theorem 16.6. (iii)] or Gänssler and Stute [GS77, Section 7.2.10].

The space $B_b(E)$. Let $(E, \mathcal{B})$ be a measurable space. By $B_b(E)$ we denote the set of all bounded $\mathcal{B}$-measurable functions $f : E \to \mathbb{R}$. Together with the pointwise defined addition and scalar multiplication this set becomes an $\mathbb{R}$-vector spaces. Endowing $B_b(E)$ with the supremum norm it becomes a Banach space.

The space $C_0(E)$. Let E be a locally compact topological space with Borel-σ-algebra $\mathcal{B}$, so $(E, \mathcal{B})$ is a measurable Borel-space. We denote by $C_0(E)$ the set of all continuous functions $f : E \to \mathbb{R}$, which vanish at infinity, i.e., for every $\varepsilon > 0$ there exists a compact set $K \subset E$ such that $|f(x)| < \varepsilon$ for all $x \in E \backslash K$. Also $C_0(E)$ is a closed subspace of $B_b(E)$.

The spaces $\mathrm{ca}(\mathcal{B})$ and $\mathrm{rca}(\mathcal{B})$. Let $(E, \mathcal{B})$ be a measurable space. By $\mathrm{ca}(\mathcal{B})$ we denote the set of all finite signed measures on $\mathcal{B}$, i.e., every mapping $\mu : \mathcal{B} \to \mathbb{R}$ that satisfy

$$\mu\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n \in \mathbb{N}} \mu\left(A_n\right),$$

for every sequence A_1, $A_2, \ldots$ of disjoint sets in $\mathcal{B}$. Together with the pointwise defined addition and scalar multiplication this set becomes an $\mathbb{R}$-vector spaces.

For $\mathrm{ca}(\mathcal{B})$ we can define a norm using the total variation of finite measures in the following way: For each $\mu \in \mathrm{ca}(\mathcal{B})$ there exists a *Hahn-Decomposition* of E into two disjoint sets $\Gamma_\mu^+, \Gamma_\mu^- \in \mathcal{B}$ such that $E = \Gamma_\mu^+ \cup \Gamma_\mu^-$ and for all $\mathcal{B}$-measurable sets $\Gamma \subset \Gamma_\mu^+$, resp. $\Gamma \subset \Gamma_\mu^-$ we have $\mu(\Gamma) \geq 0$, resp. $\mu(\Gamma) \leq 0$; see Billingsley [Bil95, Theorem 32.1]. Then the mappings defined by

$$\mu^+(\Gamma) := \mu(\Gamma \cap \Gamma_\mu^+) \quad \text{and} \quad \mu^-(\Gamma) := -\mu(\Gamma \cap \Gamma_\mu^-),$$

for $\Gamma \in \mathcal{B}$, are finite non-negative measures. The total variation of μ is then defined as

$$|\mu|(\Gamma) := \mu^+(\Gamma) + \mu^-(\Gamma), \quad \Gamma \in \mathcal{B}.$$

It is easily checked that $|\mu|$ defines a non-negative finite measure and $\|\mu\| := |\mu|(E)$ defines a norm on $\mathrm{ca}(\mathcal{B})$ such that $(\mathrm{ca}(\mathcal{B}), \|\mu\|)$ is a Banach space.

The space $\mathrm{rca}(\mathcal{B})$ is the subspace of $\mathrm{ca}(\mathcal{B})$ of all *regular* countably additive functions μ, i.e., it additionally holds that for all $\Gamma \in \mathcal{B}$ we have $\mu(\Gamma) = \sup\{\mu(K) \mid K \subset \Gamma,\ K \text{ compact}\}$ and $\mu(\Gamma) = \inf\{\mu(O) \mid \Gamma \subset O,\ O \text{ open}\}$. Note that, by taking complements, it is easily seen that the first condition actually implies the second.

2.2 Norming Dual Pairs and Weak Topologies

In certain applications of functional analysis the canonical weak topology seem to be too strong. Instead of the initial topology on X induced by the whole dual space X^* of a Banach space X we might want to consider a locally convex initial topology induced by a closed subspace $Y \subset X^*$. To generalize the weak topology to such cases the notion of norming dual pairs is used. We will give a definition of this and provide some examples that will proof especially useful for the study of transition operators. We follow a definition from Kunze [Kun11].

Definition 2.1. Let X and Y be a Banach spaces and $\langle \cdot, \cdot \rangle$ be a bilinear product on $X \times Y$. The triple $(X, Y, \langle \cdot, \cdot \rangle)$ is called *norming dual pair* if

$$\|x\|_X = \sup\{|\langle x, y\rangle| \mid y \in Y,\ \|y\|_Y \leq 1\}$$

and

$$\|y\|_Y = \sup\{|\langle x, y\rangle| \mid x \in X,\ \|x\|_X \leq 1\}.$$

For brevity the norming dual pairs are usually denoted by (X, Y). Note that if (X, Y) is a norming dual pair then (Y, X) can be made one by transposing the arguments of the bilinear product.

To every norming dual pair (X, Y) and every $y \in Y$ consider the semi-norm $p_y(x) = |\langle x, y\rangle|$, $x \in X$.

Definition 2.2. The locally convex topology on X generated by the family $\{p_y \mid y \in Y\}$ is called the *$\sigma(X, Y)$-topology.*

An important characterization of the $\sigma(X, Y)$-continuous functionals on X is the following, cf. Werner [Wer18, Corollary VIII.3.4].

Proposition 2.3. *A functional on X is $\sigma(X, Y)$-continuous if and only if it is of the form $x \mapsto \langle x, y\rangle$ for some $y \in Y$.*

A natural example, which shows that norming dual pairs indeed generalize the concept of duality, is the canonical duality.

Example 2.4. Is X a Banach space, then (X, X^*), together with the canonical duality pairing $\langle \cdot, \cdot \rangle$, is a norming dual pair.

Less immediate examples are provided in the following paragraphs.

The Duality between $B_b(E)$ and $\mathrm{ca}(\mathcal{B})$. Let $(E, \mathcal{B})$ be a measurable space. We define a bilinear form

$$\begin{aligned} \langle \cdot, \cdot \rangle &: B_b(E) \times \mathrm{ca}(\mathcal{B}) \to \mathbb{R}, \\ \langle f, \mu \rangle &= \int_E f(x)\, \mu(dx). \end{aligned} \tag{2.1}$$

This mapping satisfies for $f \in B_b(E)$, $\mu \in \mathrm{ca}(\mathcal{B})$

$$|\langle f, \mu\rangle| \leq \int_E |f|\, d|\mu| \leq \sup_{x\in E}|f(x)||\mu|(E) = \|f\|_{B_b}\|\mu\|_{\mathrm{ca}} \tag{2.2}$$

and

$$\|\langle f, \cdot\rangle\| = \sup_{\mu\in\mathrm{ca}} \frac{|\langle f,\mu\rangle|}{\|\mu\|} \geq \sup_{x\in E}\frac{|\langle f,\delta_x\rangle|}{\|\delta_x\|} = \sup_{x\in E}|f(x)| = \|f\|_{B_b}. \tag{2.3}$$

Similarly, letting A_μ^+ and A_μ^- be a Hahn-Decomposition of E with respect to μ and setting $f_\mu := 1_{A_\mu^+} - 1_{A_\mu^-}$, we obtain for all $\mu \in \mathrm{ca}(\mathcal{B})$ that

$$\|\langle \cdot, \mu\rangle\| = \sup_{f\in B_b(E)} \frac{|\langle f,\mu\rangle|}{\|f\|} \geq \frac{|\langle f_\mu,\mu\rangle|}{\|f_\mu\|} = |\mu|(E) = \|\mu\|_{\mathrm{ca}}. \tag{2.4}$$

Thus, the space $B_b(E)$ can be viewed as a subspace of the dual space $\mathrm{ca}(\mathcal{B})^*$ of $\mathrm{ca}(\mathcal{B})$, via the linear isometric embedding $B_b(E) \to \mathrm{ca}(\mathcal{B})^*$, $f \mapsto \langle f, \cdot\rangle$ and $\mathrm{ca}(\mathcal{B})$ can be viewed as a subspace of the dual space $B_b(E)^*$ of $B_b(E, \mathcal{B})$, via the linear isometric embedding $\mathrm{ca}(\mathcal{B}) \to B_b(E)^*$, $\mu \mapsto \langle \cdot, \mu\rangle$. Furthermore, the pair $(B_b(E), \mathrm{ca}(\mathcal{B}))$ defines a norming dual pair with pairing (2.1).

We may further specify the convergence in $B_b(E)$ with respect to the topology $\sigma(B_b(E), \mathrm{ca}(\mathcal{B}))$.

Lemma 2.5. *A sequence of functions $(f_n)_{n\in\mathbb{N}}$ in $B_b(E)$ converges in the $\sigma(B_b(E), \mathrm{ca}(\mathcal{B}))$-topology if and only if for each $x \in E$ the sequence $(f_n(x))_{n\in\mathbb{N}}$ converges in $\mathbb{R}$ and there is a constant $M > 0$ such that $\|f_n\| \leq M$ for all $n \in \mathbb{N}$.*

Proof. Assume $(f_n)_{n\in\mathbb{N}}$ converges in the $\sigma(B_b(E), \mathrm{ca}(\mathcal{B}))$-topology to some $f \in B_b(E)$. Then, in particular, for every $x \in E$ we obtain the convergence

$$|f_n(x) - f(x)| = \langle f_n - f, \delta_x\rangle \to 0,$$

as $n \to \infty$. The equiboundedness now follows from the fact that for each $\mu \in \mathrm{ca}(\mathcal{B})$ the sequence $(|\langle f_n, \mu\rangle|)_{n\in\mathbb{N}}$ is convergent and thus bounded. So, the uniform convergence principle implies the equiboundedness of $\|f_n\| = \sup\{|\langle f_n, \mu\rangle| \mid \mu \in \mathrm{ca}(\mathcal{B}),\ \|\mu\| \leq 1\}$.

Conversely, assume the pointwise convergence and equiboundedness. Denote by f the pointwise limit and note that $f \in B_b(E)$. Then an application of Lebesgue's dominated convergence theorem[1] yields the convergence of $|\langle f_n - f, \mu\rangle| \to 0$ for all $\mu \in \mathrm{ca}(\mathcal{B})$. ■

If we assume further structure on E we obtain the following technical lemma.

Lemma 2.6. *Let E be a separable metric space with Borel-σ-algebra $\mathcal{B}$. Then there exists a generator $\mathcal{E}$ of $\mathcal{B}$ that is closed under taking intersections and such that for each $B \in \mathcal{E}$ there exists a sequence $(f_n)_{n\in\mathbb{N}}$ in $C_0(E)$ such that*

$$\lim_{n\to\infty} f_n = 1_B$$

in the $\sigma(B_b(E), \mathrm{ca}(\mathcal{B}))$-topology.

[1] The dominated convergence theorem can be generalized to signed measures, e.g. by means of a Jordan decomposition of the signed measure and linearity.

Proof. Set $\mathcal{E} := \{\bigcap_{i=1}^m B_{r_i}(x_i) \subset E \mid r_i > 0,\, x_i \in E,\, i = 1,\dots,m,\, m \in \mathbb{N}\}$, where $B_r(x)$ is an open ball. As E is separable and metric, this set is a generator of the Borel-σ-algebra $\mathcal{B}$ that is stable under taking finite intersections. Moreover for each $r > 0$, $x \in E$ and $\theta \in (0,1)$ we find, by Urysohn's Lemma, a continuous function $f^{r,x,\theta}$ with $0 \leq f^{r,x,\theta} \leq 1$, $\operatorname{supp}(f^{r,x,\theta}) = \overline{B_r(x)}$ and $f^{r,x,\theta} \equiv 1$ on $B_{\theta r}(x)$. Now for each $C = \bigcap_{i=1}^m B_{r_i}(x_i) \in \mathcal{E}$ set $f^\theta := \min_{i=1,\dots,m} f^{r_i,x_i,\theta}$ then $f^\theta \in C(E)$ and for $\theta \to 1$ it converges pointwise to 1_C. Furthermore, $\|f^\theta\| \leq 1$ for each $\theta \in (0,1)$. The assertion now follows from Lemma 2.5. ∎

The Duality between $C_0(E)$ and $\mathrm{rca}(\mathcal{B})$ for locally compact Hausdorff E. Let E be a locally compact Hausdorff space with Borel-σ-algebra $\mathcal{B}$. Let $C_0(E)$ and $\mathrm{ca}(\mathcal{B})$ be as in the previous paragraphs.

It can be shown as in the previous paragraph that the pair $(C_0(E), \mathrm{rca}(\mathcal{B}))$ is a norming dual pair with duality pairing (2.1). So, the space $C_0(E)$ can be viewed as a subspace of the dual space $\mathrm{rca}(\mathcal{B})^*$ of $\mathrm{rca}(\mathcal{B})$, via the linear isometric embedding $C_0(E) \to \mathrm{rca}(\mathcal{B})^*$, $f \mapsto \langle f, \cdot \rangle$.

For the spaces $\mathrm{rca}(\mathcal{B})$ and $C_0(E)^*$ we can provide a much stronger result. This result is known as Riesz representation theorem.

Theorem 2.7 (Riesz representation theorem)**.** *Let E be a locally compact Hausdorff space. Then the map*

$$\Phi : \mathrm{rca}(\mathcal{B}) \to C_0(E)^*,$$
$$\Phi(\mu)(f) \mapsto \int_E f\, d\mu, \quad f \in C_0(E)$$

is an order-preserving isometric isomorphism.

The proof of this result is not trivial and would fill a paper on its own. A detailed treatment of this and similar results can be found in Elstrodt [Els05, Chapter VIII, Theorem 2.26].

2.3 Regularization of Supermartingales

A central role in the regularization of a Feller process is played by supermartingales. We will quote two important regularity results for supermartingales, which we will be using later to treat Feller processes.

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $\mathcal{F}$ a filtration in $\mathcal{A}$.

Definition 2.8 (martingale)**.** A real values process $(X_t)_{t\geq 0}$ is called *$\mathcal{F}$-(sub-/super-) martingale* if for all $t \geq 0$ we have $\mathbb{E}[|X_t|] < \infty$ and for all $s, t \geq 0$

$$\mathbb{E}[X_{t+s} \mid \mathcal{F}_s] = X_s, \quad \mathbb{P}\text{-a.s.},$$

with $=$ replaced by $\geq$ resp. $\leq$ in the sub- resp. supermartingale case.

Recall that a process is *càdlàg* if its paths are right-continuous and have left limits. For any process $Y = (Y_t)_{t\in\mathbb{Q}_+}$ we set the process $Y^+ = (Y_{t+})_{t\in\mathbb{R}_+}$ to be the process of the right-hand limits of Y along $\mathbb{Q}_+$ provided that they exist. The next regularization property is from Kallenberg [Kal02, Theorem 2.27].

Theorem 2.9 (regularization, Doob). *Let $X = (X_t)_{t\in\mathbb{R}_+}$ be a $\mathcal{F}$-supermartingale with restriction Y to $\mathbb{Q}_+$. Then Y^+ exists and is càdlàg outside some fixed $\mathbb{P}$-null set.*

For supermartingales the paths reaching 0 are almost surely absorbing. More, precisely we have from Kallenberg [Kal02, Lemma 7.31]:

Lemma 2.10 (absorption). *Let $X \geq 0$ be a right-continuous supermartingale, and put $\tau := \inf\{t \geq 0 \mid X_t \wedge X_{t-} = 0\}$. Then $\mathbb{P}$-a.s. we have $X = 0$ on $[\tau, \infty)$.*

3 Transition Semigroups and Markov Processes

In this chapter we give a basic introduction to the theory of Markov processes and some related notions. Our objective is to obtain an understanding on how Markov processes correspond to semigroups.

3.1 Transition functions and Transition Semigroups

We introduce the notions of Markov kernels, transition functions, operators and semigroups.

Markov kernels and transition functions.

Definition 3.1 (Markov kernel). Let $(E, \mathcal{B})$ be a measurable space. A map $p : E \times \mathcal{B} \to [0, 1]$ is called *(sub-) Markov kernel on* $(E, \mathcal{B})$ if

(i) for each $x \in E$ the map $p(x, \cdot)$ is a (sub-) probability measure on the σ-algebra $\mathcal{B}$;

(ii) for each $\Gamma \in \mathcal{B}$ the map $p(\cdot, \Gamma)$ is $\mathcal{B}$-measurable.

Definition 3.2 (transition function). Let $(E, \mathcal{B})$ be a measurable space. A family of sub-Markov kernels $(p_t)_{t\geq 0}$ is called *transition function* if

(i) for each $x \in E$ we have $p_0(x, \cdot) = \delta_x$;

(ii) for each $s, t \geq 0$ the Chapman-Kolmogorov equation

$$p_s p_t = p_{s+t} \tag{3.1}$$

holds, where the product of two kernels p and q on $(E, \mathcal{B})$ is given by

$$(pq)(x, \Gamma) = \int_E p(x, dy) q(y, \Gamma), \quad x \in E, \, \Gamma \in \mathcal{B}.$$

We will call a transition function *conservative*, if p_t is a Markov kernel for each $t \geq 0$.

Transition operators corresponding to Markov kernels. Let p be a sub-Markov kernel on a measurable space $(E, \mathcal{B})$. We define the transition operator

$$T : B_b(E) \to B_b(E), \quad Tf(x) = (Tf)(x) = \int_E f(y)p(x, dy), \quad x \in E \tag{3.2}$$

and the Perron-Frobenius operator

$$U : \mathrm{ca}(\mathcal{B}) \to \mathrm{ca}(\mathcal{B}), \quad U\mu(\Gamma) = (U\mu)(\Gamma) = \int_E p(x, \Gamma)\mu(dx), \quad \Gamma \in \mathcal{B}. \tag{3.3}$$

Note that both operators are linear. Approximating f by simple functions we see that Tf is again measurable. It is easily seen that T is a *contraction* in the sense that it satisfies $\|Tf\|_{B_b} \le \|f\|_{B_b}$ and is thus well defined. Using the Jordan decomposition $\mu = \mu^+ - \mu^-$ we see from monotonous convergence (or dominated convergence) that $T\mu$ is again a countably additive function on $\mathcal{B}$. Also U is a contraction. Indeed,

$$\begin{aligned}
\|U\mu\|_{\mathrm{ca}} &= \sup_{\Gamma \in \mathcal{B}} U\mu(\Gamma) - \inf_{\Gamma \in \mathcal{B}} U\mu(\Gamma) \\
&\le \sup_{\Gamma \in \mathcal{B}} \int_E p(x, \Gamma)\mu^+(dx) + \sup_{\Gamma \in \mathcal{B}} \int_E p(x, \Gamma)\mu^-(dx) \\
&\le \mu^+(E) + \mu^-(E) = \|\mu\|_{\mathrm{ca}}.
\end{aligned}$$

Note also that T and U are *positive operators*, i.e., $f \in B_b(E, \mathcal{B})$, $f \ge 0$ and $\mu \in \mathrm{ca}(E, \mathcal{B})$, $\mu \ge 0$ implies $Tf \ge 0$ and $U\mu \ge 0$, respectively.

Recall from Section 2.2, that $(B_b(E), \mathrm{ca}(\mathcal{B}))$ is a norming dual pair. This duality yields that the transition operator and the Perron-Frobenius operator can be viewed as partial adjoint operators in the following sense: for each $f \in B_b(E)$ and $\mu \in \mathrm{ca}(\mathcal{B})$ we have

$$\langle Tf, \mu \rangle = \int_E \int_E f(y)\, p(x, dy)\, \mu(dx) = \langle f, U\mu \rangle,$$

implying that the adjoint T^* of T extends U and, conversely, the adjoint U^* of U extends T.

One consequence of this is that the transition operator is continuous with respect to the locally convex topology $\sigma(B_b(E), \mathrm{ca}(\mathcal{B}))$ on $B_b(E)$. Indeed, for each $\mu \in \mathrm{ca}(\mathcal{B})$ the map

$$B_b(E) \ni f \mapsto \langle Tf, \mu \rangle = \langle f, U\mu \rangle$$

is $\sigma(B_b(E), \mathrm{ca}(\mathcal{B}))$-continuous, and so also T is $\sigma(B_b(E), \mathrm{ca}(\mathcal{B}))$-continuous.

Summarizing the previous discussion we obtain that a transition operator is a linear, positive, contraction, which is also $\sigma(B_b(E), \mathrm{ca}(\mathcal{B}))$-continuous. We find that these properties are enough to fully characterize the transition operators.

Proposition 3.3. *Let T be a linear, positive contraction on $B_b(E)$ that is $\sigma(B_b(E), \mathrm{ca}(\mathcal{B}))$-continuous. Then*

$$p(x, \Gamma) := (T1_\Gamma)(x), \quad x \in E, \Gamma \in \mathcal{B} \tag{3.4}$$

defines a sub-Markov kernel. In particular, the formulas (3.2) *and* (3.4) *define a one-to-one correspondence between the linear, positive contractions on $B_b(E)$ that are $\sigma(B_b(E), \mathrm{ca}(\mathcal{B}))$-continuous and the sub-Markov kernels.*

Proof. Let $x \in E$ and $\Gamma \in \mathcal{B}$. By contractivity of T we obtain that $p(x, \Gamma) \in [0, 1]$. Moreover, we have $p(\cdot, \Gamma) = T1_\Gamma \in B_b(E)$, so it is measurable. Consider a sequence $\Gamma_1, \Gamma_2, \dots \in \mathcal{B}$ of pairwise disjoint sets, set $\Gamma := \bigcup_{n\in\mathbb{N}} \Gamma_n$ and define $f_n := \sum_{i=1}^n 1_{\Gamma_i}$. Then pointwise $f_n \to 1_\Gamma$ and $\|f_n\| \le 1$. So, Lemma 2.5 implies that f_n converges to 1_Γ in the $\sigma(B_b(E), \mathrm{ca}(\mathcal{B}))$-topology. Thus, the linearity and $\sigma(B_b(E), \mathrm{ca}(\mathcal{B}))$-continuity of T yield

$$\sum_{i\in\mathbb{N}} p(x, \Gamma_i) = \lim_{n\to\infty} (Tf_n)(x) = (T1_\Gamma)(x) = p(x, \Gamma).$$

So $p(x, \cdot)$ is a σ-additive sub-probability measure. It is straightforward to check that (3.2) and (3.4) define the claimed one-to-one correspondence. ■

Transiton semigroups corresponding to transitions functions. Given a family of sub-Markov kernels $(p_t)_{t\ge0}$. Using (3.2), we can now associate a family of transition operators $(T_t)_{t\ge0}$ to it. We investigate the case in which $(p_t)_{t\ge0}$ is a transition function.

A family of operators $(S_t)_{t\ge0}$ is called the *semigroup (of operators)* if $S_{t+s} = S_s S_t$ for all $s, t \ge 0$ and S_0 is the identity operator.

Lemma 3.4. *A family of sub-Markov kernels $(p_t)_{t\ge0}$ on $(E, \mathcal{B})$ is a transition function iff the corresponding family of transition operators $(T_t)_{t\ge0}$ is a semigroup of contractions.*

Proof. For $x \in E$, $\Gamma \in \mathcal{B}$ and $s, t \ge 0$ we have $T_{s+t}1_\Gamma(x) = p_{s+t}(x, \Gamma)$ and

$$\begin{aligned}(T_sT_t)1_\Gamma(x) = T_s(T_t1_\Gamma(x)) &= \int_E (T_t1_\Gamma)(y)p_s(x, dy)\\ &= \int_E p_t(y, \Gamma)p_s(x, dy) = (p_sp_t)(x, \Gamma).\end{aligned}$$

This implies that the Chapman-Kolmogorov relation (3.1) is equivalent to $T_{s+t}1_\Gamma = (T_sT_t)1_\Gamma$ for every $\Gamma \in \mathcal{S}$. The latter relation extends to every bounded $\mathcal{B}$-measurable function by linearity and monotone convergence.

Now note that T in (3.2) is the identity operators iff $p(x, \cdot) = \delta_x$ for all $x \in E$. ■

This motivates the following definition.

Definition 3.5 (Transition semigroup)**.** Let $(p_t)_{t\ge0}$ be a transition function. The family $(T_t)_{t\ge0}$ of operators $T_t : B_b(E, \mathcal{B}) \to B_b(E, \mathcal{B})$, $t \ge 0$, defined by

$$T_tf(x) := \int_E f(y)p_t(x, dy), \quad f \in B_b(E, \mathcal{B}),\ x \in E, \tag{3.5}$$

is called *transition semigroup for $(p_t)_{t\ge0}$*.

Let $(T_t)_{t\ge0}$ be a semigroup of positive linear contractions on $B_b(E, \mathcal{B})$ that are $\sigma(B_b(E), \mathrm{ca}(\mathcal{B}))$-continuous. Define a family of sub-Markov kernels $(p_t)_{t\ge0}$ on $(E, \mathcal{B})$ by

$$p_t(x, \Gamma) := (T_t1_\Gamma)(x), \quad x \in E,\ \Gamma \in \mathcal{B}. \tag{3.6}$$

As an immediate consequence of Proposition 3.3 and Lemma 3.4 we obtain the following.

Corollary 3.6. *The semigroups of positive linear contractions on $B_b(E, \mathcal{B})$ that are $\sigma(B_b(E), \mathrm{ca}(\mathcal{B}))$-continuous are in one-to-one correspondence with the transition functions on $(E, \mathcal{B})$ via the equations* (3.5) *and* (3.6)*.*

In light of this we may generalize the definition of a transition semigroup to:

Definition 3.7. A *transition semigroup* is a semigroup of positive linear contractions on $B_b(E, \mathcal{B})$ that are $\sigma(B_b(E), \mathrm{ca}(\mathcal{B}))$-continuous.

Conservativity. One important reason why we treat transition semigroups is that they naturally correspond to Markov processes. An important condition for this is that of conservativity. Recall that a transition function $(p_t)_{t\geq 0}$ is called conservative if it is a family of Markov kernels.

Definition 3.8. The transition semigroup $(T_t)_{t\geq 0}$ is called *conservative* if its corresponding transition function is conservative.

Lemma 3.9. *A transition semigroup $(T_t)_{t\geq 0}$ is conservative iff $T_t 1 = 1$ for all $t \geq 0$.*

Proof. Let $(p_t)_{t\geq 0}$ be the corresponding transition function. Then the statement follows from $(T_t 1)(x) = p_t(x, E)$ for all $x \in E$. ■

3.2 Markov Processes

We follow a definition of Markov processes given in Dynkin [Dyn65].

Definition 3.10 (Markov process)**.** Suppose we are given:

(i) a measurable space $(E, \mathcal{B})$ called *state space*,

(ii) a measurable space $(\Omega, \mathcal{A})$,

(iii) a filtration $\mathcal{F} = (\mathcal{F}_t)_{t\geq 0}$ in $\mathcal{A}$,

(iv) a process $X : \mathbb{R}_+ \times \Omega \to E$, $X(t, \omega) = X_t(\omega)$,

(v) for each $x \in E$ a probability measure $\mathbb{P}^x$ on $\mathcal{A}$.

We say that the triple $((X_t)_{t\geq 0}, \mathcal{F}, \{\mathbb{P}^x\}_{x\in E})$ is a *Markov process* if

(M_1) the process $(X_t)_{t\geq 0}$ is $\mathcal{F}$-adapted,

(M_2) for all $A \in \mathcal{A}$ the mapping $E \ni x \mapsto \mathbb{P}^x(A)$ is $\mathcal{B}$-$\mathcal{B}([0, 1])$-measurable,

(M_3) for all $x \in E$ we have $\mathbb{P}^x X_0^{-1} = \delta_x$,

(M_4) for all $s, t \geq 0$, $x \in E$ and $\Gamma \in \mathcal{B}$ we have

$$\mathbb{P}^x(\{X_{s+t} \in \Gamma\} \mid \mathcal{F}_s) = \mathbb{P}^{X_s}(\{X_t \in \Gamma\}), \quad \mathbb{P}^x\text{-a.s.},$$

(M_5) for each $t \geq 0$ and $\omega \in \Omega$ we find an $\omega' \in \Omega$ such that

$$X_{s+t}(\omega) = X_s(\omega'), \quad s \geq 0.$$

For brevity we usually write "X is a $\mathcal{F}$-Markov process" instead of the triple. If the filtration $\mathcal{F}$ is clear, irrelevant or is the filtration generated by X, we omit it. Also note that it follows from the tower property of the conditional expectation that any $\mathcal{F}$-Markov process is always a Markov process w.r.t. its canonical filtration $\mathcal{F}^X$. By $\mathcal{F}_\infty$ we denote the smallest σ-algebra containing the filtration $\mathcal{F}$.

Remark 3.11.

(i) We may replace property (M_4) by the two conditions

($M_{4,1}$) for all $s, t \geq 0$, $x \in E$ and $\Gamma \in \mathcal{B}$ we have

$$\mathbb{P}^x(\{X_{s+t} \in \Gamma\} \mid \mathcal{F}_s) = \mathbb{P}^x(\{X_{s+t} \in \Gamma\} \mid X_s), \quad \mathbb{P}^x\text{-a.s.},$$

($M_{4,2}$) for all $s, t \geq 0$, $x \in E$ and $\Gamma \in \mathcal{B}$ we have

$$\mathbb{P}^x(\{X_{s+t} \in \Gamma\} \mid X_s = y) = \mathbb{P}^y(\{X_t \in \Gamma\}), \quad \text{for } \mathbb{P}^x X_s^{-1}\text{-a.e. } y \in E,$$

Indeed, assuming (M_4) we obtain from the $\sigma(X_s)$-$\mathcal{B}([0,1])$-measurability of $\mathbb{P}^{X_s}(\{X_{s+t} \in \Gamma\})$ that there is a $\sigma(X_s)$-$\mathcal{B}([0,1])$-measurable version of $\mathbb{P}^x(\{X_{s+t} \in \Gamma\} \mid \mathcal{F}_s)$. So $\mathbb{P}^x$-a.s.

$$\begin{aligned}\mathbb{P}^x(\{X_{s+t} \in \Gamma\} \mid \mathcal{F}_s) &= \mathbb{P}^x(\mathbb{P}^x(\{X_{s+t} \in \Gamma\} \mid \mathcal{F}_s) \mid X_s) \\ &= \mathbb{P}^x(\{X_{s+t} \in \Gamma\} \mid X_s) = \mathbb{P}^{X_s}(\{X_{s+t} \in \Gamma\}),\end{aligned}$$

and thus, factoring by X_s yields ($M_{4,2}$). Conversely, assume that ($M_{4,1}$) and ($M_{4,2}$) hold. Then $\mathbb{P}^x$-a.s.

$$\begin{aligned}\mathbb{P}^x(\{X_{s+t} \in \Gamma\} \mid \mathcal{F}_s) &= \mathbb{P}^x(\{X_{s+t} \in \Gamma\} \mid X_s) \\ &= \mathbb{P}^x(\{X_{s+t} \in \Gamma\} \mid X_s = \cdot) \circ X_s = \mathbb{P}^{X_s}(\{X_{s+t} \in \Gamma\}).\end{aligned}$$

(ii) The property (M_5) implies that we may construct a family of $\mathcal{F}^X_\infty$-$\mathcal{F}^X_\infty$-measurable *shift operators* $(\theta_t)_{t \geq 0}$ such that $\theta_t : \Omega \to \Omega$ satisfies

$$X_{s+t}(\omega) = X_s(\theta_t \omega), \quad \omega \in \Omega,\ s, t \geq 0. \tag{3.7}$$

Thus, this property requires the set of trajectories of the Markov process to be invariant under all shifts. If not already so, a triple that satisfies all conditions but (M_5) can be modified by enlarging the measurable space $(\Omega, \mathcal{A})$ such that this invariance is met. This is illustrated in a footnote of Dynkin [Dyn65, p. 79].

It is sometimes conceptually helpful to not have a whole family of probability measures but only a single probability space. We therefore introduce the Markov random functions.

Definition 3.12 (Markov random function)**.** Suppose we are given items (i) to (iv) of Definition 3.10 and a probability measure $\mathbb{P}$ on $\mathcal{A}$. We say that the triple $((X_t)_{t\geq 0}, \mathcal{F}, \mathbb{P})$ is a *Markov random function* if conditions (M_1) and (M_5) in Definition 3.10 are fulfilled and we have for all $s,t \geq 0$, $x \in E$ and $\Gamma \in \mathcal{B}$ that

$$\mathbb{P}(\{X_{s+t} \in \Gamma\} \mid \mathcal{F}_s) = \mathbb{P}(\{X_{s+t} \in \Gamma\} \mid X_s), \quad \mathbb{P}\text{-a.s.}, \tag{3.8}$$

The pushforward measure $\nu := \mathbb{P}X_0^{-1}$ is called *initial distribution* of the Markov random function.

If we want to emphasize that a Markov random function has an given initial distribution ν we sometimes write $\mathbb{P}^\nu$ instead of $\mathbb{P}$. The next proposition exploits a connection of Markov processes and Markov random functions.

Proposition 3.13 (mixtures)**.** *Let X be a Markov process and ν be a probability measure on $\mathcal{B}$. Then the triple $((X_t)_{t\geq 0}, \mathcal{F}, \mathbb{P}^\nu)$, where*

$$\mathbb{P}^\nu(A) = \int_E \mathbb{P}^x(A)\,\nu(dx), \quad A \in \mathcal{A}, \tag{3.9}$$

is a probability measure, defines a Markov random function with initial distribution ν. Moreover, we have for $t \geq 0$, $\Gamma \in \mathcal{B}$ that

$$\mathbb{P}^\nu(\{X_{t+s} \in \Gamma\} \mid \mathcal{F}_s) = \mathbb{P}^{X_s}(\{X_t \in \Gamma\}), \quad \mathbb{P}^\nu\text{-a.s.} \tag{3.10}$$

Proof. It is easily checked that $\mathbb{P}^\nu X_0^{-1} = \nu$. Note that to show the Markov property (3.8) it suffices to proof equation (3.10). Let $s,t \geq 0$, $\Gamma \in \mathcal{B}$ and $A \in \mathcal{F}_s$. We obtain by definition of $\mathbb{P}^\nu$ and from property (M_4) that

$$\begin{aligned}\mathbb{P}^\nu(A \cap \{X_{s+t} \in \Gamma\}) &= \int_E \mathbb{P}^x(A \cap \{X_{s+t} \in \Gamma\})\,\nu(dx) = \int_E \int_A 1_{\{X_{s+t}\in\Gamma\}}\,d\mathbb{P}^x\,\nu(dx)\\ &= \int_E \int_A \mathbb{P}^{X_s}(\{X_t \in \Gamma\})\,d\mathbb{P}^x\,\nu(dx) = \int_A \mathbb{P}^{X_s}(\{X_t \in \Gamma\})\,d\mathbb{P}^\nu.\end{aligned}$$

Thus, $\mathbb{P}^\nu(\{X_{t+s} \in \Gamma\} \mid \mathcal{F}_s) = \mathbb{P}^{X_s}(\{X_t \in \Gamma\})$, $\mathbb{P}^\nu$-a.s. ∎

As a consequence of Proposition 3.13 we can not only associate Dirac initial distributions to Markov processes but may give it arbitrary *initial distribution* ν.

In the next result we extend the property given in equation (3.10) to more general sets than $\{X_{s+t} \in \Gamma\}$. A proof of this can be found in Dynkin [Dyn61, §2, Theorem 2.2].

Proposition 3.14 (extended Markov property)**.** *Let X be a Markov process and let ν be some probability measure on $\mathcal{B}$. Then for $B \in \mathcal{B}^{\mathbb{R}_+}$ and each $t \geq 0$ we have*

$$\mathbb{P}^\nu(\{X \circ \theta_t \in B\} \mid \mathcal{F}_t) = \mathbb{P}^{X_t}(\{X \in B\}), \quad \mathbb{P}^\nu\text{-a.s.} \tag{3.11}$$

and

$$\mathbb{E}^\nu\left[\xi \circ \theta_t \mid \mathcal{F}_t\right] = \mathbb{E}^{X_t}\left[\xi\right], \quad \mathbb{P}^\nu\text{-a.s.},$$

for any bounded or nonnegative $\mathcal{F}_\infty^X$-$\mathcal{B}(\mathbb{R})$-measurable $\xi : \Omega \to \mathbb{R}$.

Markov processes and corresponding transition functions. We say that a Markov process X corresponds to a transition function $(p_t)_{t\geq 0}$, if for all $x \in E$

$$p_t(X_s, \Gamma) = \mathbb{P}^x(\{X_{s+t} \in \Gamma\} \mid X_s) = \mathbb{P}^x(\{X_{s+t} \in \Gamma\} \mid \mathcal{F}_s), \quad \mathbb{P}^x\text{-a.s.}, \Gamma \in \mathcal{B}. \tag{3.12}$$

The next result states that any Markov process corresponds to a transition function.

Proposition 3.15. *Let X be a Markov process. Then the family of functions $(p_t)_{t\geq 0}$ defined by*

$$p_t(x, \Gamma) = \mathbb{P}^x(\{X_t \in \Gamma\}), \quad t \geq 0,\, x \in E,\, \Gamma \in \mathcal{B}$$

is a conservative transition function for which equation (3.12) *holds.*

Proof. By condition (M_2) it is clear that for each $t \geq 0$ the mapping $p_t : E \times \mathcal{B} \to [0, 1]$ is a Markov kernel. So it remains to show that the Chapman-Kolmogorov equation holds. For $s, t \geq 0$, $x \in E$ and $\Gamma \in \mathcal{B}$ we infer from property ($M_{4,2}$), the transformation theorem, cf. Theorem 1.20.2 in Gänssler and Stute [GS77], and the law of total expectation that

$$\begin{aligned}
(p_t p_s)(x, \Gamma) &= \int_E p_s(y, \Gamma)\, p_t(x, dy) = \int_E \mathbb{P}^y(\{X_s \in \Gamma\})\, \mathbb{P}^x(\{X_s \in dy\}) \\
&= \int_E \mathbb{P}^x(\{X_{s+t} \in \Gamma\} \mid X_s = y)\, \mathbb{P}^x(\{X_s \in dy\}) = \int_\Omega \mathbb{P}^x(\{X_{s+t} \in \Gamma\} \mid X_s)\, d\mathbb{P}^x \\
&= \mathbb{P}^x(\{X_{s+t} \in \Gamma\}) = p_{s+t}(x, \Gamma).
\end{aligned}$$

That equation (3.12) holds follows immediately from property (M_4). ■

Representation of the Transition Semigroup as a Koopman Semigroup. In the last paragraph we have seen how we can associate a conservative transition function, and so also a transition semigroup, to a Markov process. We want to exploit this connection to give an alternative representation of transition semigroups as a stochastic Koopman type semigroup.

Proposition 3.16. *Let X be a Markov process and let $(T_t)_{t\geq 0}$ be its corresponding transition semigroup. Let $t \geq 0$, $x \in E$ and $f \in B_b(E)$. Then we have*

$$T_t f(x) = \mathbb{E}^x \left[f(X_t)\right].$$

Proof. Denote by $(p_t)_{t\geq 0}$ the corresponding transition function. Let $s, t \geq 0$, $x \in E$ and $\Gamma \in \mathcal{B}$ and choose the initial distribution $\nu = \delta_x$. Then it follows from the correspondence (3.12) and the Markov property (M_4) that

$$p_t(X_0, \Gamma) = \mathbb{P}^{X_0}(\{X_t \in \Gamma\}), \quad \mathbb{P}^x\text{-a.s.},$$

and so after factoring by X_0

$$p_t(y, \Gamma) = \mathbb{P}^y(\{X_t \in \Gamma\}), \quad \mathbb{P}^x X_0^{-1}\text{-a.e. } y \in E,$$

which implies for $\mathbb{P}^x X_0^{-1}$-a.e. $y \in E$

$$T_t 1_\Gamma(y) = p_t(y, \Gamma) = \mathbb{P}^y(\{X_t \in \Gamma\}) = \mathbb{E}^y \left[1_{\{X_t \in \Gamma\}}\right] = \mathbb{E}^y \left[1_\Gamma(X_t)\right].$$

Note that $\mathbb{P}^x X_0^{-1} = \delta_x$. As x lies outside of every null set of δ_x it follows that $T_t 1_\Gamma(x) = \mathbb{E}^x \left[1_\Gamma(X_t)\right]$. Now the assertion is deduced using a monotone class type argument. ■

3.3 The Canonical Markov Process

Assuming that the space E is Polish, we may construct a Markov process corresponding to a transition function. We will first illustrate the construction for Markov random functions and will later extend this to Markov processes.

The canonical Markov random function. Set $I_n := \{(t_1, \dots, t_n) \in \mathbb{R}_+^n \mid t_1 \leq \dots \leq t_n\}$ and $I := \bigcup_{n \in \mathbb{N}} I_n$. For $n \in \mathbb{N}$, let $\{\mu_t\}_{t \in I_n}$ be a family of probability measure on $\mathcal{B}^n$. We call the family $\{\mu_t\}_{t \in I}$ *projective*, if for all $n \in \mathbb{N}_{\geq 2}$, $(t_1, \dots, t_n) \in I_n$, $i \in \{1, \dots, n\}$ and $B \in \mathcal{B}^{n-1}$ it holds that

$$\mu_{t_1, \dots, t_n}(B_i) = \mu_{t_1, \dots, t_{i-1}, t_{i+1}, \dots, t_n}(B),$$

where $B_i := \{(x_1, \dots, x_n) \in E^n \mid (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) \in B\}$.

Theorem 3.17 (existence of Markov random function)**.** *Let E be a Polish space and let $(p_t)_{t \geq 0}$ be a conservative transition function on $(E, \mathcal{B})$. Then there exists a Markov random function $X^\nu = ((X_t^\nu)_{t \geq 0}, \mathcal{F}, \mathbb{P}^\nu)$ with initial distribution ν, which satisfies for all $s, t \geq 0$ and $\Gamma \in \mathcal{B}$ the property*

$$\mathbb{P}^\nu(\{X_{s+t} \in \Gamma\} \mid \mathcal{F}_s) = p_t(X_s, \Gamma), \quad \mathbb{P}^\nu\text{-a.s.} \tag{3.13}$$

Proof. Choose $\Omega = E^{\mathbb{R}_+}$, $\mathcal{A} = \mathcal{B}^{\mathbb{R}_+}$, where the latter denotes the product-σ-algebra, and X as the identity on $(\Omega, \mathcal{A})$. Note that then $X_t : \Omega \to E$, $\omega \mapsto \omega(t)$ is measurable, as it is a projection. Moreover take for $\mathcal{F}$ the filtration generated by X.

We construct for each $(t_1, \dots, t_n) \in I$ a probability measure $\mu_{t_1, \dots, t_n}$ on $\mathcal{B}^n$ such that on the cuboid sets $B_0 \times B_1 \times \dots \times B_n$, where $B_0, \dots, B_n \in \mathcal{B}$, we have

$$\begin{aligned}
&\mu_{0, t_1, \dots, t_n}(B_0 \times B_1 \times \dots \times B_n) \\
&= \int_{B_0} \int_{B_1} \cdots \int_{B_n} 1 \, p_{t_n - t_{n-1}}(x_{n-1}, dx_n) \dots p_{t_2 - t_1}(x_1, dx_2) \, p_{t_1}(x_0, dx_1) \, \nu(dx_0)
\end{aligned}$$

and

$$\mu_{t_1, \dots, t_n}(B_1 \times \dots \times B_n) = \mu_{0, t_1, \dots, t_n}(E \times B_1 \times \dots \times B_n).$$

It follows from the Chapman-Kolmogorov equation that this family is projective. Hence, by Kolmogorov's existence theorem there exists a measure $\mathbb{P}^\nu$ on $\mathcal{A}$ with finite dimensional marginal distributions $\{\mu_t\}_{t \in I}$.

It is easily checked that $\mathbb{P}^\nu X_0^{-1} = \nu$. Also properties ($\mathrm{M}_1$) and ($\mathrm{M}_5$) are fulfilled by construction. To see that X is a Markov random function with initial distribution ν we show (3.13) from which (3.8) follows by the tower property of the conditional expectation. Fix $(s_1, \dots, s_n, s_{n+1}) \in I$ and set $B \in \mathcal{B}^n$ and $C \in \mathcal{B}$. We obtain from the fact that X is the identity on Ω that

$$\begin{aligned}
&\mathbb{P}^\nu(\{(X_{t_1}, \dots, X_{s_n}, X_{s_{n+1}}) \in B \times C\}) = \mu_{s_1, \dots s_n, s_{n+1}}(B \times C) \\
&= \int_B p_{s_{n+1} - s_n}(x_n, C) \, \mu_{s_1, \dots, s_n}(dx) = \int_{\{(X_{t_1}, \dots, X_{s_n}) \in B\}} p_{s_{n+1} - s_n}(X_{s_n}, C) \, d\mathbb{P}^\nu,
\end{aligned}$$

and so we obtain by a monotone class argument that for all $A \in \mathcal{F}_{s_n}$

$$\int_A 1_{\{X_{s_{n+1}} \in C\}}\, d\mathbb{P}^\nu = \int_A p_{s_{n+1}-s_n}(X_{s_n}, C)\, d\mathbb{P}^\nu.$$

Thus, $\mathbb{P}^\nu(\{X_{s_{n+1}} \in C\} \mid \mathcal{F}_{s_n}) = p_{s_{n+1}-s_n}(X_{s_n}, C)$, $\mathbb{P}^\nu$-a.s. ■

The canonical Markov Process. We will now argue how the canonical construction of a Markov random function can be used to construct a Markov process corresponding to a transition function $(p_t)_{t\geq 0}$.

First note that from Theorem 3.17 we get for each $x \in E$ a Markov random function $((X_t^x)_{t\geq 0}, \mathcal{F}^x, \mathbb{P}^x)$[2]. Also note that by construction neither the processes X^x nor the filtrations $\mathcal{F}^x$ depend on x, as the former were chosen to be identity mappings in $E^{\mathbb{R}_+}$ and the latter were chosen to be the natural filtrations. For families of this type we obtain the following result.

Theorem 3.18 (existence of Markov process)**.** *Let $(\Omega, \mathcal{A})$, $(E, \mathcal{B})$ measurable spaces. Assume that Ω is a subspace of the function space $E^{\mathbb{R}_+}$ and $\mathcal{A}$ is the trace-σ-algebra of $\mathcal{B}^{\mathbb{R}_+}$. Moreover, let $(p_t)_{t\geq 0}$ be a conservative transition function on $(E, \mathcal{B})$ and suppose we are given a family of Markov random functions $\{(X, \mathcal{F}, \mathbb{P}^x)\}_{x\in E}$ satisfying for all $s, t \geq 0$ and $\Gamma \in \mathcal{B}$*

$$\mathbb{P}^x(\{X_{s+t} \in \Gamma\} \mid \mathcal{F}_s) = p_t(X_s, \Gamma), \quad \mathbb{P}^x\text{-a.s., } x \in E, \tag{3.14}$$

where $X = id_\Omega$ and $\mathcal{F}$ is the canonical filtration of X. Then the triple $((X_t)_{t\geq 0}, \mathcal{F}, \{\mathbb{P}^x\}_{x\in E})$ is a Markov process corresponding to the transition function $(p_t)_{t\geq 0}$.

In particular, if E is Polish and $(\Omega, \mathcal{A}) = (E^{\mathbb{R}_+}, \mathcal{B}^{\mathbb{R}_+})$, we can associate a Markov process to each conservative transition function $(p_t)_{t\geq 0}$ on $(E, \mathcal{B})$.

In order to construct a Markov process from a family of Markov random functions we generally do not require that the processes do not depend on x or that the filtration is canonical. For a more general construction we refer the reader to Dynkin [Dyn65, §1, Section 3.2]. The benefit of these assumption is that the construction becomes much simpler.

Proof. Immediately, the properties (M_1), (M_3) and (M_5) are clear. So it remains to show property (M_2) and (M_4).

(M_2): First note that $\mathcal{F}_\infty = \mathcal{A}$, as the process X was chosen to be the identity on Ω. For $0 < t_1 \leq \cdots \leq t_n$, $\Gamma_0, \ldots, \Gamma_n \in \mathcal{B}$ and $n \in \mathbb{N}$ the measurability in x for cylinder sets follows from the representation

$$\mathbb{P}^x(\{X_0 \in \Gamma_0, X_{t_1} \in \Gamma_1 \ldots, X_{t_n} \in \Gamma_n\}) = \int_{\Gamma_0} f(x_0)\delta_x(dx_0) = 1_{\Gamma_0}(x)f(x), \tag{3.15}$$

[2] For dirac initial distributions it is customary to replace the superscript δ_x by x.

where

$$f : E \to \mathbb{R},\ x_0 \mapsto \int_{\Gamma_1} \cdots \int_{\Gamma_n} 1\, p_{t_n - t_{n-1}}(x_{n-1}, dx_n) \ldots p_{t_2 - t_1}(x_1, dx_2) p_{t_1}(x_0, dx_1)$$

is well-defined and by Lemma 1.8.7 in Gänssler and Stute [GS77] it is $\mathcal{B}$-$\mathcal{B}([0,1])$-measurable. Note that the representation in equation (3.15) follows by induction from the Chapman-Kolmogorov equation (3.1) and relation (3.14).

It is straightforward to verify that the set $\{A \in \mathcal{F}_\infty \mid E \ni x \mapsto \mathbb{P}^x(A) \text{ is measurable}\}$ is a Dynkin system, which contains the cuboid sets. Also the set of all cuboid sets is stable under taking finite intersections and generates $\mathcal{F}_\infty$. Hence, the measurability follows from Gänssler and Stute [GS77, Theorem 1.1.22].

(M_4): Choose any $s, t \in \mathbb{R}_+$, $x \in E$ and $\Gamma \in \mathcal{B}$ then from (3.14) for $s = 0$, the law of total expectation and the transformation theorem

$$\begin{aligned} \mathbb{P}^x(\{X_t \in \Gamma\}) &= \int_\Omega \mathbb{P}^x(\{X_t \in \Gamma\} \mid \mathcal{F}_0)\, d\mathbb{P}^x = \int_\Omega p_t(X_0, \Gamma)\, d\mathbb{P}^x \\ &= \int_E p_t(y, \Gamma)\, \underbrace{\mathbb{P}^x X_0^{-1}}_{=\delta_x}(dy) = p_t(x, \Gamma) \end{aligned}$$

and so equation (3.14) yields

$$\mathbb{P}^{X_s}(\{X_t \in \Gamma\}) = p_t(X_s, \Gamma) = \mathbb{P}^x(\{X_{s+t} \in \Gamma\} \mid \mathcal{F}_s), \quad \mathbb{P}^x\text{-a.s.}$$

From this it is also clear that (3.12) is satisfied.

The last assertion now follows from Theorem 3.17 and the above. ■

The triple $((X_t)_{t\geq 0}, \mathcal{F}, \{\mathbb{P}^x\}_{x\in E})$ constructed in Theorem 3.18 is called the *canonical Markov process*. For it the shift operators for canonical Markov processes are given by $\theta_t : E^{\mathbb{R}_+} \to E^{\mathbb{R}_+}$,

$$(\theta_t \omega)(s) := \omega(t+s), \quad \omega \in E^{\mathbb{R}_+},\ t, s \geq 0. \tag{3.16}$$

4 Feller Semigroups and Feller Processes

A particularly interesting class of Markov processes are those whose transition semigroups leave the space $C_0(E)$ invariant, i.e., $T_t C_0(E) \subset C_0(E)$ for all $t \geq 0$. For Markov processes belonging to such transition semigroups some very useful properties and regularization theorems hold. We treat Feller semigroups, show their strong continuity and associate them to Markov processes. We then proof a regularization for these processes and construct the canonical Feller process. Then we present Dynkin's formula, which prepares the investigation of Feller diffusions with which we conclude this chapter.

In the following chapter we assume E to be a locally compact Polish space with Borel-σ-algebra $\mathcal{B}$.

4.1 Feller Semigroups

Definition 4.1 (Feller (transition) semigroups and processes)**.** Let E be a locally compact Polish space.

(i) A semigroup of positive contraction operators $(T_t)_{t\geq 0}$ on $C_0(E)$ is called *Feller semigroup* if

$$\lim_{t\to 0^+} T_t f(x) = f(x), \quad f \in C_0(E),\, x \in E. \tag{4.1}$$

(ii) A transition semigroup $(T_t)_{t\geq 0}$ is called Feller if $T_t C_0(E) \subset C_0(E)$ and its restriction to $C_0(E)$ defines a Feller semigroup.

(iii) A Markov process $(X)_{t\geq 0}$ is called Feller process if the corresponding transition semigroup is Feller.

Extension to transition semigroups on $B_b(E)$. We show that there is a one-to-one correspondence between Feller semigroups and transition semigroups that are Feller, i.e., the restriction of a Feller transition group to $C_0(E)$ already carries all information in order to reconstruct the transition semigroup.

We will first look at this for single transition operators $T : B_b(E) \to B_b(E)$ that satisfy $TC_0(E) \subset C_0(E)$. Clearly, their restriction to $C_0(E)$ is a positive linear contraction. We find that these properties are enough to characterize the transition operators on $B_b(E)$ that leave $C_0(E)$ invariant.

We call an operator $T : C_0(E) \to C_0(E)$ *conservative* if for all $x \in E$ we have $\sup_{f\in C_0,\, 0\leq f\leq 1} Tf(x) = 1$.

Lemma 4.2. *Let $T : C_0(E) \to C_0(E)$ be a positive linear contraction. Then there exists a unique sub-Markov kernel p on $(E, \mathcal{B})$ such that*

$$Tf(x) = \int_E f(y)\, p(x, dy), \quad x \in E,\ f \in C_0(E). \tag{4.2}$$

In particular, T extends to a unique transition operator on $B_b(E)$. The operator T is conservative iff p is a Markov kernel.

Proof. For each $x \in E$ the mapping $\varphi : C_0(E) \to \mathbb{R}$, $f \mapsto Tf(x)$ is a positive linear functional on $C_0(E)$ with $\|\varphi\| \leq 1$. By the Riesz representation theorem, cf. Theorem 2.7, there is a sub-probability measure $p(x, \cdot)$ on $\mathcal{B}$ such that (4.2) holds. Note that by the isometry property of the Riesz representation $p(x, \cdot)$ is a probability measure iff T is conservative.

We prove that for each $\Gamma \in \mathcal{B}$ the map $E \ni x \mapsto p(x, \Gamma)$ is measurable. It is straightforward to check that $\mathcal{D} := \{B \in \mathcal{B} \mid E \ni x \mapsto p(x, \Gamma) \text{ is measurable}\}$ is a Dynkin system. From Lemma 2.6 we obtain the existence of a generator $\mathcal{E}$ of $\mathcal{B}$ that is closed under taking finite intersections and is such that for $C \in \mathcal{E}$ there exists an equibounded sequence $(f_n)_{n\in\mathbb{N}}$ in $C_0(E)$ converging pointwise to 1_C[3]. Thus, by dominated convergence

[3] cf. Lemma 2.5.

and (4.2) we have $Tf_n(x) \to p(x, C)$ for each $x \in E$. As the pointwise limit of the measurable functions Tf_n is measurable we obtain that $\mathcal{E} \subset \mathcal{D}$, whence Theorem 1.1.22 in Gänssler and Stute [GS77] yields $\mathcal{B} = \mathcal{D}$. ■

Similarly, we call a semigroup of linear positive contractions on $C_0(E)$ *conservative* if for all $t \geq 0$ and all $x \in E$ we have $\sup_{f \in C_0,\, 0 \leq f \leq 1} T_t f(x) = 1$.

Proposition 4.3. *Let $(T_t)_{t \geq 0}$ be a semigroup of linear positive contractions on $C_0(E)$. Then there exists a unique transition function $(p_t)_{t \geq 0}$ on $(E, \mathcal{B})$ such that*

$$T_t f(x) = \int_E f(y)\, p_t(x, dy), \quad x \in E,\ f \in C_0(E),\ t \geq 0. \tag{4.3}$$

In particular, $(T_t)_{t \geq 0}$ extends to a unique transition semigroup on $B_b(E)$. The semigroup $(T_t)_{t \geq 0}$ is conservative iff $(p_t)_{t \geq 0}$ is conservative.

Proof. To each family of linear positive contractions on $C_0(E)$ Lemma 4.2 assures the existence and uniqueness of a family $(p_t)_{t \geq 0}$ of sub-Markov kernels satisfying (4.3). So, by this equation, $(T_t)_{t \geq 0}$ extends to a unique family of transition operators on $B_b(E)$, which we denote by $(\tilde{T}_t)_{t \geq 0}$. Showing that $(p_t)_{t \geq 0}$ is a transition function is now equivalent to proving that $\mathcal{B} = \mathcal{D} := \{B \in \mathcal{B} \mid \tilde{T}_0 1_B = 1_B \text{ and } \forall s, t \geq 0 : \tilde{T}_s \tilde{T}_t 1_B = \tilde{T}_{s+t} 1_B\}$[4].

It is straightforward to check that $\mathcal{D}$ is a Dynkin system. From Lemma 2.6 we obtain the existence of a generator $\mathcal{E}$ of $\mathcal{B}$ that is closed under taking finite intersections and is such that for $C \in \mathcal{E}$ there exists sequence $(f_n)_{n \in \mathbb{N}}$ in $C_0(E)$ converging to 1_C in the $\sigma(B_b(E), \mathrm{ca}(\mathcal{B}))$-topology. As transition operators are $\sigma(B_b(E), \mathrm{ca}(\mathcal{B}))$-continuous and $(\tilde{T}_t)_{t \geq 0}$ is a semigroup on $C_0(E)$ we obtain $\tilde{T}_0 1_C = 1_C$ and $\tilde{T}_s \tilde{T}_t 1_C = \tilde{T}_{s+t} 1_C$, and thus, $\mathcal{E} \subset \mathcal{D}$, whence Theorem 1.1.22 in Gänssler and Stute [GS77] yields $\mathcal{B} = \mathcal{D}$. The last assertion follows from Lemma 4.2, as p_t is a Markov kernel iff T_t is conservative. ■

Stochastic continuity. To explain condition (4.1) imposed on Feller semigroups we proceed to investigate a related property of the transition function $(p_t)_{t \geq 0}$.

We call a transition function $(p_t)_{t \geq 0}$ *stochastically continuous* if for every $x \in E$ and every $\varepsilon > 0$ we have $\lim_{t \downarrow 0} p_t(x, \bar{B}_\varepsilon(x)) = 1$. The following is Lemma II.2.2 in Dynkin [Dyn65].

Lemma 4.4. *Let $(T_t)_{t \geq 0}$ be a semigroup of linear positive contractions on $C_0(E)$. Then condition* (4.1) *holds iff the corresponding transition function*[5] *$(p_t)_{t \geq 0}$ is stochastically continuous.*

Proof. Assume that $(p_t)_{t \geq 0}$ is stochastically continuous. Let $x \in E$, $\varepsilon > 0$ and $f \in$

[4] cf. Lemma 3.4.

[5] cf. Proposition 4.3.

$C_0(E)$. Then

$$\begin{aligned}|T_t f(x) - f(x)| &\le \int_{B_\varepsilon(x)} |f(y) - f(x)|\, p_t(x, dy) + |f(x)||p_t(x, B_\varepsilon(x)) - 1| \\ &+ \int_{E \setminus B_\varepsilon(x)} |f(y)|\, p_t(x, dy) \\ &\le \sup_{y \in B_\varepsilon(x)} |f(y) - f(x)| + 2\|f\|(1 - p_t(x, B_\varepsilon(x))),\end{aligned}$$

which for ε and t small enough becomes arbitrarily small.

Conversely, assume that condition (4.1) holds and let $\varepsilon > 0$. By Urysohn's lemma there is a $f \in C_0(E)$ with $0 \le f \le 1$ and $f \equiv 1$ on $B_\varepsilon(x)$. Then

$$0 \le 1 - p_t(x, B_\varepsilon(x)) \le 1 - \int_E f(y)\, p_t(x, dy) = f(x) - T_t f(x),$$

which vanishes for $t \downarrow 0$. ■

4.2 Strong Continuity of Feller Semigroups

In the following our objective is to show that every Feller semigroup $(T_t)_{t \ge 0}$ is strongly continuous. To do this we consider for each $\lambda > 0$ an operator R_λ on $C_0(E)$ as the Laplace transform

$$R_\lambda f(x) := \int_0^\infty e^{-\lambda t} T_t f(x)\, dt, \quad x \in E,\ f \in C_0(E).$$

Note that this is well-defined, as $T_t f(x)$ is right-continuous and bounded uniformly in $t \ge 0$ for each fixed $x \in E$. This is Theorem 19.4 in Kallenberg [Kal02].

Theorem 4.5 (resolvent and generator)**.** *Let $(T_t)_{t \ge 0}$ be a Feller semigroup on $C_0(E)$. Then for each $\lambda > 0$ the operator λR_λ is an injective contraction on $C_0(E)$ such that strongly $\lambda R_\lambda \to I$ as $\lambda \to \infty$. Moreover, the range $\mathcal{D} = R_\lambda C_0(E)$ is independent of λ and dense in $C_0(E)$ and there exists an operator A with domain $\mathcal{D}$ such that $R_\lambda^{-1} = \lambda - A$ for every $\lambda > 0$. Also, A commutes with T_t on $\mathcal{D}$ for every $t \ge 0$.*

Proof. Fix $\lambda > 0$ and let $f \in C_0(E)$. As $T_t f \in C_0(E)$ it follows from Lebesgue's dominated convergence theorem that $R_\lambda f \in C_0(E)$. The contraction property follows from

$$\|\lambda R_\lambda f\| \le \lambda \int_0^\infty e^{-\lambda t} \|T_t f\|\, dt \le \lambda \int_0^\infty e^{-\lambda t}\, dt \|f\| = \|f\|.$$

For $\lambda, \mu > 0$ and $f \in C_0(E)$ we obtain from Fubini's theorem and a simple substitution the *resolvent equation*

$$
\begin{aligned}
R_\lambda f - R_\mu f &= \int_0^\infty e^{-\mu s}(e^{-(\lambda-\mu)s} - 1)T_s f\, ds \\
&= \int_0^\infty e^{-\mu s} \int_0^s (\mu - \lambda) e^{-(\lambda-\mu)r}\, dr\, T_s f\, ds \\
&= (\mu - \lambda) \int_0^\infty \int_r^\infty e^{-\mu s} e^{-(\lambda-\mu)r} T_s f\, ds\, dr \\
&= (\mu - \lambda) \int_0^\infty \int_0^\infty e^{-\mu(s+r)} e^{-(\lambda-\mu)r} T_{s+r} f\, ds\, dr \\
&= (\mu - \lambda) \int_0^\infty e^{-\lambda r} T_r \left(\int_0^\infty e^{-\mu s} T_s f\, ds \right) dr \\
&= (\mu - \lambda) R_\lambda R_\mu f.
\end{aligned}
\tag{4.4}
$$

This shows that the operators R_λ commute. Also for any f in the range of R_λ there is $g \in C_0(E)$ such that $f = R_\lambda g$. By equation (4.4) we then have $f = R_\mu(g + (\mu - \lambda)R_\lambda g)$, which shows that the range $\mathcal{D}$ is independent of λ. For $f \in \mathcal{D}$ and $g \in C_0(E)$ with $f = R_1 g$ we obtain from the resolvent equation

$$
\|\lambda R_\lambda f - f\| = \|(\lambda R_\lambda - I)R_1 g\| = \|(R_1 - I)R_\lambda g\| \le \lambda^{-1}\|R_1 - I\|\|g\|,
$$

which vanishes as $\lambda \to \infty$. By approximating $f \in \overline{\mathcal{D}}$ by a sequence in $\mathcal{D}$ we may extend this convergence to the closure of $\mathcal{D}$. Now assume that $\overline{\mathcal{D}} \neq C_0(E)$. Then by Hahn-Banach's extension theorem[6] we find a functional $\varphi \not\equiv 0$ with $\varphi = 0$ on $\overline{\mathcal{D}}$. By the Riesz representation theorem, see Theorem 2.7, we find $\mu \in \mathrm{rca}(\mathcal{B})$ such that $\varphi(f) = \int f\, d\mu$, $f \in C_0(E)$. Letting $f \in C_0(E)$ we obtain from the pointwise convergence in condition (4.1) and dominated convergence that

$$
\begin{aligned}
0 = \varphi(\lambda R_\lambda f) &= \int_E \int_0^\infty \lambda e^{-\lambda t} T_t f(x)\, dt\, \mu(dx) \\
&= \int_E \int_0^\infty e^{-t} T_{t/\lambda} f(x)\, dt\, \mu(dx) \to \varphi(f),
\end{aligned}
$$

as $\lambda \to \infty$. This implies $\varphi \equiv 0$ contradicting the assumption that $\overline{\mathcal{D}} \neq C_0(E)$, and so $\mathcal{D}$ is dense in $C_0(E)$.

We show that the operators R_λ are injective. To see this, let $f \in C_0(E)$ be such that $R_{\lambda_0} f = 0$ for some $\lambda_0 > 0$. Then the resolvent equation (4.4) yields $R_\lambda f = 0$ for all $\lambda > 0$. For $\lambda \to \infty$ we therefore obtain $0 = \lambda R_\lambda f \to f$, and so $f = 0$. Therefore, R_λ is injective and the inverse R_λ^{-1} exists and is defined on $\mathcal{D}$. Multiplying the resolvent equation by R_λ^{-1} from the left and by R_μ^{-1} from the right we obtain after rearranging $\lambda - R_\lambda^{-1} = \mu - R_\mu^{-1}$, and so $A := \lambda - R_\lambda$ does not depend on λ.

Finally, as T_t and R_λ commute we obtain

$$
T_t(\lambda - A)R_\lambda = T_t = (\lambda - A)R_\lambda T_t = (\lambda - A)T_t R_\lambda, \quad t, \lambda > 0,
$$

and so after subtracting by $\lambda T_t R_\lambda$ and multiplying R_λ^{-1} from the right we obtain $T_t A = A T_t$ on $\mathcal{D}$. ■

[6] Rather a corollary of it, cf. Corollary 2.3.5 in Pedersen [Ped89].

For any strongly continuous semigroup $(S_t)_{t\geq 0}$ on $C_0(E)$ we call the linear operator $A : \mathcal{D}(A) \to C_0(E)$, $\mathcal{D}(A) \subset C_0(E)$ *generator* if

$$Af(x) = (Af)(x) = \lim_{t\to 0^+} \frac{S_t f(x) - f(x)}{t}, \quad f \in \mathcal{D}(A),\, x \in E,$$

where $\mathcal{D}(A) := \{f \in C_0(E) \mid \lim_{t\to 0^+} \frac{S_t f - f}{t} \in C_0(E)\}$.

Corollary 4.6 (strong continuity)**.** *Every Feller semigroup $(T_t)_{t\geq 0}$ is strongly continuous with generator $A = \lambda - R_\lambda^{-1}$, $\lambda > 0$ and $\mathcal{D}(A) = R_\lambda C_0(E)$.*

Proof. Theorem 4.5 and the theorem of Hille-Yosida, see Engel and Nagel [EN06, Theorem 3.5], yields that A is the generator of a strongly continuous contraction semigroup $(S_t)_{t\geq 0}$ on $C_0(E)$. By Theorem II.1.10 in Engel and Nagel [EN06] we obtain that

$$R_\lambda f(x) = (\lambda - A)^{-1} f(x) = \int_0^\infty e^{-\lambda} S_t f(x)\, dt, \quad \lambda > 0,\, f \in C_0(E),\, x \in E.$$

Now the uniqueness theorem for Laplace transforms and the fact that the mappings $\mathbb{R}_+ \ni t \mapsto T_t f(x)$ and $\mathbb{R}_+ \ni t \mapsto S_t f(x)$ are right-continuous implies that $T_t = S_t$ for all $t \geq 0$. ■

We recall a properties of the generator $(A, \mathcal{D}(A))$ that generalizes the Kolmogorov forward and backward equations for certain diffusion processes[7]. For a proof see Engel and Nagel [EN06, Lemma II.1.3].

Remark 4.7 (forward and backward equation). Let be $(S_t)_{t\geq 0}$ strongly continuous semigroup on $C_0(E)$ with generator $(A, \mathcal{D}(A))$. For $f \in C_0(E)$ and $t \geq 0$ the semigroup $(S_t)_{t\geq 0}$ satisfies

$$S_t f - f = S_t A f = \int_0^t S_s A f\, ds.$$

Moreover, if $f \in \mathcal{D}(A)$ the map $\mathbb{R}_+ \ni t \mapsto S_t f$ is differentiable and satisfies

$$\frac{d}{dt}(S_t f) = S_t A f = A S_t f.$$

Positive-maximum principle. We say that a linear operator $(A, \mathcal{D}(A))$ on $C_0(E)$ satisfies the *positive-maximum principle*, if we have $Af(x) \leq 0$ whenever $f \in \mathcal{D}(A)$ satisfies $f \vee 0 \leq f(x)$ for $x \in E$. The next lemma will assure that this fundamental property is satisfied by generators of Feller semigroups.

Lemma 4.8 (positive-maximum principle)**.** *Let $(A, \mathcal{D}(A))$ be the generator of a Feller semigroup on $C_0(E)$. Then A satisfies the positive-maximum principle.*

[7] cf. Theorem 4.22

Proof. Let $f \in \mathcal{D}(A)$ and $x \in E$ such that $f^+ = f \vee 0 \leq f(x)$. Let $t \geq 0$. Then by linearity, positivity and contractivity of T_t implies

$$T_t f(x) \leq T_t f^+(x) \leq \|T_t f^+\| \leq \|f^+\| = f(x).$$

Thus, $Af(x) = \lim_{t \to 0^+} t^{-1}(T_t f(x) - f(x)) \leq 0$. ∎

An easy observation is that every linear operator $(A, \mathcal{D}(A))$ on $C_0(E)$ satisfying the positive-maximum principle must be *dissipative*, i.e, it satisfies

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

Lemma 4.9. *Let $(A, \mathcal{D}(A))$ be a linear operator on $C_0(E)$ satisfying the positive-maximum principle. Then A is dissipative.*

Proof. Assume that there exists $\lambda > 0$ and $f \in \mathcal{D}(A)$ such that $\|(\lambda - A)f\| < \lambda\|f\|$. We may assume that there exists $x \in E$ such that $f \vee 0 \leq f(x)$. Indeed, as $f \in C_0(E)$, it attains its maximum. If f attains some nonnegative value then this property is always fulfilled. If f only takes on negative values, we may consider $-f$ instead. We thus obtain

$$(\lambda - A)f(y) < \lambda f(x), \quad y \in E,$$

whereas the positive-maximum principle yields $(\lambda - A)f(x) \geq \lambda f(x)$. ∎

We'll find that in the class of linear operators $(A, \mathcal{D}(A))$, which satisfy the positive-maximum principle, the generators of Feller semigroups are maximal. This is Lemma 19.12 in Kallenberg [Kal02].

Lemma 4.10 (maximality)**.** *Let $(A, \mathcal{D}(A))$ be the generator of a Feller semigroup. Assume that A extends to a linear operator $(A', \mathcal{D}(A'))$ satisfying the positive-maximum principle. Then $\mathcal{D}(A) = \mathcal{D}(A')$.*

Proof. Let $f \in \mathcal{D}(A')$ and set $g = (I - A')f$. As A' is dissipative by Lemma 4.9 and $(I - A)R_1 = I$ on $C_0(E)$ we obtain

$$\|f - R_1 g\| \leq \|(I - A')(f - R_1 g)\| = \|g - (I - A)R_1 g\| = 0,$$

and so $f = R_1 g \in \mathcal{D}(A)$. ∎

4.3 The Canonical Construction of a Feller Process

Our goal is to construct a Feller process corresponding to a given Feller semigroup. As Feller processes are Markov processes it therefore suffices to show that we can associate a Markov process to each Feller semigroup. This is not obvious, as Feller semigroups don't necessarily correspond to conservative transition functions on $(E, \mathcal{B})$. In fact this is only true if the Feller semigroup is already conservative, cf. Proposition 4.3.

Compactification and Existence. We introduce a new state $\Delta \notin E$ and let $\hat{E} = E \cup \{\Delta\}$ be the one-point compactification of E. As E is a locally compact separable metric space the compact space $\hat{E}$ is separable and has a metric, cf. Mandelkern [Man89, Theorem 2]. In particular, as every compact metric space is complete[8], the space $\hat{E}$ is a compact Polish space. We denote by $\hat{\mathcal{B}}$ its Borel-σ-algebra.

Any function $f \in C_0(E)$ is continuously extended to $\hat{E}$ by setting $f(\Delta) = 0$. In this way we can embed the space $C_0(E)$ into $C(\hat{E})$, the space of all continuous functions on $\hat{E}$. We now extend the original Feller semigroup on $C_0(E)$ to a conservative semigroup on $C(\hat{E})$. The following is Lemma 19.13 in Kallenberg [Kal02].

Lemma 4.11 (compactification)**.** *Any Feller semigroup $(T_t)_{t\geq 0}$ on $C_0(E)$ can be extended to a conservative Feller semigroup on $C(\hat{E})$, which is given by*

$$\hat{T}_t f = f(\Delta) + T_t \left[f - f(\Delta) \right], \quad t \geq 0,\ f \in C(\hat{E}).$$

Proof. Well-definition and the semigroup property are straightforward to check. Note that from the compactness of $\hat{E}$ it follows that $C(\hat{E}) = C_0(\hat{E})$. Condition (4.1) follows easily from the definition of $\hat{T}_t$. For $f \in C(\hat{E})$, $f \geq 0$ the function $g := f(\Delta) - f \in C_0(E)$ satisfies $g \leq f(\Delta)$. Hence, by positivity of T_t,

$$T_t g \leq T_t g^+ \leq \|T_t g^+\| \leq \|g^+\| \leq f(\Delta)$$

and thus $\hat{T}_t f = f(\Delta) - T_t g \geq 0$ for all $t \geq 0$. The contraction property and conservativity follow from the fact that $\hat{T}_t 1 = 1$. ∎

Let $(p_t)_{t\geq 0}$ be a transition function. We call a state $x \in \hat{E}$ *absorbing for* $(p_t)_{t\geq 0}$ if $p_t(x, \{x\}) = 1$ for every $t \geq 0$. Proposition 4.3 shows that we can associate to $(\hat{T}_t)_{t\geq 0}$ a unique conservative transition function $(p_t)_{t\geq 0}$ on $(\hat{E}, \hat{\mathcal{B}})$. In particular, this transition function satisfies

$$T_t f(x) = \int_E f(y)\, p_t(x, dy), \quad f \in C_0(E),\ x \in E, \tag{4.5}$$

and that Δ is also absorbing for $(p_t)_{t\geq 0}$. In fact, we will see that $(p_t)_{t\geq 0}$ is even uniquely determined by these properties. This is Proposition 19.14 in Kallenberg [Kal02].

Proposition 4.12 (existence)**.** *For any Feller semigroup $(T_t)_{t\geq 0}$ on $C_0(E)$ there exists a unique conservative transition function $(p_t)_{t\geq 0}$ on $(\hat{E}, \hat{\mathcal{B}})$ satisfying* (4.5) *for which Δ is absorbing.*

Proof. Let $(p_t)_{t\geq 0}$ be the unique conservative transition function[9] on $(\hat{E}, \hat{\mathcal{B}})$ satisfying

$$\hat{T}_t f(x) = \int_{\hat{E}} f(y) p_t(x, dy), \quad f \in C(\hat{E}),\ x \in \hat{E}. \tag{4.6}$$

[8] To see this note that every sequence $(x_n)_{n\in\mathbb{N}}$ in $\hat{E}$ has a subsequence $(x_{n_m})_{m\in\mathbb{N}}$ converging to some x. Assuming that$(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence an application of the triangle inequality yields $d(x_n, x) \leq d(x_n, x_{n_m}) + d(x_{n_m}, x)$, which becomes arbitrarily small for large enough n.

[9] cf. Proposition 4.3.

Equation (4.5) is now a special case of equation (4.6). Moreover, equation (4.6) implies that

$$\int_E f(y)p_t(\Delta, dy) = \hat{T}_t f(\Delta) = 0, \quad f \in C_0(E),$$

which in turn implies that $p_t(\Delta, \{\Delta\}) = 1$. Hence, Δ is absorbing for $(p_t)_{t\geq 0}$.

The uniqueness of $(p_t)_{t\geq 0}$ is also a consequence Proposition 4.3. ■

Regularization. Proposition 4.12 together with Theorem 3.18 now imply that we can associate to each Feller semigroup $(T_t)_{t\geq 0}$ a canonical Markov process X, which by Definition 4.1 is a Feller process.

As $\mathbb{R}_+$ is uncountable, choosing $\Omega = E^{\mathbb{R}_+}$ and $\mathcal{A} = \mathcal{B}^{\mathbb{R}_+}$ for the canonical construction of the Markov process has some defects, e.g.,

(i) the sets whose description depend on an uncountable number of distinct times are not in the σ-algebra $\mathcal{B}^{\mathbb{R}_+}$,

(ii) the mapping $\mathbb{R}_+ \times \Omega \ni (t, \omega) \mapsto X_t(\omega)$ is not in general jointly measurable.

Being able to assume more regularity on the paths of a Feller process would overcome the above deficiencies. Hence, the aim of this subsection is to proof the following theorem.

Let X be a process with state space $\hat{E}$. We say for a path $\omega \in \Omega$ that Δ is *absorbing for* $X^\pm$ if

$$\forall t \geq 0 : (X_t(\omega) = \Delta \vee \lim_{s \to t^-} X_s(\omega) = \Delta) \implies (\forall u \geq t : X_u(\omega) = \Delta).$$

We say Δ is *absorbing for* $X^\pm$ if this holds for all $\omega \in \Omega$.

Theorem 4.13. *Let $(T_t)_{t\geq 0}$ be a Feller semigroup then there exists a Feller process X in $\hat{E}$ such that X has càdlàg paths and Δ is absorbing for $X^\pm$. If the semigroup on $C_0(E)$ is already conservative, we can choose the state space to be E.*

The proof will require some insights on how a canonical Feller process can be constructed. We will start by following an idea in Arnold [Arn98, Appendix A.2].

Let $(\hat{E}^{\mathbb{R}_+}, \hat{\mathcal{B}}^{\mathbb{R}_+}, \mathbb{P})$ be a probability space. For $\Omega_0 \subset \hat{E}^{\mathbb{R}_+}$ not necessarily measurable and $\mathbb{P}^*(\Omega_0) = 1$, where $\mathbb{P}^*(\Omega_0)$ denotes the *outer measure* of Ω_0

$$\mathbb{P}^*(\Omega_0) := \inf\{\mathbb{P}(A) \mid \Omega_0 \subset A,\ A \in \mathcal{B}^{\mathbb{R}_+}\}, \tag{4.7}$$

we can restrict $(\hat{E}^{\mathbb{R}_+}, \mathcal{B}^{\mathbb{R}_+}, \mathbb{P})$ to $(\Omega_0, \mathcal{A}_0, \mathbb{P}_0)$, where $\mathbb{P}_0$ is the unique probability measure on the trace-σ-algebra $\mathcal{A}_0 := \Omega_0 \cap \mathcal{B}_+^{\mathbb{R}}$ defined by

$$\mathbb{P}_0(\Omega \cap A) := \mathbb{P}(A), \quad A \in \mathcal{B}^{\mathbb{R}_+}, \tag{4.8}$$

see Gänssler and Stute [GS77, Theorem 7.1.18]. As the finite dimensional distribution of both, $\mathbb{P}$ and $\mathbb{P}_0$, coincide, the functions $X_t(\omega) := \omega(t)$ define stochastic processes with the same law but the the paths of the restricted system lie in Ω_0.

In the case of Feller processes, the next result will motivate this restriction to a smaller path space. Recall that for a process $Y = (Y_t)_{t\geq 0}$ on a measurable space $(\Omega, \mathcal{A})$ is a *modification* with respect to a probability measure $\mathbb{P}$ on $\mathcal{A}$ of another process $X = (X_t)_{t\geq 0}$ on the same space, if

$$\forall t \geq 0 : \mathbb{P}(\{Y_t = X_t\}) = 1.$$

The following is Theorem 19.15 in Kallenberg [Kal02].

Theorem 4.14 (Regularization)**.** *Let X be a Feller process in $\hat{E}$ and let ν be an initial distribution on $\hat{\mathcal{B}}$. Then there exists a càdlàg process $\tilde{X}$ for which holds that*

(i) $\tilde{X}$ is a modification of X with respect to $\mathbb{P}^\nu$,

(ii) Δ is absorbing for $\tilde{X}^\pm$.

If X's corresponding Feller semigroup $(T_t)_{t\geq 0}$ on $C_0(E)$ is conservative and ν restricted to $\mathcal{B}$ is a probability measure, we can choose $\tilde{X}$ to have state space E.

For the sake of coherence we will postpone the proof of this theorem to continue the discussion before.

In view of the regularization theorem we may choose

$$\Omega_0 = \{\omega \in D(\mathbb{R}_+; \hat{E}) \mid \forall t \geq 0 : ((\omega(t) = \Delta \vee \omega(t-) = \Delta) \implies (\forall u \geq t : \omega(t) = \Delta))\}$$

and $\mathcal{A}_0$ to be the trace-σ-algebra of $\hat{\mathcal{B}}^{\mathbb{R}_+}$ in Ω_0[10], for which obtain the following corollary of the regularization theorem.

Corollary 4.15. *Let X be a $\hat{E}$-valued Feller process on $(\Omega, \mathcal{A})$ and let ν be an initial distribution on $\hat{\mathcal{B}}$. Then we have for the* outer measure $(\mathbb{P}^\nu X^{-1})^*$ *of the pushforward measure $\mathbb{P}^\nu X^{-1}$ on $\hat{\mathcal{B}}^{\mathbb{R}_+}$ that $(\mathbb{P}^\nu X^{-1})^*(\Omega_0) = 1$.*

If X's corresponding Feller semigroup $(T_t)_{t\geq 0}$ on $C_0(E)$ is conservative we may replace the $\hat{E}$ resp. $\hat{\mathcal{B}}$ in the definition of $(\Omega_0, \mathcal{A}_0)$ by E resp. $\mathcal{B}$.

Proof. Let a $A \in \hat{\mathcal{B}}^{\mathbb{R}_+}$ be such that $\Omega_0 \subset A$. Then there exists a countable index set $I \subset \mathbb{R}_+$ and $B \in \hat{\mathcal{B}}^I$ such that

$$A = B \times \bigtimes_{t \in \mathbb{R}_+ \setminus I} \hat{E}.$$

Let π_I be the projection map $\pi_I : \hat{E}^{\mathbb{R}_+} \to \hat{E}^I$. By the regularization theorem, cf. Theorem 4.14, we find a modification $\tilde{X}$ of X with respect to $\mathbb{P}^\nu$ and with paths in Ω_0. Note that then $\tilde{X} : (\Omega, \mathcal{A}) \to (\Omega_0, \mathcal{A}_0)$. Hence, we obtain

$$\begin{aligned} \mathbb{P}^\nu X^{-1}(A) &= \mathbb{P}^\nu(\{X \in A\}) = \mathbb{P}^\nu(\{\pi_I(X) \in B\}) = \mathbb{P}^\nu(\{\pi_I(X) \in B\} \cap \bigcap_{t\in I}\{X_t = \tilde{X}_t\}) \\ &= \mathbb{P}^\nu(\{\pi_I(\tilde{X}) \in B\} \cap \bigcap_{t\in I}\{Y_t = \tilde{X}_t\}) = \mathbb{P}^\nu(\{\pi_I(\tilde{X}) \in B\}) = \mathbb{P}^\nu(\{\tilde{X} \in A\}) = 1. \end{aligned}$$

The last statement follows analogously using the last statement of Theorem 4.14. ■

[10] This incidentally coincides with the trace-σ-algebra of $\mathcal{D}$ in Ω_0, see Section 2.1.

We are now ready to state the proof of Theorem 4.13.

Proof of Theorem 4.13. Let $\hat{X} = ((\hat{X}_t)_{t\geq 0}, \hat{\mathcal{F}}, \{\mathbb{P}^x\}_{x\in E})$ be a canonical $\hat{E}$-valued Markov process corresponding to the Feller semigroup $(T_t)_{t\geq 0}$. Set X to be the identity mapping on the measurable space $(\Omega_0, \mathcal{A}_0)$ so that $X_t(\omega) = \omega(t)$ for $t \geq 0$ and $\omega \in \Omega_0$, and let $\mathcal{F}$ to be the canonical filtration of X.

By the previous discussion and Corollary 4.15 the triple $(X, \mathcal{F}, \mathbb{P}_0^x)$, where $\mathbb{P}_0^x$ is the probability measure on $\mathcal{A}_0$ constructed from $\mathbb{P}^x$ as in (4.8), is an $\hat{E}$-valued Markov random function for each $x \in \hat{E}$. Now Theorem 3.18 implies that $(X, \mathcal{F}, \{\mathbb{P}_0^x\}_{x\in\hat{E}})$ is a càdlàg Markov process corresponding to the same transition semigroup as $\hat{X}$, as it has the same probability law.

If the Feller semigroup on $C_0(E)$ is conservative, then proof goes analogously by replacing $\hat{E}$ with E and $\hat{\mathcal{B}}$ with $\mathcal{B}$. ■

The process constructed in the proof is called the *canonical Feller process* for $(T_t)_{t\geq 0}$.

Terminal time. As a corollary of the regularization theorem we obtain the following.

Corollary 4.16. *Let X be a Feller process in $\hat{E}$ with right-continuous paths and let ν be an initial distribution on $\hat{\mathcal{B}}$. Then X has $\mathbb{P}^\nu$-a.s. càdlàg paths for which Δ is absorbing for $X^\pm$.*

This allows us to define for each right-continuous Feller process a *terminal time* ζ by

$$\zeta = \inf\{t \geq 0 \mid X_t = \Delta \text{ or } X_{t-} = \Delta\}, \quad \mathbb{P}^\nu\text{-a.s.}$$

In particular, given any initial distribution ν, we obtain that the probability law of ζ with respect to $\mathbb{P}^\nu$ is independent of the chosen right-continuous Feller process X. Furthermore, if the corresponding Feller semigroup is conservative, then $\zeta = \infty$, $\mathbb{P}^\nu$-a.s.

Proof of the regularization theorem. For the proof of the regularization theorem we construct a rich class of supermartingales to which we can apply the regularity results in Section 2.3.

Lemma 4.17. *Let A be the generator of a Feller semigroup and let X be a Feller process associated to the semigroup. Let $f \in C_0(E)$ with $f \geq 0$. Then the Process $Y = (Y_t)_{t\geq 0}$ defined by*

$$Y_t := e^{-t} R(1, A) f(X_t), \quad t \geq 0,$$

is a supermartingale with respect to $\mathbb{P}^\nu$ for every initial distribution ν.

Proof. Proposition 3.14, Proposition 3.16 yield $\mathbb{P}^\nu$-a.s.

$$\begin{aligned}\mathbb{E}^\nu\left[Y_{t+h} \mid \mathcal{F}_t^X\right] &= e^{-t-h}\mathbb{E}^\nu\left[R_1(A)f(X_t\circ\theta_h) \mid \mathcal{F}_t^X\right]\\ &= e^{-t-h}\mathbb{E}^{X_t}\left[R_1(A)f(X_h)\right] = e^{-t-h}T_hR_1(A)f(X_t)\\ &= e^{-t-h}T_h\int_0^\infty e^{-s}T_sf(X_t)\,ds = e^{-t}\int_0^\infty e^{-s-h}T_{s+h}f(X_t)\,ds\\ &= e^{-t}\int_h^\infty e^{-s}T_sf(X_t)\,ds \le e^{-t}\int_0^\infty e^{-s}T_sf(X_t)\,ds\\ &= e^{-t}R_1(A)f(X_t) = Y_t,\end{aligned}$$

as $T_tf \ge 0$ for all $t \ge 0$, by positivity of T_t. So Y is a supermartingale. ■

Proof of Theorem 4.14. 1. Step: We first construct a càdlàg process $\tilde{X}$ that is a candidate for a modification of X.

Fix ν and let $f \in \mathcal{D}(A)$. Then there exists $g \in C_0(E)$ such that $f = R_1(A)g$. Note that $g^+, g^- \in C_0(E)$ with $g^+, g^- \ge 0$. So we may apply Lemma 4.17 to obtain that $Y_t^+ := e^{-t}R_1(A)g^+(X_t)$ and $Y_t^- := e^{-t}R_1(A)g^-(X_t)$ define supermartingales under $\mathbb{P}^\nu$. Hence, Theorem 2.9 implies that $f(X_t) = e^t(Y_t^+ - Y_t^-)$ has right- and left-hand limits along $\mathbb{Q}_+$ outside of some $\mathbb{P}^\nu$-null set A_f. Recall that $\mathcal{D}(A)$ is dense in $C_0(E)$[11]. So for $f \in C_0(E)$ there exists a sequence $(f_k)_{k\in\mathbb{N}}$ in $\mathcal{D}(A)$ such that $f_n \to f$. Then for every $t \ge 0$ and any sequence $t_n \downarrow t$ in $\mathbb{Q}_+$ we have

$$\begin{aligned}|f(X_{t_n}) - f(X_{t_m})| &\le |f(X_{t_n}) - f_k(X_{t_n})| + |f_k(X_{t_n}) - f_k(X_{t_m})| + |f_k(X_{t_m}) - f(X_{t_m})|\\ &\le 2\|f - f_k\| + |f_k(X_{t_n}) - f_k(X_{t_m})|,\end{aligned} \tag{4.9}$$

which becomes small by first fixing k large enough so that the first term is small enough and then choosing m, n large enough so that the second term becomes small enough. So, by completeness, $(f(X_t))_{t\ge0}$ has right-hand limits along $\mathbb{Q}_+$ outside of the $\mathbb{P}^\nu$-null set $A_f := \bigcup_{n\in\mathbb{N}} A_{f_n}$. Similarly, one shows this for the left-hand side limits.

As $C_0(E)$ is separable we may choose a countable dense subset $\mathcal{D} \subset C_0(E)$. Then $N := \bigcup_{f\in\mathcal{D}} A_f$ is a $\mathbb{P}^\nu$-null set independent of f outside of which, by the same argument as in equation (4.9), the left- and right-hand limits of $(f(X_t))_{t\ge0}$ exist for all $f \in C_0(E)$.

If $(x_n)_{n\in\mathbb{N}}$ is a sequence in $\hat{E}$ such that $f(x_n)$ converges for every $f \in C_0(E)$, then also x_n converges in the topology of $\hat{E}$[12]. So, X has itself right- and left-hand limits $X_{t\pm}$ along $\mathbb{Q}_+$ on N^c; we may redefine X to be 0 on N. Now, as $\mathbb{Q}_+$ is dense in $\mathbb{R}_+$, the process defined by $\tilde{X}_t = X_{t+}$ exists and is càdlàg.

[11] cf. Theorem 4.5 and Corollary 4.6.

[12] Indeed, by compactness of $\hat{E}$, $(x_n)_{n\in\mathbb{N}}$ has at least one accumulation point and if it is unique the sequence converges to it. Let $x^{(1)}$ and $x^{(2)}$ be accumulation points. Then there exist two subsequences $(n_k^{(1)})_{k\in\mathbb{N}}$ and $(n_k^{(2)})_{k\in\mathbb{N}}$ of $(n)_{n\in\mathbb{N}}$ such that $x_{n_k^{(i)}} \to x^{(i)}$, $i = 1, 2$. But then by continuity $f(x_{n_k^{(i)}}) \to f(x^{(i)})$, $i = 1, 2$. Hence, by the assumed convergence $f(x^{(1)}) = f(x^{(2)})$ for all $f \in C_0(E)$. As, by Urysohn's lemma, $C_0(E)$ separates points in $\hat{E}$, we obtain $x^{(1)} = x^{(2)}$ and thus the claimed convergence.

2. Step: We show that $\tilde{X}$ is indeed a modification of X.

Fix $t \geq 0$ and let $t_n \downarrow t$ in $\mathbb{Q}_+$. Note, that $X_{t_n} \to X_t$ in probability. To show this, denote by d a metric on $\hat{E}$. Let $x \in \hat{E}$ and $\varepsilon > 0$. Then we may choose $f \in C(\hat{E})$ with $0 \leq f \leq 1$, $f \equiv 1$ on $\{y \in E \mid d(y,x) \leq \varepsilon/2\}$ and $f \equiv 0$ on $\{y \in E \mid d(y,x) \geq \varepsilon\}$, which exists by Urysohn's Lemma. So by Proposition 3.16 we get the convergence

$$\mathbb{P}^x(\{d(X_{t_n-t},x) < \varepsilon\}) \geq \mathbb{E}^x\left[f(X_{t_n-t})\right] = \hat{T}_{t_n-t}f(x) \to f(x) = 1,$$

as $n \to \infty$, from condition (4.1). Moreover, by contractivity of the semigroup we have $\|T_{t_n-t}f\| \leq 1$ for every n. From this together with the law of total expectation, the mixture property in Proposition 3.13 and Lebesgue's dominated convergence theorem we infer

$$\mathbb{P}^\nu(\{d(X_{t_n},X_t) \geq \varepsilon\}) = \int_E \underbrace{\mathbb{P}^x(\{d(X_{t_n},X_t) \geq \varepsilon\})}_{\to 0} \mu(dx) \to 0,$$

as $n \to \infty$. So $X_{t_n} \to X_t$ in probability as claimed. Then, as $X_{t_n} = \tilde{X}_{t_n}$, by definition, and $\tilde{X}_{t_n} \to \tilde{X}_t$ with $\mathbb{P}^\nu$-probability 1, we obtain that $\mathbb{P}^\nu$-a.s. $\tilde{X}_t = X_t$. As t was arbitrary, it follows that $\tilde{X}$ is a modification of X.

3. Step: It remains to show that we can redefine $\tilde{X}$ in such a way that Δ is absorbing for $\tilde{X}^\pm$.

Fix any $f \in C_0(E)$ with $f > 0$. Then, by positivity $T_s f \geq 0$ for all $s \geq 0$ and by condition (4.1) we find for each $x \in E$ an $\varepsilon_x > 0$ such that $T_s f(x) > 0$ for all $s \in [0,\varepsilon_x]$. Hence, we obtain from the Laplace representation of the resolvent that

$$R_1(A)f(x) = \int_0^\infty e^{-s} T_s f(x)\, ds > 0, \quad x \in E$$

Applying Lemma 2.10 to the right-continuous, nonnegative supermartingale defined by $Y_t = e^t R_1(A) f(\tilde{X}_t)$, $t \geq 0$, and noting that $Y_t = 0$ iff $\tilde{X}_t = \Delta$ yields $\mathbb{P}^\nu$-a.s. $\tilde{X} = \Delta$ on $[\zeta,\infty)$, where $\zeta = \inf\{t \geq 0 \mid \tilde{X}_t = \Delta \text{ or } \tilde{X}_{t-} = \Delta\}$. By setting $\tilde{X} \equiv \Delta$ on the exceptional $\mathbb{P}^\nu$-null set we can make this hold identically.

4. Step: Lastly, we show that $\tilde{X}$ can be chosen to be E-valued, if $(T_t)_{t\geq 0}$ is conservative.

Let $(T_t)_{t\geq 0}$ be conservative and ν restricted to $\mathcal{B}$ be a probability measure. Then we have for all $t \geq 0$ and $x \in E$ by dominated convergence and Proposition 3.16 that $\mathbb{P}^x(\{\tilde{X}_t \in E\}) = \sup_{f \in C_0(E), 0 \leq f \leq 1} \mathbb{E}^x[f] = \sup_{f \in C_0(E), 0 \leq f \leq 1} T_t f(x) = 1$. Thus, by the mixture property in Proposition 3.13 we have that $\tilde{X}_t \in E$ outside of some $\mathbb{P}^\nu$-null set N_t. As Δ is absorbing for $\tilde{X}^\pm$, we also have $\tilde{X}_{t-} \in E$ outside of N_t, and so $\zeta > t$ outside of N_t. Hence, we obtain $\zeta = \infty$ outside the $\mathbb{P}^\nu$-null set $N := \bigcup_{n\in\mathbb{N}} N_n$. By redefining $\tilde{X}_t$ for each $t \geq 0$ outside of N, we may assume that $\tilde{X}_t$ and $\tilde{X}_{t-}$ take values in E.

■

4.4 Dynkin's Formula

Let $X = (X_t)_{t\geq 0}$ be a Feller process with state space $\hat{E}$ and denote by $(T_t)_{t\geq 0}$ the corresponding Feller semigroup on $C_0(E)$, let $(A, \mathcal{D}(A))$ be the semigroup's generator and $f \in \mathcal{D}(A)$. An important role in the analysis of Feller processes is played by the process

$$M_t^f = f(X_t) - f(X_0) - \int_0^t Af(X_s)\,ds, \quad t \geq 0.$$

Recall that an $\mathcal{F}$-*optional time* is a random variable $\tau : \Omega \to [0, \infty]$ such that $\{\tau \leq t\} \in \mathcal{F}_t$ for every $t \geq 0$. This is Lemma 19.21 in Kallenberg [Kal02].

Lemma 4.18 (Dynkin's formula)**.** *Let X be a Feller process. The processes $M^f = (M_t^f)_{t\geq 0}$, $f \in \mathcal{D}(A)$ are martingales under every initial distribution ν on $\hat{\mathcal{B}}$. In particular, we have for every bounded $\mathcal{F}$-optional time τ*

$$\mathbb{E}^x\left[f(X_\tau)\right] = f(x) + \mathbb{E}^x\left[\int_0^\tau Af(X_s)\,ds\right], \quad f \in \mathcal{D}(A),\, x \in E. \tag{4.10}$$

Proof. For $f \in \mathcal{D}(A)$ and $t, h \geq 0$ we have

$$M_{t+h}^f - M_t^f = f(X_{t+h}) - f(X_t) - \int_t^{t+h} Af(X_s)\,ds = M_h^f \circ \theta_t,$$

where θ_t is the shift operator introduced in (3.7). So, the $\mathcal{F}_t$-$\mathcal{B}(\mathbb{R})$-measurability, the $\mathcal{F}_\infty^X$-$\mathcal{B}(\mathbb{R})$-measurability of M_t^f together with Proposition 3.14, Fubini's theorem, Proposition 3.16 and Remark 4.7 imply that $\mathbb{P}^\nu$-a.s.

$$\begin{aligned}
\mathbb{E}^\nu\left[M_{t+h}^f \mid \mathcal{F}_t\right] - M_t^f &= \mathbb{E}^\nu\left[M_h^f \circ \theta_t \mid \mathcal{F}_t\right] = \mathbb{E}^{X_t}\left[M_h^f\right] \\
&= \mathbb{E}^{X_t}\left[f(X_h)\right] - \mathbb{E}^{X_t}\left[f(X_0)\right] - \int_0^t \mathbb{E}^{X_t}\left[Af(X_s)\right]\,ds \\
&= T_h f(X_t) - f(X_t) - \int_0^t T_s Af(X_t)\,ds = 0.
\end{aligned}$$

Hence, M^f is a martingale.

From Doob's optional sampling theorem, e.g. Kallenberg [Kal02, Theorem 7.12], we obtain

$$0 = M_0^f = \mathbb{E}^x\left[M_\tau^f \mid \mathcal{F}_0\right],$$

whence follows by the law of total expectation and the definition of M^f that

$$0 = \mathbb{E}^x\left[f(X_\tau)\right] - \underbrace{\mathbb{E}^x\left[f(X_0)\right]}_{=f(x)} - \mathbb{E}^x\left[\int_0^\tau Af(X_s)\,ds\right],$$

which is equation (4.10). ∎

4.5 A stochastic description of the Generator

In the following we want to further describe some properties of the characteristic operator A of a Feller process. For this let ρ be a metric in E, which we may extend to a symmetric function on $\hat{E}$ into $\mathbb{R}_+ \cup \{\infty\}$ by setting $\rho(x, \Delta) = \infty$ for all $x \in \hat{E}$. For a Feller process X in $\hat{E}$ we introduce the optional times

$$\tau_h = \inf\{t \geq 0 \mid \rho(X_t, X_0) > h\}, \quad h > 0.$$

Note that a state $x \in E$ is absorbing for $(p_t)_{t\geq 0}$ iff for every $h > 0$ we have $\mathbb{P}^x$-a.s. $\tau_h = \infty$.

Lemma 4.19 (escape times). *Let X be a Feller process in $\hat{E}$. Then we have for any nonabsorbing state $x \in E$ that $\mathbb{E}^x[\tau_h] < \infty$ for all sufficiently small $h > 0$. If X has right-continuous paths then also $\mathbb{E}^x[\tau_h] > 0$ for every $x \in E$.*

Proof. To show the first assertion assume that x is absorbing. Then we have $p_t(x, B_\varepsilon(x)) = 1$ for all $t \geq 0$ and all $\varepsilon > 0$, where $B_\varepsilon(x) = \{y \in S \mid \rho(x, y) \leq \varepsilon\}$. Hence, for a nonabsorbing state x we find $\varepsilon > 0$, a time $t > 0$ and a constant $p < 1$ such that $p_t(x, B_\varepsilon(x)) < p$.

By Urysohn's lemma there exists a function $g \in C_0(E)$ with $1_{B_{\varepsilon/2}(x)} \leq g \leq 1_{B_\varepsilon(x)}$. Then we have for each $y \in E$

$$p_t(y, B_{\varepsilon/2}(x)) \leq \int_E g(z)\, p_t(y, dz) = T_t g(y) \leq p_t(y, B_\varepsilon(x)). \tag{4.11}$$

Set $\hat{\varepsilon} := p - p_t(x, B_\varepsilon(x))$. By continuity there exists $\delta > 0$ such that for all $y \in B_\delta(x)$

$$T_t g(y) - T_t g(x) \leq \hat{\varepsilon}.$$

Thus, choosing $h := \min\{\delta, \varepsilon/2\}$, equation (4.11) yields

$$\begin{aligned} &p_t(y, B_h(x)) - p_t(x, B_\varepsilon(x)) \\ \leq\, &\underbrace{p_t(y, B_{\varepsilon/2}(x)) - T_t g(y)}_{\leq 0} + \underbrace{T_t g(y) - T_t g(x)}_{\leq \hat{\varepsilon}} + \underbrace{T_t g(x) - p_t(x, B_\varepsilon(x))}_{\leq 0} \leq \hat{\varepsilon}, \end{aligned}$$

whence follows

$$p_t(y, B_h(x)) \leq p_t(x, B_\varepsilon(x)) + \hat{\varepsilon} = p, \quad y \in B_h(x).$$

Thus,

$$\begin{aligned} &\mathbb{P}^x(\{\tau_h > nt\}) = \mathbb{P}^x(\{\inf\{s \geq 0 \mid \rho(X_s, x) > h\} > nt\}) \\ &\leq \mathbb{P}^x(\{X_0 \in B_h(x), X_t \in B_h(x), \ldots, X_{nt} \in B_h(x)\}) \\ &= \int_{B_h(x)} \int_{B_h(x)} \cdots \int_{B_h(x)} 1\, p_t(x_{n-1}, dx_n) \ldots p_t(x_1, dx_2)\, p_t(x_0, dx_1)\, \delta_x(dx_0) \leq p^n. \end{aligned}$$

Using a Cavalieri principle type argument, see Kallenberg [Kal02, Lemma 3.4], and the monotonicity $\{\tau_h > s\} \subset \{\tau_h > t\}$ for $t \leq s$, we get

$$\mathbb{E}^x[\tau_h] = \int_0^\infty \mathbb{P}^x(\{\tau_h > s\})\, ds \leq t \sum_{n\geq 0} \mathbb{P}^x(\{\tau_h \geq nt\}) \leq t \sum_{n \geq 0} p^n = \frac{t}{1-p} < \infty.$$

Let $x \in E$. If $\mathbb{E}^x[\tau_h] = 0$ then, as $\tau_h \geq 0$, we have $\mathbb{P}^x(\{\tau_h = 0\}) = 1$. But then there exists a sequence $(t_n)_{n\in\mathbb{N}}$ in $\mathbb{R}_{>0}$ converging to 0 such that $\mathbb{P}^x$-a.s.

$$\mathbb{P}^x(\{\rho(X_{t_n}, X_0) > h\}) = \mathbb{P}^x(\{\rho(X_{t_n}, x) > h\}) = 1, \quad n \in \mathbb{N}.$$

This contradicts the right-continuity of X. ∎

The next theorem will give a probabilistic description of the generator of a Feller process. An operator is said to be *maximal* in a class of operators if it extends every operator of that class.

Theorem 4.20 (probabilistic characteristic operator, Dynkin)**.** *Let X be a Feller process in $\hat{E}$ with right-continuous paths and generator $(A, \mathcal{D}(A))$. Let $f \in \mathcal{D}(A)$ and $x \in E$. If x is absorbing for $(p_t)_{t\geq 0}$, we have $Af(x) = 0$, and otherwise*

$$Af(x) = \lim_{h\to 0^+} \frac{\mathbb{E}^x[f(X_{\tau_h})] - f(x)}{\mathbb{E}^x[\tau_h]}. \tag{4.12}$$

Moreover, A is the maximal operator satisfying these properties.

Proof. Let $f \in \mathcal{D}(A)$ and $x \in E$ be absorbing. Then we have $T_t f(x) = f(x)$ for all $t \geq 0$, and so $Af(x) = 0$. If x is nonabsorbing, Lemma 4.18 yields

$$\mathbb{E}^x[f(X_{\tau_h\wedge t})] - f(x) = \mathbb{E}^x\left[\int_0^{\tau_h\wedge t} Af(X_s)\,ds\right], \quad t, h > 0. \tag{4.13}$$

By Lemma 4.19 we have $\mathbb{P}^x$-a.s. $\tau_h < \infty$ for sufficiently small h. So equation (4.13) extends by dominated convergence to $t = \infty$. Note for all $h > 0$ we have $\mathbb{P}^x$-a.s. that for all $s < \tau_h$

$$\rho(X_s, x) \leq h,$$

and thus, it follows from the continuity of Af that for every $\varepsilon > 0$ there is exists $h > 0$ sufficiently small such that

$$\left|\mathbb{E}^x[f(X_{\tau_h})] - f(x) - \mathbb{E}^x\left[\int_0^{\tau_h} Af(x)\,ds\right]\right| \leq \varepsilon \cdot \mathbb{E}^x[\tau_h],$$

which implies (4.12), as by right-continuity and Lemma 4.19, we may divide by $\mathbb{E}^x[\tau_h]$.

Assume there exists a linear operator $(A', \mathcal{D}(A'))$ that extends A and satisfies these properties. Let $f \in \mathcal{D}(A')$ be such that there exists a point $x \in E$ with $f^+ = f \ \vee 0 \leq f(x)$. Then, by equation (4.12), we have $A'f(x) \leq 0$. So A' satisfies the positive-maximum principle, which implies $A = A'$, see Lemma 4.10. ∎

4.6 Feller Diffusions

A further description of the generator $(A, \mathcal{D}(A))$ of a Feller process X in $\hat{E}$ can be given for processes with $E = \mathbb{R}^d$.

First we will need some new nomenclature. Denote by $C_c^\infty(\mathbb{R}^d)$ the set of all infinitely differentiable functions on $\mathbb{R}^d$ with bounded support. We call an operator $(A, \mathcal{D}(A))$ *local* *on $\mathcal{D}$*, if $\mathcal{D} \subset \mathcal{D}(A)$ and we have $Af(x) = 0$ whenever $f \in \mathcal{D}$ vanishes in a neighborhood of x. We say that a linear operator $(A, \mathcal{D}(A))$ satisfies a *local positive-maximum principle on $\mathcal{D}$* if $\mathcal{D} \subset \mathcal{D}(A)$ and we have $Af(x) \leq 0$ whenever $f \in \mathcal{D}$ has a local maximum at $x \in \mathbb{R}^d$ with $f(x) \geq 0$.

Lemma 4.21 (local positive-maximum principle)**.** *Let $(A, \mathcal{D}(A))$ be a linear operator on $C_0(\mathbb{R}^d)$ that is local on $C_c^\infty(\mathbb{R}^d)$ and satisfies the positive-maximum principle. Then A also satisfies the local positive-maximum principle on $C_c^\infty(\mathbb{R}^d)$.*

Proof. Let $f \in C_c^\infty(\mathbb{R}^d)$ be a function that has a local maximum at $x \in \mathbb{R}^d$ such that $f(x) \geq 0$. Choose a function $\tilde{f} \in C_c\infty(\mathbb{R}^d)$ such that $\tilde{f}$ agrees with f on some neighborhood U of x and satisfies $\tilde{f} \leq f(x)$. It can be seen that such a function exists by choosing two open bounded neighborhoods U and V of x with $\overline{U} \subset V$ small enough such that $f(y) \leq f(x)$ for all $y \in V$. By a smooth version of Urysohn's Lemma, e.g. Lee [Lee03, Proposition 2.25], we find a function ψ with $\psi \equiv 1$ on U and $\operatorname{supp}(\psi) \subset V$. Now set $\tilde{f} := f\psi$. As $f - \tilde{f}$ vanishes on U we obtain by the positive-maximum principle and the locality of A that

$$Af(x) = \underbrace{A\tilde{f}(x)}_{\leq 0} + \underbrace{A(f - \tilde{f})(x)}_{=0} \leq 0.$$

So A satisfies the local positive-maximum principle. ■

The next result establishes a connection between nonnegative differential operators and the generator of Feller processes. This is Theorem 19.24 in Kallenberg [Kal02].

Theorem 4.22 (Feller diffusions and nonnegative differential operators, Dynkin)**.** *Let $E = \mathbb{R}^d$ and let $(A, \mathcal{D}(A))$ be the generator of a Feller process X in $\hat{E}$ that has càdlàg paths. Assume that $C_c^\infty(\mathbb{R}^d) \subset \mathcal{D}(A)$. Then X is $\mathbb{P}^\nu$-a.s. continuous on $[0, \zeta)$ for every initial distribution ν on $\hat{\mathcal{B}}$ iff A is local on $C_c^\infty(\mathbb{R}^d)$. In this case there exist continuous functions $a : \mathbb{R}^d \to \mathbb{R}^{d \times d}$, $b : \mathbb{R}^d \to \mathbb{R}^d$ and $c : \mathbb{R}^d \to \mathbb{R}$, such that a takes nonnegative definite values, $c \geq 0$ and for every $f \in C_c^\infty(\mathbb{R}^d)$, $x \in \mathbb{R}^d$ we have*

$$Af(x) = \frac{1}{2} \sum_{i,j=1}^{d} a_{ij}(x) \partial_{x_i, x_j} f(x) + \nabla f(x) b(x) - c(x) f(x). \tag{4.14}$$

Proof. <u>1. Step:</u> We show that X is $\mathbb{P}^\nu$-a.s. continuous on $[0, \zeta)$ iff A is local.

Let X be $\mathbb{P}^\nu$-a.s. continuous on $[0, \zeta)$ for all initial distributions ν and assume that $f \in \mathcal{D}(A)$ vanishes in some neighborhood U of $x \in \mathbb{R}^d$. For $h > 0$ such that $B_h(x) \subset U$ we have $\mathbb{P}^x$-a.s. $f(X_{\tau_h}) = 0$, and thus, Theorem 4.20 implies $Af(x) = 0$. So, A is local on $C_c^\infty(\mathbb{R}^d)$.

Conversely, assume that A is local on $C_c^\infty(\mathbb{R}^d)$. Denote by ρ the metric induced by some norm $|\cdot|$ on $\mathbb{R}^d$ and extend ρ to $\hat{E}$ as in Subsection 4.5. First we show that

$$\mathbb{P}^x(\{\rho(X_{\tau_h}, x) \leq h\} \cup \{X_{\tau_h} = \Delta\}) = 1, \quad x \in \mathbb{R}^d \cup \{\Delta\},\, h > 0. \tag{4.15}$$

Assume that (4.15) does not hold for $x \in \mathbb{R}^d$ and a real number $h > 0$. By the monotone convergence theorem, there exists an $m > h$ such that $\mathbb{P}^x(\{X_{\tau_h} \in \mathbb{R}^d,\ h < |X_{\tau_h} - x| < m\}) > 0$. And by the dominated convergence theorem we find $f \in C_c^\infty(\mathbb{R}^d)$ with $f \geq 0$ and support in the closed annulus $\{y \in \mathbb{R}^d \mid h \leq |y - x| \leq m\}$ with center x such that $\mathbb{E}^x[f(X_{\tau_h})] > 0$. Then the locality of A on $C_c^\infty(\mathbb{R}^d)$ yields $Af(y) = 0$ for all $y \in B_h(x)$. Thus, an application of Dynkin's formula, see Lemma 4.18, and Fubini's theorem yield

$$\mathbb{E}^x[f(X_{\tau_h \wedge t})] = f(x) + \int_0^{\tau_h \wedge t} \mathbb{E}^x[Af(X_s)]\, ds = 0, \quad t \geq 0.$$

As by Lemma 4.19 $\mathbb{P}^x$-a.s. $\tau_h < \infty$, it follows from Lebesgue's theorem that for $t \to \infty$ we have $\mathbb{E}^x[f(X_{\tau_h})] = 0$. This is a contradiction. So, equation (4.15) must hold for all $x \in \mathbb{R}^d$. For $x = \Delta$ we obtain $\mathbb{P}^\Delta$-a.s. that $\tau_h = 0$, and so, $\mathbb{P}^\Delta(\{\rho(X_{\tau_h}, \Delta) \leq h\} \cup \{X_{\tau_h} = \Delta\}) = \mathbb{P}^\Delta(\{X_0 = \Delta\}) = 1$.

For $t \geq 0$ and $h > 0$ set

$$B_{t,h} := \theta_t^{-1}(\{\rho(X_{\tau_h}, X_0) \leq h\} \cup \{X_{\tau_h} = \Delta\}).$$

Note that $B_{t,h} \in \mathcal{F}_\infty^X$. Using the law of total expectation, the extended Markov property given in Proposition 3.14 and equation (4.15), we now obtain for every $t \geq 0$ and every initial distribution ν

$$\begin{aligned}\mathbb{P}^\nu(B_{t,h}) &= \mathbb{P}^\nu(\theta_t^{-1}(B_{0,h})) = \mathbb{E}^\nu\left[1_{B_{0,h}} \circ \theta_t\right] = \mathbb{E}^\nu\left[\mathbb{E}^\nu\left[1_{B_{0,h}} \circ \theta_t \mid \mathcal{F}_t\right]\right] \\ &= \mathbb{E}^\nu\left[\mathbb{E}^{X_t}\left[1_{B_{0,h}}\right]\right] = \mathbb{E}^\nu\left[\mathbb{P}^{X_t}(B_{0,h})\right] = 1,\end{aligned}$$

and thus,

$$\mathbb{P}^\nu\left(\bigcap_{t \in \mathbb{Q}_+} B_{t,h}\right) = 1, \quad h > 0.$$

We aim to show that for every $h > 0$

$$B := \bigcap_{t \in \mathbb{Q}_{\geq 0}} B_{t,h} \subset \{\omega \in \Omega_0 \mid \zeta(\omega) > 0,\ \sup_{t < \zeta(\omega)} |X_t(\omega) - X_{t-}(\omega)| \leq h\} \cup \{\zeta = 0\}, \quad (4.16)$$

which together with the right continuity of X yields the $\mathbb{P}^\nu$-a.s. continuity on $[0, \zeta)$. Fix $h > 0$, let $\omega \in B$ and assume that ω is no element of the right-hand side of (4.16). Then in particular $\zeta(\omega) > 0$ and and for some $t \in (0, \zeta(\omega))$

$$|X_t(\omega) - X_{t-}(\omega)| > h.$$

As X has left limits, we may choose $s \in \mathbb{Q}_{\geq 0}$ with $s \leq t$ such that for all $\tilde{s} \in [s, t)$

$$|X_t(\omega) - X_s(\omega)| > h \quad \text{and} \quad |X_{\tilde{s}}(\omega) - X_s(\omega)| < h. \quad (4.17)$$

Also, as $\omega \in B_{s,h}$,

$$\rho(X_{\tau_h(\theta_s\omega)}(\theta_s\omega), X_0(\theta_s\omega)) \leq h \quad \text{or} \quad X_{\tau_h(\theta_s\omega)}(\theta_s\omega) = \Delta.$$

In case the first equality is fulfilled we obtain that $X_{\tau_h(\theta_s\omega)+s}(\omega), X_s(\omega) \in \mathbb{R}^d$ and

$$|X_{\tau_h(\theta_s\omega)+s}(\omega) - X_s(\omega)| \leq h. \tag{4.18}$$

Note $\tau_h \circ \theta_s = \inf\{t \geq 0 \mid \rho(X_{t+s}, X_s) > h\}$. So the first inequality in (4.17) implies $\tau(\theta_s\omega) \leq t-s$. Now assume that $\tau(\theta_s\omega) < t-s$. Then $\tau_h(\theta_s\omega)+s \in [s,t)$ and the second inequality in (4.17) implies $|X_{\tau_h(\theta_s\omega)+s}(\omega) - X_s(\omega)| < h$ but from the definition of τ_h and the right-continuity of X we obtain $|X_{\tau_h(\theta_s\omega)+s}(\omega) - X_s(\omega)| \geq h$. So $\tau_h(\theta_s\omega) = t-s$ and from (4.17) and (4.18)

$$h < |X_t(\omega) - X_s(\omega)| = |X_{\tau_h(\theta_s\omega)+s}(\omega) - X_s(\omega)| \leq h.$$

Thus, we must have that $X_{\tau_h(\theta_s\omega)+s}(\omega) = X_{\tau_h(\theta_s\omega)}(\theta_s\omega) = \Delta$. Then $\tau_h(\theta_s\omega)+s \geq \zeta(\omega) > t$ and so $\tau_h(\theta_s\omega) > t-s$, which we have shown to contradict the first equality in (4.17). For h was arbitrarily chosen, this proves (4.16), and thus, the P^ν-a.s. continuity of X on $[0,\zeta)$.

2. Step: Assuming the locality of A, we show the representation (4.14).

Fix an open bounded neighborhood U and choose for each $x \in U$ some functions f_0^x, f_i^x, $f_{ij}^x \in C_c^\infty$ with

$$f_0^x(y) = 0, \quad f_i^x(y) = y_i - x_i, \quad f_{ij}^x(y) = (y_i - x_i)(y_j - x_j), \quad y \in U,\, i,\, j \in \{1,\dots,d\}.$$

Recall that such functions may be constructed using a smooth version of Urysohn's Lemma, e.g. Lee [Lee03, Proposition 2.25]. Set

$$c(x) = -Af_0^x(x), \quad b(x) = (Af_i^x(x))_{i\in\{1,\dots,d\}}, \quad a(x) = \left(Af_{ij}^x(x)\right)_{i,j\in\{1,\dots,d\}}, \quad x \in U.$$

By the locality of A the coefficients are well-defined, i.e., the values $a(x)$, $b(x)$ and $c(x)$ do not depend on values of f_0^x, f_i^x and $f_{i,j}^x$ outside of some arbitrarily small open neighborhood of x. Note that $a(x)$ is symmetric.

For functions $f_0, f_i, f_{i,j} \in C_c^\infty(\mathbb{R}^d)$ satisfying $f_0(x) = 1$, $f_i(x) = y_i$ and $f_{i,j}(x) = y_i y_j$ for all $x \in U$, $i, j \in \{1,\dots,d\}$ we obtain from the locality of A that

$$Af_0(x) - Af_0^x(x) = A(f_0 - f_0^x)(x) = 0, \quad x \in U,$$

and thus, $Af_0(x) = -c(x)$, $x \in U$. Similarly, we obtain for $x \in U$ and $i, j \in \{1,\dots,d\}$

$$\begin{aligned} Af_i(x) &= b_i(x) - x_i c(x), \\ Af_{ij}(x) &= a_{ij}(x) + x b_j(x) + x_j b_i(x) - x_i x_j c(x), \end{aligned}$$

whence we obtain continuity of a, b and c on U.

Let $x \in U$. The local positive-maximum principle applied to f_0^x now yields that $c(x) \geq 0$. The same principle applied to the function

$$f = -\left(\sum_{i=1}^d u_i f_i^x\right)^2 = -\sum_{i,j=1}^d u_i u_j f_{ij}^x,$$

where $(u_1, \dots, u_d) \in \mathbb{R}^d$, we obtain $\sum_{i,j=1}^d u_i u_j A f_{ij}^x \geq 0$, as f has a nonnegative maximum in x. This shows that a only takes on nonnegative definite values.

Let $f \in C_c^\infty(\mathbb{R}^d)$ with a second-order Taylor expansion $\tilde{f}$ around x. Then functions $g_+^\varepsilon, g_-^\varepsilon \in C_c^\infty(\mathbb{R}^d)$ satisfying

$$g_\pm^\varepsilon(y) = \pm(f(y) - \tilde{f}(y)) - \varepsilon \underbrace{(x-y)^\top (x-y)}_{=|x-y|^2}, \quad y \in U,\, \varepsilon > 0$$

have local maxima 0 at x. Hence, the local positive-maximum principle implies

$$A g_\pm^\varepsilon(x) = \pm(Af(x) - A\tilde{f}(x)) - \varepsilon \sum_{i=1}^d a_{ii}(x) \leq 0, \quad \varepsilon > 0.$$

As $\varepsilon \to 0$ we obtain $|Af(x) - A\tilde{f}(x)| \leq 0$, and so $Af(x) = A\tilde{f}(x)$. Since x was chosen arbitrarily, this shows equation (4.14).

■

To illustrate the last result we give an example of a well-known diffusion process.

Example 4.23 (Brownian motion)**.** For the conservative transition semigroup

$$T_t f(x) := \int_{\mathbb{R}} f(y) \frac{1}{\sqrt{2\pi t}} e^{-\frac{(y-x)^2}{2t}} dy, \quad t > 0,\, f \in B_b(\mathbb{R}),\, x \in E,$$

the restriction to $C_0(\mathbb{R})$ exists and is a Feller semigroup. Moreover, its generator A is local on $C_c^\infty(\mathbb{R})$ and satisfies

$$Af = \frac{1}{2}\Delta f, \quad f \in C_c^\infty(\mathbb{R}),$$

where Δ denotes the Laplace operator.

As $(T_t)_{t\geq 0}$ is conservative we may choose the Feller process in $\mathbb{R}$, so we have $\zeta = \infty$, $\mathbb{P}^\nu$-a.s., for all initial distributions ν on $\mathcal{B}(\mathbb{R})$. The corresponding Feller process can be chosen continuous and is the triple $((W_t)_{t\geq 0}, \mathcal{F}^W, \{\mathbb{P}^x\}_{x\in E})$, where $((W_t)_{t\geq 0}, \mathcal{F}^W, \mathbb{P}^x)$ are Brownian motions with initial point x.

For $f \in C_0(\mathbb{R})$ we have

$$\frac{d}{dt} T_t f = \frac{1}{2} \Delta T_t f, \quad t \geq 0.$$

So the map $[0, \infty) \ni t \mapsto T_t f$ is a solution of the heat equation

$$\begin{cases} \dfrac{\partial}{\partial t} u(t,x) = \dfrac{1}{2} \Delta_x u(t,x), \quad t \geq 0,\, x \in \mathbb{R}, \\ u(0,x) = f(x), \end{cases}$$

with boundary condition imposed by the choice of the function space $C_0(\mathbb{R})$.

5 Notes and References

Many modern textbooks on probability theory include some treatment of Markov processes but few are as detailed as the books [Dyn61; Dyn65] of Dynkin. In addition to the treatment of time-homogeneous Markov processes, in [Dyn61] also includes a detailed treatment of the inhomogeneous case. In [Dyn65] a detailed connection of Markov processes and transition functions is drawn. Consequentially, many results of Section 3 can be found in these books. In particular, the definitions of Markov processes and Markov random functions are based on these given by Dynkin. In younger textbooks these concepts are sometimes also referred to as Markov family resp. Markov process, e.g. Karatzas and Shreve [KS14].

Although, Corollary 3.6 was inspired by Theorem 1.2.1 in Dynkin [Dyn65], unfortunately, the conditions on the semigroup of the latter result are to weak for the statement to hold true. This was corrected in Corollary 3.6 by also assuming $\sigma(B_b(E), \mathrm{ca}(\mathcal{B}))$-continuity for which we needed the treatment of weak topologies in Section 2.2. For more information on when linear operator are transition or kernel operators we refer to Kunze [Kun11].

Contrary to the order presented in Section 4, historically, the treatment of Feller processes has begun by investigating differential diffusion equations and their connection to semigroups, e.g. Feller [Fel52]. The modern treatment given here is due to Dynkin [Dyn65] and Kallenberg [Kal02]. In addition to the definition of Feller processes presented here the literature knows many deviations, variants and generalizations that were not presented here.

Due to this paper being supposed to be a brief overview of the topic, we haven't been able to present some interesting traits associated to Feller processes. To name a few: the strong Markov property of the canonical Feller processes and that the filtration of the canonical Feller processes can be chosen right-continuous.

References

[Arn98] Ludwig Arnold. *Random Dynamical Systems*. Springer Monographs in Mathematics. Berlin Heidelberg: Springer-Verlag, 1998.

[Bil68] Patrick Billingsley. *Convergence of Probability Measures*. Wiley Series in Probability and Mathematical Statistics. Wiley, 1968.

[Bil95] Patrick Billingsley. *Probability and Measure*. Anniversary. Wiley Series in Probability and Mathematical Statistics. Wiley, 1995.

[Dyn61] Evgenij B. Dynkin. *Die Grundlagen der Theorie der Markoffschen Prozesse*. Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen mit besonderer Berücksichtigung der Anwendungsgebiete Bd. 1. Academic Press, 1961.

[Dyn65] Evgenij B. Dynkin. *Markov Processes I*. Springer Science & Business Media, 1965.

[Els05] Jürgen Elstrodt. *Maß- und Integrationstheorie*. 4th edition. Springer-Verlag, 2005.

[EN06] Klaus-Jochen Engel and Rainer Nagel. *One-Parameter Semigroups for Linear Evolution Equations*. Springer Science & Business Media, 2006.

[Fel52] William Feller. “The Parabolic Differential Equations and the Associated Semi-Groups of Transformations”. In: *Annals of Mathematics* 55.3 (1952). Publisher: Annals of Mathematics, pp. 468–519.

[GS77] Peter Gänssler and Winfried Stute. *Wahrscheinlichkeitstheorie*. 1st edition. Hochschultext. Springer-Verlag, 1977.

[Kal02] Olav Kallenberg. *Foundations of Modern Probability*. 2nd edition. Probability and Its Applications. Springer New York, 2002.

[KS14] Ioannis Karatzas and Steven E. Shreve. *Brownian Motion and Stochastic Calculus*. Graduate Texts in Mathematics. Springer New York, 2014.

[Kun11] Markus Kunze. “A Pettis-type integral and applications to transition semigroups”. In: *Czechoslovak Mathematical Journal* 61.2 (2011), pp. 437–459.

[Lee03] John M. Lee. *Introduction to Smooth Manifolds*. 2nd edition. Graduate Texts in Mathematics. Springer, 2003.

[Man89] Mark Mandelkern. “Metrization of the one-point compactification”. In: *Proceedings of the American Mathematical Society* 107.4 (1989), pp. 1111–1115.

[Ped89] Gert K. Pedersen. *Analysis Now*. 2nd edition. Graduate Texts in Mathematics. Springer-Verlag, 1989.

[Wer18] Dirk Werner. *Funktionalanalysis*. Springer-Lehrbuch. Berlin, Heidelberg: Springer Berlin Heidelberg, 2018.

Index

Eidesstattliche Erklärung

Hiermit erkläre ich, **Julian Hölz** (mit der Matrikelnummer 77903), gegenüber der Fakultät für Informatik und Mathematik der Universität Passau, dass ich die Seminararbeit mit dem Thema “Markov and Feller Processes” selbstständig angefertigt und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Alle wörtlich und sinngemäß übernommenen Ausführungen wurden als solche gekennzeichnet. Weiterhin erkläre ich, dass ich diese Arbeit in gleicher oder ähnlicher Form nicht bereits einer anderen Prüfungsbehörde vorgelegt habe.

Passau, den 29. Juli 2020

Julian Hölz