Chapter 2
Markov Processes

In this chapter, we provide a background material that is needed to define and study Markov processes in discrete and continuous time. We start by giving basic examples of transition probabilities and the corresponding operators on the spaces of functions and measures. An emphasis is put on stochastic operators on the spaces of integrable functions. The importance of transition probabilities is that the distribution of a stochastic process with Markov property is completely determined by transition probabilities and initial distributions. The Markov property simply states that the past and the future are independent given the present. We refer the reader to Appendix A for the required theory on measure, integration, and basic concepts of probability theory.

2.1 Transition Probabilities and Kernels

2.1.1 Basic Concepts

In this section, we introduce transition kernels, stressing the interplay between analytic and stochastic interpretations. Let X be a set and let Σ be a σ-algebra of subsets of X. The pair (X, Σ) is called a measurable space. A function $P\colon X \times \Sigma \to [0, 1]$ is said to be a *transition kernel* if

(1) for each set $B \in \Sigma$, $P(\cdot, B)\colon X \to [0, 1]$ is a measurable function;
(2) for each $x \in X$ the set function $P(x, \cdot)\colon \Sigma \to [0, 1]$ is a measure.

If $P(x, X) = 1$ for all $x \in X$ then P is said to be a *transition probability* on (X, Σ).

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space. A mapping $\xi\colon \Omega \to X$ is said to be an *X-valued random variable* if it is measurable, i.e. for any $B \in \Sigma$

R. Rudnicki and M. Tyran-Kamińska, *Piecewise Deterministic Processes in Biological Models*, SpringerBriefs in Mathematical Methods,
DOI 10.1007/978-3-319-61295-9_2

$$\{\xi \in B\} = \xi^{-1}(B) = \{\omega \in \Omega : \xi(\omega) \in B\} \in \mathscr{F}.$$

The *(probability) distribution* or the *(probability) law* of ξ is a probability measure μ_ξ on (X, Σ) defined by

$$\mu_\xi(B) = \mathbb{P}(\xi \in B), \quad B \in \Sigma.$$

Given an X-valued random variable ξ on the probability space $(\Omega, \mathscr{F}, \mathbb{P})$ the family

$$\sigma(\xi) = \{\xi^{-1}(B) : B \in \Sigma\}$$

is a σ-algebra of subsets of Ω and it is called the *σ-algebra generated by* ξ. Note that ξ is an X-valued random variable if and only if $\sigma(\xi) \subseteq \mathscr{F}$. If $\mathscr{G}$ is a sub-σ-algebra of $\mathscr{F}$ then ξ is said to be $\mathscr{G}$-measurable if $\sigma(\xi) \subseteq \mathscr{G}$. Thus one can also say that $\sigma(\xi)$ is the smallest σ-algebra $\mathscr{G}$ such that ξ is $\mathscr{G}$-measurable. Recall that sub-σ-algebras $\mathscr{F}_1$ and $\mathscr{F}_2$ of $\mathscr{F}$ are called *independent* if $\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1)\mathbb{P}(A_2)$ for any $A_1 \in \mathscr{F}_1$ and $A_2 \in \mathscr{F}_2$. If a random variable ϑ with values in any measurable space is *independent of the σ-algebra* $\mathscr{G}$, i.e. if $\sigma(\vartheta)$ and $\mathscr{G}$ are independent, and if ξ is an $\mathscr{G}$-measurable X-valued random variable, then the conditional expectation of $g(\xi, \vartheta)$ with respect to $\mathscr{G}$ is given by

$$\mathbb{E}(g(\xi, \vartheta)|\mathscr{G}) = \mathbb{E}(g(\xi, \vartheta)|\xi) \quad \text{and} \quad \mathbb{E}(g(\xi, \vartheta)|\xi = x) = \mathbb{E}(g(x, \vartheta)) \tag{2.1}$$

for all $x \in X$ and for any measurable non-negative function g; see Lemma A.3.

Let Y be a metric space. Consider a probability measure ν on the Borel σ-algebra $\mathscr{B}(Y)$. Given a measurable mapping $\kappa : X \times Y \to X$, we define

$$P(x, B) = \int \mathbf{1}_B(\kappa(x, y))\nu(dy), \quad x \in X, \quad B \in \Sigma, \tag{2.2}$$

where $\mathbf{1}_B$ is the indicator function, equal to one on the set B and zero otherwise. Clearly, P is a transition probability. For a Y-valued random variable ϑ with distribution ν and for every $x \in X$, the X-valued random variable $\kappa(x, \vartheta)$ has distribution $P(x, \cdot)$, since

$$\mathbb{P}(\kappa(x, \vartheta) \in B) = \mathbb{E}(\mathbf{1}_B(\kappa(x, \vartheta))) = \int \mathbf{1}_B(\kappa(x, y))\nu(dy) = P(x, B), \quad B \in \Sigma.$$

As a particular example, we can take Y to be the unit interval $[0, 1]$ and ν to be the Lebesgue measure on $[0, 1]$; in that case ϑ is said to be uniformly distributed on the unit interval $[0, 1]$. Now let ξ_0 be an X-valued random variable independent of ϑ and let $\xi_1 = \kappa(\xi_0, \vartheta)$. We can write

$$\mathbb{P}(\xi_1 \in B|\xi_0 = x) = \mathbb{P}(\kappa(\xi_0, \vartheta) \in B|\xi_0 = x) = \mathbb{E}(\mathbf{1}_B(\kappa(\xi_0, \vartheta))|\xi_0 = x)$$

and if we take $g(x, y) = \mathbf{1}_B(\kappa(x, y))$ in (2.1) then $\mathbb{E}(g(x, \vartheta)) = P(x, B)$. Thus

$$\mathbb{P}(\xi_1 \in B | \xi_0 = x) = P(x, B),$$

which gives the *conditional distribution* of ξ_1 given ξ_0. Moreover, if μ_{ξ_0} is the distribution of ξ_0 then the joint distribution $\mu_{(\xi_0, \xi_1)}$ of ξ_0 and ξ_1 satisfies

$$\mu_{(\xi_0, \xi_1)}(B_0 \times B_1) = \mathbb{P}(\xi_0 \in B_0, \xi_1 \in B_1) = \int_{B_0} P(x, B_1) \mu_{\xi_0}(dx)$$

for all $B_0, B_1 \in \Sigma$.

We need an extension of the concept of transition probabilities. Let (X_0, Σ_0) and (X_1, Σ_1) be two measurable spaces. A function $P : X_0 \times \Sigma_1 \to [0, 1]$ is said to be a *transition probability from* (X_0, Σ_0) *to* (X_1, Σ_1) if $P(\cdot, B)$ is measurable for every $B \in \Sigma_1$ and $P(x, \cdot)$ is a probability measure on (X_1, Σ_1) for every $x \in X_0$. The next result ensures existence of probability measures on product spaces.

Proposition 2.1 *Let P be a transition probability from (X_0, Σ_0) to (X_1, Σ_1). If μ_0 is a probability measure on (X_0, Σ_0) then there exists a unique probability measure μ on the product space $(X_0 \times X_1, \Sigma_0 \otimes \Sigma_1)$ such that*

$$\mu(B_0 \times B_1) = \int_{B_0} P(x_0, B_1)\, \mu_0(dx_0), \quad B_0 \in \Sigma_0, B_1 \in \Sigma_1.$$

Proof Let $B^{x_0} = \{x_1 \in X_1 : (x_0, x_1) \in B\}$ for every $x_0 \in X_0$ and $B \in \Sigma_0 \otimes \Sigma_1$. Clearly, if $B = B_0 \times B_1$ then we have $B^{x_0} = B_1$ for $x_0 \in B_0$ and $B^{x_0} = \emptyset$ for $x_0 \notin B_0$. By the monotone class theorem (see Theorem A.1) for each $B \in \Sigma_0 \otimes \Sigma_1$ the function $x_0 \mapsto P(x_0, B^{x_0})$ is measurable. Define

$$\mu(B) = \int_{X_0} P(x_0, B^{x_0})\, \mu_0(dx_0), \quad B \in \Sigma_0 \otimes \Sigma_1,$$

which is clearly a measure on the product space. Since the product σ-algebra is generated by rectangles, i.e. $\Sigma_0 \otimes \Sigma_1 = \sigma(\mathscr{C})$, where $\mathscr{C} = \{B_0 \times B_1 : B_0 \in \Sigma_0, \ B_1 \in \Sigma_1\}$, and $\mathscr{C}$ is a π-system, the result follows from the $\pi - \lambda$ lemma.

We now show that, for a large class of measurable spaces, any transition probability is as in (2.2). We call a measurable space (X, Σ) a *Borel space* if X is isomorphic to a Borel subset of $[0, 1]$, i.e. there exists a Borel subset Y of $[0, 1]$ and a measurable bijection $\psi : X \to Y$ such that its inverse ψ^{-1} is also measurable. Any Borel subset of a *Polish space*, i.e. a complete separable metric space, with $\Sigma = \mathscr{B}(X)$ is a Borel space (see e.g. [55, Theorem A1.2]). Recall that a metric space (X, ρ) is *complete* if every sequence (x_n) satisfying the Cauchy condition $\lim_{m,n\to\infty} \rho(x_n, x_m) = 0$ is convergent to some $x \in X$: $\lim_{n\to\infty} \rho(x_n, x) = 0$, and it is *separable* if there is a countable set $X_0 \subset X$ which is dense in X, so that for any $x \in X$ there exists a sequence from X_0 converging to x.

Proposition 2.2 *Let P be a transition probability from a measurable space* (X_0, Σ_0) *to a Borel space* (X_1, Σ_1). *Then there exists a measurable function* $\kappa : X_0 \times [0, 1] \to X_1$ *such that if a random variable ϑ is uniformly distributed on the unit interval* $[0, 1]$ *then $\kappa(x, \vartheta)$ has distribution $P(x, \cdot)$ for every $x \in X_0$.*

Proof Since $\psi \colon X_1 \to \psi(X_1)$ with $\psi(X_1) \subset [0, 1]$ is a bijection and we have $\psi^{-1}([0, t]) = \psi^{-1}([0, t] \cap \psi(X_1))$, we can define

$$F_x(t) = P(x, \psi^{-1}([0, t])), \quad t \in (0, 1].$$

For each x the function $t \mapsto F_x(t)$ is right-continuous with $F_x(1) = 1$. Consider its generalized inverse

$$F_x^{\leftarrow}(q) := \inf\{t \geq 0 \colon F_x(t) \geq q\}, \quad q \geq 0,$$

with the convention that the infimum of an empty set is equal to $+\infty$, and define $\kappa(x, q) = \psi^{-1}(F_x^{\leftarrow}(q))$ if $F_x^{\leftarrow}(q) \in \psi(X_1)$ and $\kappa(x, q) = x_1$ if $F_x^{\leftarrow}(q) \notin \psi(X_1)$, where x_1 is an arbitrary point from X_1. We have

$$\mathrm{Leb}\{q \in [0, 1] \colon \kappa(x, q) \in B\} = P(x, B) \quad \text{for } x \in X_0,\ B \in \Sigma_1,$$

where Leb denotes the Lebesgue measure on $[0, 1]$.

2.1.2 Transition Operators

It is convenient to associate with each transition kernel two linear mappings acting in two Banach spaces, one which is a space of bounded functions and the other which is a space of measures.

Let P be a transition kernel on a measurable space (X, Σ). Consider first the Banach space $B(X)$ of bounded, measurable, real-valued functions on X with the supremum norm

$$\|g\|_u = \sup_{x \in X} |g(x)|.$$

Given $g \in B(X)$ we define

$$Tg(x) = \int_X g(y) P(x, dy), \quad x \in X.$$

If $B \in \Sigma$ then the indicator function $\mathbf{1}_B \in B(X)$ and

$$T\mathbf{1}_B(x) = P(x, B), \quad x \in X.$$

Thus $T\mathbf{1}_B \in B(X)$. By the monotone class theorem, we obtain $Tg \in B(X)$ for $g \in B(X)$. From linearity of the integral it follows that the mapping $B(X) \ni g \mapsto Tg \in B(X)$ is a linear operator and that

$$\|Tg\|_u \le \|g\|_u, \quad g \in B(X).$$

Now, let $\mathscr{M}^+(X)$ be the collection of all finite measures on Σ. For each $\mu \in \mathscr{M}^+(X)$ we define

$$P\mu(B) = \int_X P(x, B)\,\mu(dx), \quad B \in \Sigma.$$

Then $P\mu$ is a finite measure and $P\mu(X) \le \mu(X)$. The space $\mathscr{M}^+(X)$ is a *cone*, i.e. $c_1\mu_1 + c_2\mu_2 \in \mathscr{M}^+(X)$, if c_1, c_2 are non-negative constants and $\mu_1, \mu_2 \in \mathscr{M}^+(X)$, and we have

$$P(c_1\mu_1 + c_2\mu_2) = c_1 P\mu_1 + c_2 P\mu_2.$$

Therefore we can extend P so that it is a linear operator on the vector space of all finite signed measures

$$\mathscr{M}(X) = \{\mu_1 - \mu_2 \colon \mu_1, \mu_2 \in \mathscr{M}^+(X)\},$$

by setting $P\mu = P\mu_1 - P\mu_2$ for $\mu = \mu_1 - \mu_2 \in \mathscr{M}(X)$. Define the *total variation norm* on $\mathscr{M}(X)$ by

$$\|\mu\|_{TV} = |\mu|(X), \quad \mu \in \mathscr{M}(X),$$

where the *total variation* of a finite signed measure μ is given by

$$|\mu|(X) = \sup_{B\in\Sigma} \mu(B) - \inf_{B\in\Sigma} \mu(B) = \sup\{|\int g d\mu| \colon g \in B(X),\ \|g\|_u \le 1\}.$$

Then $(\mathscr{M}(X), \|\cdot\|_{TV})$ is a Banach space and we have

$$\|P\mu\|_{TV} \le \|\mu\|_{TV}, \quad \mu \in \mathscr{M}(X).$$

In particular, a transition probability P provides the following stochastic interpretation. If X-valued random variables ξ_0 and ξ_1 are such that

$$\mathbb{P}(\xi_1 \in B | \xi_0 = x) = P(x, B)$$

and ξ_0 has a distribution μ_0, then $\mu_1 = P\mu_0$ is the distribution of ξ_1 and

$$\mathbb{E}(g(\xi_1)|\xi_0 = x) = \int_X g(y) P(x, dy) = Tg(x), \quad x \in X,\ g \in B(X).$$

2.1.3 Substochastic and Stochastic Operators

In this section, we suppose that a measurable space (X, Σ) carries a σ-finite measure m. The set of all measurable and m-integrable functions is a linear space and it becomes a Banach space, written $L^1(X, \Sigma, m)$, by defining the norm

$$\|f\| = \int_X |f(x)|\, m(dx), \quad f \in L^1(X, \Sigma, m),$$

and by identifying functions that are equal almost everywhere. Given a measurable function $f: X \to \mathbb{R}$ the *essential supremum* of f is defined by

$$\operatorname{ess\,sup}|f| = \inf\{c > 0\colon\ m\{x\colon |f(x)| > c\} = 0\} =: \|f\|_\infty.$$

The set of all functions with a finite essential supremum is denoted by $L^\infty(X, \Sigma, m)$ and it becomes a Banach space when we identify functions that are equal almost everywhere and take $\|\cdot\|_\infty$ as the norm.

A linear operator $P\colon L^1(X, \Sigma, m) \to L^1(X, \Sigma, m)$ is called *substochastic* if it is *positive*, i.e. if $f \ge 0$ then $Pf \ge 0$ for $f \in L^1(X, \Sigma, m)$, and if $\|Pf\| \le \|f\|$ for $f \in L^1(X, \Sigma, m)$. A substochastic operator is called *stochastic* if $\|Pf\| = \|f\|$ for $f \in L^1(X, \Sigma, m)$ and $f \ge 0$. We denote by $P^*\colon L^\infty(X, \Sigma, m) \to L^\infty(X, \Sigma, m)$ the *adjoint operator* of P, i.e. for every $f \in L^1(X, \Sigma, m)$ and $g \in L^\infty(X, \Sigma, m)$

$$\int_X Pf(x)g(x)\, m(dx) = \int_X f(x)P^*g(x)\, m(dx).$$

Let P be a transition kernel on (X, Σ). Take $f \in L^1(X, \Sigma, m)$ with $f \ge 0$ and $\|f\| = 1$. The measure $\mu \in \mathscr{M}^+(X)$ given by $d\mu = f dm$, i.e.

$$\mu(B) = \int_B f(x)\, m(dx), \quad B \in \Sigma,$$

is absolutely continuous with respect to m and we call f the *density* of μ (with respect to m). If the measure $\nu = P\mu$ given by

$$\nu(B) = P\mu(B) = \int_X P(x, B) f(x)\, m(dx), \quad B \in \Sigma,$$

is absolutely continuous with respect to m then, by the Radon–Nikodym theorem, there is an essentially unique $g \in L^1$, $g \ge 0$, such that

$$\nu(B) = \int_B g(x)\, m(dx), \quad B \in \Sigma.$$

The function g is the Radon–Nikodym derivative $d\nu/dm$ of the measure ν with respect to m. If P is a transition probability then we have $P\mu(X)=1$. Thus f is mapped to g, and we can write

$$Pf = g, \quad \text{if } f = \frac{d\mu}{dm},$$

which is the density of the measure $\nu = P\mu$

$$P\mu(B) = \int_B Pf(x)\, m(dx).$$

This motivates the following definitions.

Denote by D the subset of $L^1 = L^1(X, \Sigma, m)$ which contains all *densities*, i.e.

$$D = \{f \in L^1 \colon\ f \ge 0,\ \|f\| = 1\}.$$

Observe that a linear operator P on $L^1(X, \Sigma, m)$ is a stochastic operator if and only if $P(D) \subseteq D$. The transition kernel P corresponds to a substochastic operator P if

$$\int_X P(x, B) f(x)\, m(dx) = \int_B Pf(x)\, m(dx) \quad \text{for all } B \in \Sigma,\ f \in D,$$

or, equivalently, the adjoint of P is of the form

$$P^* g(x) = \int g(y)\, P(x, dy) \quad \text{for } g \in L^\infty(X, \Sigma, m). \tag{2.3}$$

On the other hand, if a transition kernel P has the following property

$$m(B) = 0 \Longrightarrow P(x, B) = 0 \quad \text{for } m\text{-a.e. } x \text{ and } B \in \Sigma, \tag{2.4}$$

then there exists a substochastic operator P such that (2.3) holds.

Remark 2.1 There exists a stochastic operator which does not have a transition kernel [32]. But if X is a Polish space (a complete separable metric space) and Σ is a σ-algebra of Borel subsets of X, then each substochastic operator on $L^1(X, \Sigma, m)$ has a transition kernel [48]. The same result holds if X is a Borel subset of a Polish space and Σ is the σ-algebra of Borel subsets of X.

2.1.4 Integral Stochastic Operators

Let (X, Σ, m) be a σ-finite measure space. Any measurable function $k \colon X \times X \to [0, \infty)$ such that

$$\int_X k(x, y)\, m(dx) = 1 \quad \text{for all } y \in X$$

defines a transition probability by

$$P(x, B) = \int_B k(y, x)\, m(dy), \quad B \in \Sigma,$$

called an *integral kernel*. If we set

$$Pf(x) = \int_X k(x, y) f(y)\, m(dy), \quad f \in L^1(X, \Sigma, m),$$

then for $f \geq 0$ we have $Pf \geq 0$ and

$$\int_X Pf(x)\, m(dx) = \int_X \int_X k(x, y)\, m(dx) f(y)\, m(dy) = \int_X f(y)\, m(dy),$$

which shows that P is a stochastic operator. If, instead, k is such that

$$\int_X k(x, y)\, m(dx) \leq 1 \quad \text{for all } y \in X,$$

then $P(x, B)$ is a transition kernel. It defines a substochastic operator on $L^1(X, \Sigma, m)$, called an integral operator.

Suppose that the state space is *discrete*, i.e. the set X is finite or countable and Σ is the family of all subsets of X. Since any probability measure on Σ is uniquely defined by its values on the singleton sets $\{y\}$, $y \in X$, any transition probability P satisfies

$$P(x, B) = \sum_{y \in B} P(x, \{y\}), \quad x \in X, \; B \subseteq X.$$

If the state space is discrete and the measure m is the counting measure then (X, Σ, m) is a σ-finite measure space and every kernel is an integral kernel.

We now consider the case of $X = \mathbb{N} = \{0, 1, \ldots\}$, where we use the notation $l^1 = L^1(X, \Sigma, m)$ with m being the counting measure on X. We represent any function f as a sequence $u = (u_i)_{i \in \mathbb{N}}$. We have $(u_i)_{i \in \mathbb{N}} \in l^1$ if and only if $\sum_{i=0}^{\infty} |u_i| < \infty$. If

$$\sum_{i=0}^{\infty} k_{ij} = 1 \quad \text{for all } j,$$

then the operator $P \colon l^1 \to l^1$ defined by

$$(Pu)_i = \sum_{j=0}^{\infty} k_{ij} u_j, \quad i \in \mathbb{N},$$

is a stochastic operator. It can be identified with a matrix $P = [k_{ij}]_{i,j\in\mathbb{N}}$, called a *stochastic matrix*. It has non-negative elements and the elements in each column sum up to one. The transposed matrix

$$P^* = [k_{ji}]_{i,j\in\mathbb{N}}, \quad k_{ji} = P(i, \{j\}),$$

is called a *transition matrix* and it is such that the elements in each row sum up to one. Note that k_{ji} is the probability of going from state i to state j.

2.1.5 Frobenius–Perron Operator

Consider a measurable transformation $S\colon X \to X$, where (X, Σ, m) is a space with a σ-finite measure m. Let μ be a probability measure on (X, Σ) and let us observe the evolution of this measure under the action of the system. For example, if we start with the probability measure concentrated at the point x, i.e. the Dirac measure δ_x, then under the action of the system we obtain the measure $\delta_{S(x)}$. In general, if a measure μ describes the distribution of points in the phase space X, then the measure ν given by the formula $\nu(B) = \mu(S^{-1}(B))$ describes the distribution of points after the action of the transformation S. Assume that μ is absolutely continuous with respect to m with density f. If the measure ν is also absolutely continuous with respect to m, and $g = d\nu/dm$, then we define an operator P_S by $P_S f = g$. This operator corresponds to the transition probability function $P(x, B)$ on (X, Σ) given by

$$P(x, B) = \begin{cases} 1, & \text{if } S(x) \in B, \\ 0, & \text{if } S(x) \notin B. \end{cases} \tag{2.5}$$

The operator P_S is correctly defined if the transition probability function $P(x, B)$ satisfies condition (2.4). Condition (2.4) now takes the form

$$m(B) = 0 \Longrightarrow m(S^{-1}(B)) = 0 \text{ for } B \in \Sigma \tag{2.6}$$

and the transformation S which satisfies (2.6) is called *non-singular*. This operator can be extended to a bounded linear operator $P_S\colon L^1 \to L^1$, and P_S is a stochastic operator. The operator P_S is called the *Frobenius–Perron operator* for the transformation S or the *transfer operator* or the *Ruelle operator*.

We now give a formal definition of the Frobenius–Perron operator. Let (X, Σ, m) be a σ-finite measure space and let S be a measurable non-singular transformation of X. An operator $P_S\colon L^1 \to L^1$ which satisfies the following condition

$$\int_B P_S f(x)\, m(dx) = \int_{S^{-1}(B)} f(x)\, m(dx) \quad \text{for } B \in \Sigma \text{ and } f \in L^1 \tag{2.7}$$

is the Frobenius–Perron operator for the transformation S. The adjoint of the Frobenius–Perron operator $P_S^*\colon L^\infty \to L^\infty$ is given by $P_S^* g(x) = g(S(x))$ and is called the *Koopman operator* or the *composition operator*. In particular, if $S\colon X \to X$ is one to one and non-singular with respect to m, then

$$P_S f(x) = \mathbf{1}_{S(X)}(x) f(S^{-1}(x)) \frac{d(m \circ S^{-1})}{dm}(x) \quad \text{for } m\text{-a.e. } x \in X,$$

where $d(m \circ S^{-1})/dm$ is the Radon–Nikodym derivative of the measure $m \circ S^{-1}$ with respect to m.

We next show how to find the Frobenius–Perron operator for piecewise smooth transformations of subsets of $\mathbb{R}^d$. Let X be a subset of $\mathbb{R}^d$ with non-empty interior and with the boundary of zero Lebesgue measure. Let $S\colon X \to X$ be a measurable transformation. We assume that there exist pairwise disjoint open subsets $U_1,\ldots,U_n$ of X having the following properties:

(a) the sets $X_0 = X \setminus \bigcup_{i=1}^n U_i$ and $S(X_0)$ have zero Lebesgue measure,

(b) maps $S_i = S\big|_{U_i}$ are diffeomorphisms from U_i onto $S(U_i)$, i.e. S_i are C^1 and invertible transformations with $\det S_i'(x) \neq 0$ at each point $x \in U_i$.

Then transformations $\psi_i = S_i^{-1}$ are also diffeomorphisms from $S(U_i)$ onto U_i and the Frobenius–Perron operator P_S exists and is given by the formula

$$P_S f(x) = \sum_{i \in I_x} f(\psi_i(x)) |\det \psi_i'(x)|, \tag{2.8}$$

where $I_x = \{i\colon \psi_i(x) \in U_i\}$. Indeed,

$$\begin{aligned}
\int_{S^{-1}(B)} f(x)\,dx &= \sum_{i=1}^n \int_{S^{-1}(B)\cap U_i} f(x)\,dx = \sum_{i=1}^n \int_{\psi_i(B)} f(x)\,dx \\
&= \sum_{i=1}^n \int_{B\cap S(U_i)} f(\psi_i(x)) |\det \psi_i'(x)|\,dx \\
&= \int_B \sum_{i\in I_x} f(\psi_i(x)) |\det \psi_i'(x)|\,dx = \int_B Pf(x)\,dx.
\end{aligned}$$

2.1.6 Iterated Function Systems

We now consider a finite set of different non-singular transformations $S_1, \ldots, S_k$ of a σ-finite measure space (X, Σ, m). Let $p_1(x), \ldots, p_k(x)$ be non-negative measurable functions defined on X such that $p_1(x) + \cdots + p_k(x) = 1$ for all $x \in X$. Take a point x. We choose a transformation S_j with probability $p_j(x)$ and the position of x after the action of the system is given by $S_j(x)$. Thus we consider the following transition probability

$$P(x, B) = \sum_{j=1}^{k} p_j(x)\delta_{S_j(x)}(B)$$

for $x \in X$ and measurable sets B. Hence, for any measure μ, we have

$$P\mu(B) = \sum_{j=1}^{k} \int_X p_j(x)\delta_{S_j(x)}(B)\mu(dx) = \sum_{j=1}^{k} \int_{S_j^{-1}(B)} p_j(x)\mu(dx).$$

If μ is absolutely continuous and $f = d\mu/dm$ then

$$P\mu(B) = \sum_{j=1}^{k} \int_{S_j^{-1}(B)} p_j(x) f(x)\, m(dx) = \sum_{j=1}^{k} \int_B P_{S_j}(p_j f)(x)\, m(dx),$$

where $P_{S_1}, \ldots, P_{S_k}$ are the corresponding Frobenius–Perron operators. Consequently, the stochastic operator on $L^1 = L^1(X, \Sigma, m)$ corresponding to P is of the form

$$Pf = \sum_{j=1}^{k} P_{S_j}(p_j f), \quad f \in L^1.$$

We can extend the iterated function system in the following way. Consider a family of measurable transformations $S_y : X \to X$, $y \in Y$, where Y is a metric space which carries a Borel measure ν, and a family of measurable functions $p_y : X \to [0, \infty)$, $y \in Y$, satisfying

$$\int_Y p_y(x)\nu(dy) = 1, \quad x \in X,$$

so that the transition probability P is of the form

$$P(x, B) = \int_Y \mathbf{1}_B(S_y(x))p_y(x)\nu(dy), \quad x \in X. \tag{2.9}$$

If each S_y is a non-singular transformation of the space (X, Σ, m) then the stochastic operator P corresponding to the transition probability as in (2.9) is of the form

$$Pf = \int_Y P_{S_y}(p_y f)\nu(dy), \quad f \in L^1,$$

where P_{S_y} is the Frobenius–Perron operator for S_y, $y \in Y$.

As a final example, we look at the transition probability describing the switching mechanism in Sects. 1.8 and 1.9. Let I be at most a countable set. We define the transformation $S_j : X \times I \to X \times I$, $j \in I$, by

$$S_j(x, i) = (x, j), \quad x \in X,\ i, j \in I.$$

Each transformation defined on $\mathbb{X} = X \times I$ is non-singular with respect to the product measure m of a measure on X and the counting measure on I. We assume that $q_{ji}(x),\ j \neq i$, are non-negative measurable functions satisfying $\sum_{j \neq i} q_{ji}(x) < \infty$ for all $i \in I$, $x \in X$. Then we define the jump rate function by

$$q_i(x) = \sum_{j \neq i} q_{ji}(x)$$

and the probabilities $p_j,\ j \in I$, by $p_i(x, i) = 0$ and

$$p_j(x, i) = \begin{cases} 1, & q_i(x) = 0,\ j \neq i, \\ \frac{q_{ji}(x)}{q_i(x)}, & q_i(x) > 0,\ j \neq i. \end{cases}$$

The stochastic operator P on $L^1 = L^1(\mathbb{X}, \Sigma, m)$ corresponding to the transition probability

$$P((x, i), \{(x, j)\}) = p_j(x, i)$$

is given by

$$Pf(x, i) = \sum_{j \neq i} p_i(x, j) f(x, j).$$

In particular, if $I = \{0, 1\}$ and $q_i(x) > 0$ for $x \in X,\ i = 0, 1$, then

$$p_j(x, i) = \begin{cases} 0, & j = i, \\ 1, & j = 1 - i, \end{cases}$$

and $Pf(x, i) = P_{S_{1-i}}(x, i) = f(x, 1 - i)$ for $(x, i) \in \mathbb{X} = X \times \{0, 1\}$.

2.2 Discrete-Time Markov Processes

2.2.1 Markov Processes and Transition Probabilities

A family $\xi_n,\ n \in \mathbb{N}$, of X-valued random variables defined on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ is called a *discrete-time stochastic process* with *state space* (X, Σ). It is said to have the (weak) *Markov property* if for each time $n \geq 0$

$$\mathbb{P}(\xi_{n+1} \in B | \xi_0, \ldots, \xi_n) = \mathbb{P}(\xi_{n+1} \in B | \xi_n), \quad B \in \Sigma. \tag{2.10}$$

The sequence $\mathscr{F}_n^\xi = \sigma(\xi_0, \dots, \xi_n)$, $n \ge 0$, is called the *natural filtration* or the *history* of the process (ξ_n). We have

$$\sigma(\xi_0, \dots, \xi_n) = \{\{\xi_0 \in B_0, \dots, \xi_n \in B_n\} \colon B_0, \dots, B_n \in \Sigma\}.$$

The monotone class theorem and properties of the conditional expectation imply that if (ξ_n) has the weak Markov property then for each $k \ge 0$ and all $g_0, \dots, g_k \in B(X)$ the following holds

$$\mathbb{E}(g_0(\xi_n) \dots g_k(\xi_{n+k})|\xi_0, \dots, \xi_n) = \mathbb{E}(g_0(\xi_n) \dots g_k(\xi_{n+k})|\xi_n).$$

We can extend it further as follows. A sequence (ξ_n) has the weak Markov property if and only if the future $\sigma(\xi_m : m \ge n)$ and the past $\sigma(\xi_m : m \le n)$ are conditionally independent given the present $\sigma(\xi_n)$ (see Lemma A.4), i.e. for each n and all $A \in \sigma(\xi_m : m \ge n)$ and $F \in \sigma(\xi_m : m \le n)$ we have

$$\mathbb{P}(A \cap F|\xi_n) = \mathbb{P}(A|\xi_n)\mathbb{P}(F|\xi_n).$$

A discrete-time stochastic process $\xi = (\xi_n)_{n\ge 0}$ is called a *(homogeneous) Markov process* with transition probability P and initial distribution μ on (X, Σ) if there is a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ such that

$$\mu(B) = \mathbb{P}(\xi_0 \in B) \quad \text{and} \quad \mathbb{P}(\xi_{n+1} \in B|\xi_0, \dots, \xi_n) = P(\xi_n, B) \tag{2.11}$$

for all $B \in \Sigma$ and $n \ge 0$. Taking the conditional expectation of (2.11) with respect to $\sigma(\xi_n)$ we obtain

$$\mathbb{P}(\xi_{n+1} \in B|\xi_n) = P(\xi_n, B)$$

which implies that $\xi = (\xi_n)_{n\ge 0}$ has the weak Markov property. Moreover, we have

$$\mathbb{P}(\xi_{n+k} \in B|\xi_0, \dots, \xi_n) = \mathbb{P}(\xi_{n+k} \in B|\xi_n) = P^k(\xi_n, B), \quad k \ge 1,$$

where P^k, called the *kth step transition probability*, is defined inductively by

$$P^0(x, B) = \delta_x(B), \quad P^{k+1}(x, B) = \int_X P(y, B)\, P^k(x, dy).$$

The kernels P^k, $k \ge 1$, satisfy the *Chapman–Kolmogorov equation*

$$P^{k+n}(x, B) = \int_X P^k(y, B)\, P^n(x, dy). \tag{2.12}$$

We can regard the stochastic process ξ_n, $n \ge 0$, as an $X^{\mathbb{N}}$-valued random variable, where $X^{\mathbb{N}}$ is the product space

$$X^{\mathbb{N}} = \{x = (x_0, x_1, \ldots) \colon x_n \in X,\ n \geq 0\}$$

with product σ-algebra $\Sigma^{\mathbb{N}} = \sigma(\mathscr{C})$ which is the smallest σ-algebra of subsets of $X^{\mathbb{N}}$ containing the family $\mathscr{C}$ of all finite-dimensional rectangles

$$\mathscr{C} = \{\{x \in X^{\mathbb{N}} \colon x_0 \in B_0, \ldots, x_n \in B_n\} \colon B_0, \ldots, B_n \in \Sigma,\ n \geq 0\}. \tag{2.13}$$

To see this let for each ω the mapping $n \mapsto \xi_n(\omega)$ denotes the sequence $\xi(\omega) = (\xi_n(\omega))_{n\geq 0} \in X^{\mathbb{N}}$. Thus we obtain the mapping $\xi \colon \Omega \to X^{\mathbb{N}}$ and ξ is measurable if and only if $\xi_n \colon \Omega \to X$ is measurable for each $n \in \mathbb{N}$. The distribution of the process ξ is a probability measure μ_ξ on $(X^{\mathbb{N}}, \Sigma^{\mathbb{N}})$ and it is uniquely defined through its values on finite-dimensional rectangles. In other words, the finite-dimensional distributions $\mu_{0,1,\ldots,n} = \mu_{(\xi_0,\ldots,\xi_n)}$

$$\mu_{0,1,\ldots,n}(B_0 \times \cdots \times B_n) = \mathbb{P}(\xi_0 \in B_0, \ldots, \xi_n \in B_n), \quad B_0, \ldots, B_n \in \Sigma,\ n \geq 0,$$

uniquely determine the law of the process, since

$$\begin{aligned}\mu_\xi(B_0 \times \cdots \times B_n \times X^{\mathbb{N}}) &= \mathbb{P}(\xi \in B_0 \times \cdots \times B_n \times X^{\mathbb{N}}) \\ &= \mu_{(\xi_0,\ldots,\xi_n)}(B_0 \times \cdots \times B_n)\end{aligned}$$

and all finite-dimensional rectangles $B_0 \times \cdots \times B_n \times X^{\mathbb{N}}$ generate $\Sigma^{\mathbb{N}}$. A transition probability and an initial distribution determine the finite-dimensional distributions of a discrete-time Markov process, as stated in the next theorem. Its proof is based on the monotone class theorem and properties of conditional expectations.

Theorem 2.1 *Let μ be a probability measure, P be a transition probability, and $(\xi_n)_{n\geq 0}$ be defined on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Then $(\xi_n)_{n\geq 0}$ is a Markov process with transition probability P and initial distribution μ if and only if*

$$\begin{aligned}&\mathbb{P}(\xi_0 \in B_0, \xi_1 \in B_1, \ldots, \xi_n \in B_n) \\ &\quad = \int_{B_0}\int_{B_1} \cdots \int_{B_{n-1}} P(x_{n-1}, B_n) P(x_{n-2}, dx_{n-1}) \ldots P(x_0, dx_1)\, \mu(dx_0)\end{aligned} \tag{2.14}$$

for all sets $B_0, B_1, \ldots, B_n \in \Sigma$, $n \geq 0$.

2.2.2 Random Mapping Representations

Suppose that $(\vartheta_n)_{n\geq 1}$ is a sequence of i.i.d. random variables with values in a metric space Y. Let ξ_0 be an X-valued random variable independent of $(\vartheta_n)_{n\geq 1}$ and with distribution μ. Consider a measurable function $\kappa \colon X \times Y \to X$ and define a sequence

ξ_n of X-valued random variables by

$$\xi_n = \kappa(\xi_{n-1}, \vartheta_n), \quad n \geq 1. \tag{2.15}$$

For $\mathscr{F}_n = \sigma(\xi_0, \ldots, \xi_n)$ we have $\mathscr{F}_n \subseteq \sigma(\xi_0, \vartheta_1, \ldots, \vartheta_n)$ and ϑ_{n+1} is independent of $\mathscr{F}_n$ and of ϑ_n. Thus, we see that (ξ_n) has the weak Markov property, by (2.1). Hence, ξ is Markov with transition probability of the form

$$P(x, B) = \mathbb{P}(\kappa(x, \vartheta_1) \in B) = \int_Y \mathbf{1}_B(\kappa(x, y))\nu(dy)$$

and initial distribution μ. A particular example is the *random walk* on $\mathbb{R}^d$ defined by

$$\xi_n = \xi_{n-1} + \vartheta_n, \quad n \geq 1.$$

From Proposition 2.2 it follows that a typical discrete-time Markov process can be defined by a recursive formula as in (2.15).

Theorem 2.2 *Let P be a transition probability on a Borel space (X, Σ). Then there exists a measurable function $\kappa : X \times [0, 1] \to X$ such that for any sequence (ϑ_n) of independent random variables with uniform distribution on $[0, 1]$ and for any $x \in X$ the sequence defined by (2.15) is a discrete-time Markov process with transition probability P and initial distribution δ_x.*

To define a discrete-time Markov process with transition probability P on a Borel space (X, Σ), we can take Ω, by Theorem 2.2, to be the countable product $[0, 1]^{\mathbb{N}}$ of the unit interval $[0, 1]$ with the product σ-algebra $\mathscr{F} = \mathscr{B}([0, 1])^{\mathbb{N}}$ and the measure $\mathbb{P}$ as a countable product of the Lebesgue measure on $[0, 1]$. Then we define a sequence of i.i.d. random variables $\vartheta_n : \Omega \to [0, 1], n \geq 0$, by $\vartheta_n(\omega) = \omega_n$ for $\omega = (\omega_n)_{n\geq 0} \in \Omega$, and ξ_n by (2.15).

2.2.3 Canonical Processes

We next construct a canonical discrete-time Markov process with a given transition probability and an initial distribution on an arbitrary measurable space (X, Σ). The existence of the process with given finite-dimensional distributions amounts to showing that there is a probability measure on the product space $(X^{\mathbb{N}}, \Sigma^{\mathbb{N}})$ satisfying (2.14) with ξ being the identity mapping. This can be realized through the discrete-time version of the Kolmogorov extension theorem. For its proof we refer the reader to [55]. Note that no regularity condition is needed on the state space.

Theorem 2.3 (Ionescu–Tulcea) *Let P be a transition probability on a measurable space (X, Σ). For every $x \in X$ there exists a unique probability measure $\mathbb{P}_x$ on the product space $(X^{\mathbb{N}}, \Sigma^{\mathbb{N}})$ such that*

$$\mathbb{P}_x(B_0 \times \cdots \times B_n \times X^{\mathbb{N}}) = \mathbf{1}_{B_0}(x) \int_{B_1} \cdots \int_{B_{n-1}} P(x_{n-1}, B_n) P(x_{n-2}, dx_{n-1}) \dots P(x, dx_1)$$

for all $B_0, \dots, B_n \in \Sigma$, $n \geq 0$. *Moreover, for every set* $A \in \Sigma^{\mathbb{N}}$, *the map* $x \mapsto \mathbb{P}_x(A)$ *is* Σ*-measurable.*

The function $\mathbb{P}_x(A)$ is a transition probability from (X, Σ) to $(X^{\mathbb{N}}, \Sigma^{\mathbb{N}})$. Thus, by Proposition 2.1, for each probability measure μ on (X, Σ) there exists a unique probability measure $\mathbb{P}_\mu$ on $(X \times X^{\mathbb{N}}, \Sigma \otimes \Sigma^{\mathbb{N}})$ such that

$$\mathbb{P}_\mu(B \times A) = \int_B \mathbb{P}_x(A)\, \mu(dx), \quad B \in \Sigma,\ A \in \Sigma^{\mathbb{N}}.$$

The sequence space

$$\Omega = X^{\mathbb{N}} = \{\omega = (x_0, x_1, \dots) \colon x_n \in X,\ n \geq 0\}$$

is called a *canonical space* and the coordinate mappings $\xi_n(\omega) = x_n$, $n \geq 0$, define a process $\xi = (\xi_n)_{n\geq 0}$ on $(\Omega, \mathscr{F}) = (X^{\mathbb{N}}, \Sigma^{\mathbb{N}})$ with distribution $\mathbb{P}_\mu$, called a *Markov canonical process*. Note that if we let the process start at $x \in X$ so that $\mu = \delta_x$ then we have $\mathbb{P}_{\delta_x} = \mathbb{P}_x$ and if $A \in \Sigma^{\mathbb{N}}$ is the rectangle

$$A = \{\omega \colon \xi_0(\omega) \in B_0, \dots, \xi_n(\omega) \in B_n\}$$

then we have

$$\begin{aligned}\mathbb{P}_\mu(A) &= \mathbb{P}_\mu(\xi_0 \in B_0, \xi_1 \in B_1, \dots, \xi_n \in B_n) \\ &= \int_{B_0} \int_{B_1} \cdots \int_{B_{n-1}} P(x_{n-1}, B_n) P(x_{n-2}, dx_{n-1}) \dots P(x_0, dx_1) \mu(dx_0).\end{aligned}$$

Consequently, for each probability measure $\mathbb{P}_\mu$ on $(\Omega, \mathscr{F})$ the process $\xi = (\xi_n)_{n\geq 0}$ is homogeneous Markov with transition probability P and initial distribution μ. We write $\mathbb{E}_x$ for the integration with respect to $\mathbb{P}_x$, $x \in X$, and we have

$$\mathbb{E}_\mu(\eta) = \int_X \mathbb{E}_x(\eta)\, \mu(dx)$$

for any bounded or non-negative random variable $\eta \colon \Omega \to \mathbb{R}$. In particular, we can rewrite the Markov property as: for each n and any bounded measurable g defined on the sequence space we have

$$\mathbb{E}_\mu(g(\xi_n, \xi_{n+1}, \dots) | \mathscr{F}_n) = h(\xi_n), \quad \text{where } h(x) = \mathbb{E}_x(g(\xi_0, \xi_1, \dots)).$$

2.3 Continuous-Time Markov Processes

2.3.1 Basic Definitions

A family $\xi(t)$, $t \in [0, \infty)$, of X-valued random variables is called a *continuous-time stochastic process* with state space X, where (X, Σ) is a measurable space. For each $t \geq 0$, let $\mathscr{F}_t = \sigma(\xi(r) \colon r \leq t)$ be the σ-algebra generated by all random variables $\xi(r), 0 \leq r \leq t$. Since $\mathscr{F}_s \subseteq \mathscr{F}_t$ for all $0 \leq s < t$, the collection $\mathscr{F}_t$, $t \geq 0$, is called a *history* of the process or a *filtration*. If X is a Borel space then the process $\xi = \{\xi(t) : t \geq 0\}$ is said to be *right-continuous* (*càdlàg*) if its sample paths $t \mapsto \xi(t)(\omega)$, $\omega \in \Omega$, are right-continuous (càdlàg).

Let $\xi = \{\xi(t) \colon t \geq 0\}$ be an X-valued process defined on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$. The process ξ is said to be a *Markov process* if for any times s, t we have

$$\mathbb{P}(\xi(s+t) \in B|\mathscr{F}_s) = \mathbb{P}(\xi(s+t) \in B|\xi(s)), \quad B \in \Sigma. \tag{2.16}$$

By the monotone class theorem the Markov property (2.16) holds if and only if

$$\mathbb{E}(g(\xi(s+t))|\mathscr{F}_s) = \mathbb{E}(g(\xi(s+t))|\xi(s))$$

for all $g \in B(X)$, where $B(X)$ is the space of bounded and measurable functions $g \colon X \to \mathbb{R}$. Note that a process is Markov if and only if for any $0 \leq s_1 < \cdots < s_n$ and $g \in B(X)$

$$\mathbb{E}(g(\xi(s_n + t))|\xi(s_1), \dots, \xi(s_n)) = \mathbb{E}(g(\xi(s_n + t))|\xi(s_n)).$$

A family $P = \{P(t, \cdot) \colon t \geq 0\}$ of transition probabilities on (X, Σ) is said to be a *transition (probability) function* if it satisfies the *Chapman–Kolmogorov equation*

$$P(s+t, x, B) = \int_X P(s, y, B) P(t, x, dy), \quad s, t \geq 0, \tag{2.17}$$

and $P(0, x, B) = \delta_x(B)$ for every $x \in X$, $B \in \Sigma$. We say that ξ is a *homogeneous Markov process* with transition function P and starting at x if

$$\mathbb{P}(\xi(0) = x) = 1 \quad \text{and} \quad \mathbb{P}(\xi(s+t) \in B|\mathscr{F}_s) = P(t, \xi(s), B) \tag{2.18}$$

for all $B \in \Sigma$ and for all times s and t. Equivalently,

$$\mathbb{E}(g(\xi(s+t))|\mathscr{F}_s) = \int_X g(y) P(t, \xi(s), dy)$$

for all $g \in B(X)$, $s, t \geq 0$. We interpret $P(t, x, B)$ as the probability that the stochastic process ξ moves from state x at time 0 to a state in B at time t. The transition function P defines a family of bounded linear operators

$$T(t)g(x) = \int_X g(y)P(t, x, dy)$$

on the Banach space $B(X)$ with supremum norm $\|\cdot\|_u$. It follows from the Chapman–Kolmogorov equation (2.17) that $T(t)$, $t \geq 0$, is a *semigroup*, i.e.

$$T(s+t)g = T(t)(T(s)g), \quad g \in B(X),\ s, t \geq 0.$$

The family $\{T(t)\}_{t \geq 0}$ is called the *transition semigroup associated to* ξ.

A random variable $\tau\colon \Omega \to [0, \infty]$ is called a *stopping time* for the filtration $(\mathscr{F}_t)$ if it satisfies $\{\tau \leq t\} \in \mathscr{F}_t$ for all $t \geq 0$. The σ-algebra $\mathscr{F}_\tau$ giving the history known up to time τ is defined as

$$\mathscr{F}_\tau = \{A \in \mathscr{F} : A \cap \{\tau \leq t\} \in \mathscr{F}_t \text{ for all } t \geq 0\}.$$

If τ_1 and τ_2 are two stopping times and $\tau_1 \leq \tau_2$ then $\mathscr{F}_{\tau_1} \subseteq \mathscr{F}_{\tau_2}$. A Markov process ξ is said to be *strong Markov* if for any stopping time τ, the *strong Markov property* at τ holds

$$\mathbb{P}(\xi(\tau + t) \in B | \mathscr{F}_\tau) = \mathbb{P}(\xi(\tau + t) \in B | \xi(\tau)) \quad \text{a.s. on } \tau < \infty \tag{2.19}$$

for all $B \in \Sigma$ and all times t. If X is a Borel space and the process has right-continuous sample paths then $\xi(\tau)$ is $\mathscr{F}_\tau$-measurable on $\{\tau < \infty\}$ for any stopping time τ for the filtration $(\mathscr{F}_t)$.

As in the discrete-time case we can consider a canonical space $\Omega = X^{\mathbb{R}_+}$, which now is the space of all functions $\omega\colon \mathbb{R}_+ \to X$, with product σ-algebra $\Sigma^{\mathbb{R}_+} = \sigma(\mathscr{C})$ which is the smallest σ-algebra of subsets of $X^{\mathbb{R}_+}$ containing the family $\mathscr{C}$ of all cylinder sets:

$$\{\omega \in X^{\mathbb{R}_+} : \omega(t_0) \in B_0, \dots, \omega(t_n) \in B_n\}, \quad B_0, \dots, B_n \in \Sigma, \quad 0 \leq t_0 < \cdots < t_n, \quad n \geq 0.$$

Let $\mathscr{F}_t = \sigma(\xi(r) : r \leq t), t \geq 0$, where the canonical process $\xi(t), t \geq 0$, is defined as the identity map on $(X^{\mathbb{R}_+}, \Sigma^{\mathbb{R}_+})$. Suppose that P is a transition probability function on a Borel space (X, Σ). It follows from a continuous-time version of the Kolmogorov extension theorem [55] that for every $x \in X$ there exists a probability measure $\mathbb{P}_x$ on $(\Omega, \mathscr{F})$ such that $\xi(t), t \geq 0$, is a homogeneous Markov process with transition function P and starting at x. Moreover, for every set $A \in \Sigma^{\mathbb{N}}$, the map $x \mapsto \mathbb{P}_x(A)$ is Σ-measurable. We have

$$\mathbb{P}_x(\xi(t) \in B) = P(t, x, B), \quad t \geq 0, \ B \in \Sigma,$$

and

$$\mathbb{P}_x(\xi(t_1) \in B_1, \dots, \xi(t_n) \in B_n)$$
$$= \int_{B_1} \dots \int_{B_{n-1}} P(t_n - t_{n-1}, x_{n-1}, B_n) P(t_{n-1} - t_{n-2}, x_{n-2}, dx_{n-1}) \dots P(t_1, x, dx_1)$$

for all $0 < t_1 < \dots < t_n$, $B_1, \dots, B_n \in \Sigma$, $n \in \mathbb{N}$.

Given a transition function on a Borel space, for each $x \in X$, there exists a probability space $(\Omega, \mathscr{F}, \mathbb{P}_x)$ and a homogeneous Markov process $\{\xi(t) : t \geq 0\}$ defined on $(\Omega, \mathscr{F}, \mathbb{P}_x)$ with transition function P and starting at x. We will also denote the process $\xi(t)$ started at x by $\xi_x(t)$. Then the family of processes $\xi = \{\xi_x(t) : t \geq 0,\ x \in X\}$ is called a *Markov family*. We will simply say that $\xi(t)$ is a Markov process with state space X defined on $(\Omega, \mathscr{F}, \mathbb{P}_x)$. However, in general, a transition function is unknown in advance and we need to construct the processes directly from other processes.

2.3.2 Processes with Stationary and Independent Increments

Our first example is a continuous time extension of random walks. An $\mathbb{R}^d$-valued process $\xi(t)$, $t \geq 0$, is said to have *independent increments*, if $\xi(s+t) - \xi(s)$ is independent of $\mathscr{F}_s = \sigma(\xi(r) : r \leq s)$ for all $t, s \geq 0$. Given such a process, we have by (2.1)

$$\mathbb{P}(\xi(s+t) \in B | \mathscr{F}_s) = \mathbb{P}(\xi(s+t) - \xi(s) + \xi(s) \in B | \xi(s)) = \mathbb{P}(\xi(s+t) \in B | \xi(s))$$

for $B \in \mathscr{B}(\mathbb{R}^d)$, since $\xi(s+t) = \xi(s+t) - \xi(s) + \xi(s)$ and $\xi(s+t) - \xi(s)$ is independent of $\mathscr{F}_s$, which shows that ξ has the Markov property.

The process ξ has *stationary increments*, if the distribution of $\xi(s+t) - \xi(s)$ is the same as the distribution of $\xi(t) - \xi(0)$ for all $s, t \geq 0$. In that case

$$\mathbb{E}\mathbf{1}_B(\xi(s+t) - \xi(s) + x) = \mathbb{E}\mathbf{1}_B(\xi(t) - \xi(0) + x)$$

for all x and s, t. In particular, we have $\xi(s+t) - \xi(0) = \xi(s+t) - \xi(t) + \xi(t) - \xi(0)$ and the random variables $\xi(s+t) - \xi(t)$ and $\xi(t) - \xi(0)$ are independent. Hence, if μ_t is the distribution of $\xi(t) - \xi(0)$, then μ_{s+t} is the convolution of μ_s and μ_t, i.e.

$$\mu_{s+t}(B) = (\mu_s * \mu_t)(B) = \int \mu_s(B - x)\mu_t(dx),$$

and

$$P(t, x, B) = \int_{\mathbb{R}^d} \mathbf{1}_B(x + y)\mu_t(dy).$$

A process is called a *Lévy process* if it has stationary independent increments, it starts at zero, i.e. $\xi(0) = 0$ a.s., and it is *continuous in probability*, i.e. for every

$\varepsilon > 0$ we have

$$\lim_{t\to 0} \mathbb{P}(|\xi(t)| > \varepsilon) = 0.$$

2.3.3 Markov Jump-Type Processes

In this section, we provide a simple construction of pure jump-type processes with bounded jump rate function. They were introduced in Sect. 1.3 and are particular examples of PDMPs defined in Sect. 1.19. Here we show that they are Markov processes. Suppose that (X, Σ) is a measurable space, P is a transition probability on X and that φ is a bounded non-negative measurable function. Set $\lambda = \sup\{\varphi(x) \colon x \in X\}$ and define the transition probability $\bar{P}$ by

$$\bar{P}(x, B) = \lambda^{-1}(\varphi(x)P(x, B) + (\lambda - \varphi(x))\delta_x(B)), \quad x \in X, B \in \Sigma. \tag{2.20}$$

Let $(\xi_n)_{n\ge 0}$ be a discrete-time Markov process with transition probability $\bar{P}$ and $(\sigma_n)_{n\ge 1}$ be a sequence of independent random variables, exponentially distributed with mean λ^{-1}, and independent of the sequence $(\xi_n)_{n\ge 0}$. We set $\tau_0 = 0$ and we define

$$\xi(t) = \xi_{n-1} \quad \text{if } \tau_{n-1} \le t < \tau_n, \quad \text{where } \tau_n = \sum_{k=1}^{n} \sigma_k, \quad n \ge 1.$$

It follows from the strong law of large numbers that $\tau_n \to \infty$, as $n \to \infty$, a.s. The sample paths of the process ξ are constant between consecutive τ_n and the random variables σ_n are called holding times.

Let $N(t)$ be the number of jump times τ_n in the time interval $(0, t]$, i.e. we have

$$N(t) = \max\{n \ge 0 \colon \tau_n \le t\} = \sum_{n=0}^{\infty} \mathbf{1}_{(0,t]}(\tau_n).$$

Then $N(t) = 0$ if $t < \tau_1$, and

$$N(t) = n \iff \tau_n \le t < \tau_{n+1}.$$

We show that $N(t)$ is Poisson distributed with parameter λt, i.e.

$$\mathbb{P}(N(t) = n) = e^{-\lambda t}\frac{(\lambda t)^n}{n!}, \quad n \ge 0.$$

We have $\mathbb{P}(N(t) = 0) = \mathbb{P}(\tau_1 > t) = e^{-\lambda t}$ and for $n \ge 1$

$$\mathbb{P}(N(t) = n) = \mathbb{P}(\tau_n \le t < \tau_{n+1}) = \mathbb{P}(\tau_{n+1} > t) - \mathbb{P}(\tau_n > t),$$

since $\{\tau_n > t\} \subseteq \{\tau_{n+1} > t\}$. The random variable τ_n, being the sum of n independent exponentially distributed random variables with parameter λ, has a gamma distribution with density

$$f_{\tau_n}(x) = \frac{\lambda^n x^{n-1}}{(n-1)!} e^{-\lambda x} \quad \text{for } x \geq 0.$$

Hence,

$$\mathbb{P}(\tau_n > t) = \int_t^\infty \frac{\lambda^n x^{n-1}}{(n-1)!} e^{-\lambda x} dx = -e^{-\lambda t} \frac{(\lambda t)^n}{n!} + \mathbb{P}(\tau_{n+1} > t).$$

We can write $\xi(t) = \xi_{N(t)}$ for $t \geq 0$. In particular, if we take $X = \mathbb{N}$ and the trivial Markov chain $\xi_n = n$ for all n, then we have $\xi(t) = N(t)$ for all $t \geq 0$ and $N(t)$, $t \geq 0$, is a Poisson process with intensity $\lambda > 0$. If $(\xi_n)_{n\geq 0}$ is a random walk then $\xi(t)$ is the compound Poisson process as in (1.3). The Poisson process N has stationary independent increments, equivalently, for any n, t, s and $A \in \sigma(N(r) \colon r \leq s)$

$$\mathbb{P}(\{N(t+s) - N(s) = n\} \cap A) = \mathbb{P}(N(t) = n)\mathbb{P}(A).$$

Consider again the general process $\xi(t) = \xi_{N(t)}$ for $t \geq 0$. We denote by $\xi_x(t)$ the particular process $\xi(t)$ starting at $\xi(0) = \xi_0 = x$. We can easily calculate the distribution of $\xi_x(t)$. By independence of $N(t)$ and (ξ_n), we obtain

$$\mathbb{P}(\xi_x(t) \in B) = \mathbb{P}(\xi(t) \in B|\xi_0 = x) = \sum_{n=0}^\infty \mathbb{P}(N(t) = n)\mathbb{P}(\xi_n \in B|\xi_0 = x).$$

Since the random variable $N(t)$ is Poisson distributed with parameter λt, we have

$$P(t, x, B) = \sum_{n=0}^\infty e^{-\lambda t} \frac{(\lambda t)^n}{n!} \bar{P}^n(x, B), \tag{2.21}$$

where $\bar{P}^n$ is the nth step transition probability.

We now check that condition (2.18) holds with $\mathscr{F}_s = \sigma(\xi(r) \colon r \leq s)$. We may write

$$\mathbb{E}(\mathbf{1}_B(\xi(t+s))|\mathscr{F}_s) = \mathbb{E}(\mathbf{1}_B(\xi_{N(t+s)})|\mathscr{F}_s) = \mathbb{E}(\mathbf{1}_B(\xi_{N(t+s)-N(s)+N(s)})|\mathscr{F}_s),$$

which gives

$$\mathbb{E}(\mathbf{1}_B(\xi(t+s))|\mathscr{F}_s) = \sum_{n=0}^\infty \mathbb{E}(\mathbf{1}_{\{N(t+s)-N(s)=n\}} \mathbf{1}_B(\xi_{n+N(s)})|\mathscr{F}_s).$$

Since $N(t+s)-N(s)$ is independent of $\sigma(N(r)\colon r\le s)$ and $\sigma(\xi_n\colon n\ge 0)$, it is also independent of $\mathscr{F}_s$. It is Poisson distributed with parameter λt. Thus

$$\begin{aligned}\mathbb{E}(\mathbf{1}_B(\xi(t+s))|\mathscr{F}_s) &= \sum_{n=0}^{\infty}\mathbb{P}(N(t+s)-N(s)=n)\mathbb{E}(\mathbf{1}_B(\xi_{n+N(s)})|\mathscr{F}_s)\\ &= \sum_{n=0}^{\infty}e^{-\lambda t}\frac{(\lambda t)^n}{n!}\mathbb{E}(\mathbf{1}_B(\xi_{n+N(s)})|\mathscr{F}_s).\end{aligned}$$

The family

$$\{A_1\cap A_2\cap\{N(s)=m\}\colon\ A_1\in\sigma(N(r)\le s),\quad A_2\in\sigma(\xi_k\colon k\le m),\quad m\ge 0\}$$

generates the σ-algebra $\mathscr{F}_s$. Thus, it is enough to show that

$$\int_{A_1\cap A_2\cap\{N(s)=m\}}\mathbf{1}_B(\xi_{n+N(s)})\,d\mathbb{P}=\int_{A_1\cap A_2\cap\{N(s)=m\}}\bar{P}^n(\xi(s),B)\,d\mathbb{P}.$$

To this end, observe that $\xi(s)=\xi_m$ on $\{N(s)=m\}$ and $A_1\cap\{N(s)=m\}$ is independent A_2 and ξ_m. This together with the Markov property for $(\xi_n)_{n\ge 0}$ leads to

$$\begin{aligned}\int_{A_1\cap A_2\cap\{N(s)=m\}}\bar{P}^n(\xi(s),B)\,d\mathbb{P} &= \mathbb{P}(A_1\cap\{N(s)=m\})\int_{A_2}\bar{P}^n(\xi_m,B)\,d\mathbb{P}\\ &= \mathbb{P}(A_1\cap\{N(s)=m\})\int_{A_2}\mathbf{1}_B(\xi_{n+m})\,d\mathbb{P}.\end{aligned}$$

Now, making use of the independence of $A_1\cap\{N(s)=m\}$ and $A_2\cap\{\xi_{n+m}\in B\}$, completes the proof of the Markov property.

2.3.4 Generators and Martingales

Assume that $\xi(t)$ is a Markov process with state space X defined on $(\Omega,\mathscr{F},\mathbb{P}_x)$. Let

$$T(t)g(x)=\mathbb{E}_x g(\xi(t))=\int_X g(y)P(t,x,dy),\quad g\in B(X),\ x\in X,\ t\ge 0,$$

where $\mathbb{E}_x$ is the expectation with respect to $\mathbb{P}_x$. Consider the class $\mathscr{D}(L)$ of all bounded and measurable functions $g\colon X\to\mathbb{R}$ such that the limit

$$\lim_{t\downarrow 0}\frac{\mathbb{E}_x(g(\xi(t)))-g(x)}{t}$$

exists uniformly for all $x \in X$. We denote the limit by $Lg(x)$ and we call L the *infinitesimal generator* of the Markov process ξ. In particular,

$$Lg(x) = \lim_{t \downarrow 0} \frac{1}{t} \int_X (g(y) - g(x)) P(t, x, dy)$$

for all $x \in X$, whenever $g \in \mathscr{D}(L)$.

For a pure jump-type process with bounded jump rate function φ as considered in Sect. 2.3.3 and with transition function of the form (2.21) we have

$$\int_X (g(y) - g(x)) P(t, x, dy) = (e^{-\lambda t} - 1) g(x) + e^{-\lambda t} \lambda t \int_X g(y) \bar{P}(x, dy) + \sum_{n=2}^{\infty} e^{-\lambda t} \frac{(\lambda t)^n}{n!} \int_X g(y) \bar{P}^n(x, dy),$$

which shows that $\mathscr{D}(L)$ consists of all bounded measurable functions and that

$$Lg(x) = \lambda \int_X (g(y) - g(x)) \bar{P}(x, dy), \quad x \in X, \ g \in B(X).$$

Using (2.20) we conclude that the infinitesimal generator of this process is of the form

$$Lg(x) = \varphi(x) \int_X (g(y) - g(x)) P(x, dy), \quad x \in X, \ g \in B(X).$$

The significance of the infinitesimal generator $(L, \mathscr{D}(L))$ is related to the *Dynkin formula*, which states that if $g \in \mathscr{D}(L)$ then the process

$$\eta(t) = g(\xi(t)) - g(\xi(0)) - \int_0^t Lg(\xi(r))\, dr$$

is a *martingale*, i.e. the random variable $\eta(t)$ is integrable, $\mathscr{F}_t$-measurable for each $t \geq 0$, and

$$\mathbb{E}(\eta(t + s) | \mathscr{F}_s) = \eta(s)$$

for all $t, s \geq 0$. In particular, the Dynkin formula holds if $\xi(t)$ is a Markov process with a metric state space and right-continuous paths. To see this take g and Lg bounded and observe that

$$\mathbb{E}_x(\eta(t + s) | \mathscr{F}_s) = \mathbb{E}_x(g(\xi(t + s)) | \mathscr{F}_s) - g(x) - \int_0^{t+s} \mathbb{E}_x(Lg(\xi(r)) | \mathscr{F}_s)\, dr.$$

Since $Lg(\xi(r))$ is $\mathscr{F}_s$-measurable for all $r \leq s$, we can write

$$\int_0^{t+s} \mathbb{E}_x(Lg(\xi(r))|\mathscr{F}_s)\,dr = \int_0^s Lg(\xi(r))\,dr + \int_0^t \mathbb{E}_x(Lg(\xi(s+r))|\mathscr{F}_s)\,dr,$$

and, by the Markov property, we have

$$\mathbb{E}_x(\eta(t+s)|\mathscr{F}_s) = T(t)g(\xi(s)) - g(x) - \int_0^s Lg(\xi(r))\,dr - \int_0^t T(r)(Lg)(\xi(s))\,dr,$$

which gives the claim, by using the identity (see (3.4) in Sect. 3.1.2)

$$T(t)g = g + \int_0^t T(r)(Lg)\,dr, \quad t \geq 0.$$

There are several different versions of generators for Markov processes. For example, one can consider instead of the uniform convergence in $B(X)$, the pointwise convergence or the so-called bounded pointwise convergence (see [35]). Another approach, given by [28], introduces the extended generator using the concept of local martingales and allowing unbounded functions in the domain of the generator; here we adopt this definition. Let $M(X)$ be the space of all measurable functions $g\colon X \to \mathbb{R}$. An operator $\widetilde{L}$ is called the *extended generator* of the Markov process ξ, if its domain $\mathscr{D}(\widetilde{L})$ consists of those $g \in M(X)$ for which there exists $f \in M(X)$ such that for each $x \in X$, $t > 0$,

$$\mathbb{E}_x(g(\xi(t))) = g(x) + \mathbb{E}_x\left(\int_0^t f(\xi(r))\,dr\right)$$

and

$$\int_0^t \mathbb{E}_x(|f(\xi(r))|)\,dr < \infty,$$

in which case we define $\widetilde{L}g = f$.

2.3.5 Existence of PDMPs

In this section, we consider the general setting from Sect. 1.19. We assume that (X, Σ) is a Borel space and that (π, Φ, P) are three characteristics representing, respectively, a semiflow, a survival function, and a jump distribution, being a transition probability from $X \cup \Gamma$ to X where Γ is the active boundary. We assume that $P(x, X \setminus \{x\}) = 1$ for all $x \in X \cup \Gamma$. Since (X, Σ) is a Borel space, we can find a measurable mapping $\kappa\colon (X \cup \Gamma) \times [0, 1] \to X$ such that

$$P(x, B) = \mathrm{Leb}\{r \in [0, 1]\colon \kappa(x, r) \in B\}, \quad x \in X \cup \Gamma, \ \ B \in \Sigma. \tag{2.22}$$

We extend the state space and the characteristics to $(X_\Delta, \Sigma_\Delta)$ as described in Sect. 1.19. We define $\kappa(\Delta, r) = \Delta$ for $r \in [0, 1]$. Thus formula (2.22) remains valid for $x \in X_\Delta \cup \Gamma$. We extend every function g defined on X to X_Δ by setting $g(\Delta) = 0$. For each $x \in X_\Delta$ we define the generalized inverse of $t \mapsto \Phi_x(t)$ by

$$\Phi_x^{\leftarrow}(q) = \inf\{t : \Phi_x(t) \le q\}, \quad q \ge 0. \tag{2.23}$$

If ϑ is a random variable uniformly distributed on $[0, 1]$, then we have

$$\mathbb{P}(\sigma > t) = \Phi_x(t), \quad t \in [0, \infty], \quad \text{where } \sigma = \Phi_x^{\leftarrow}(\vartheta).$$

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space and let $(\vartheta_n)_{n\ge 1}$ be a sequence of independent random variables with uniform distribution on $[0, 1]$. We define $\tau_0 = \sigma_0 = 0$, $\xi_0 = x$ and

$$\sigma_1 = \Phi_{\xi_0}^{\leftarrow}(\vartheta_1), \quad \tau_1 = \sigma_1 + \tau_0,$$

and we set $\xi(t) = \pi(t - \tau_0, \xi_0)$ for $t < \tau_1$. Since the function $t \mapsto \pi(t, x)$ has a left-hand limit, which belongs to the set $X_\Delta \cup \Gamma$, we can define

$$\xi_1 = \kappa(\xi(\tau_1^-), \vartheta_2) \quad \text{and} \quad \xi(\tau_1) = \xi_1.$$

On the set $\{\tau_1 = \infty\}$ the process is defined for all times t. On $\{\tau_1 < \infty\}$ we continue the construction of the process inductively. We define $\sigma_2 = \Phi_{\xi_1}^{\leftarrow}(\vartheta_3)$, $\tau_2 = \sigma_2 + \tau_1$, and we set

$$\xi(t) = \pi(t - \tau_1, \xi_1) \quad \text{if } \tau_1 \le t < \tau_2, \quad \xi(\tau_2) = \xi_2 = \kappa(\xi(\tau_2^-), \vartheta_4),$$

and so on. Consequently, we define the minimal process $\{\xi(t)\}_{t\ge 0}$ starting at $\xi(0) = x$ by

$$\xi(t) = \begin{cases} \pi(t - \tau_n, \xi_n), & \text{if } \tau_n \le t < \tau_{n+1},\ n \ge 0, \\ \Delta, & \text{if } t \ge \tau_\infty, \end{cases} \tag{2.24}$$

where

$$\tau_n = \sigma_n + \tau_{n-1}, \quad \sigma_n = \Phi_{\xi_{n-1}}^{\leftarrow}(\vartheta_{2n-1}), \quad \xi_n = \kappa(\xi(\tau_n^-), \vartheta_{2n}), \quad n \ge 1. \tag{2.25}$$

Let $N(t)$ be the number of jump times τ_n in the time interval $[0, t]$

$$N(t) = \sup\{n \ge 0 : \tau_n \le t\}. \tag{2.26}$$

Then $N(t) = 0$ if $t < \tau_1$, $N(t) = n$ if and only if $\tau_n \le t < \tau_{n+1}$, and $N(t) = \infty$ for $t \ge \tau_\infty$. If we set $\xi_\infty = \Delta$ and $\tau_{\infty+1} = \infty$ then we have

$$\xi(t) = \pi(t - \tau_n, \xi_n) \quad \text{on} \quad \{\tau_n \le t < \tau_{n+1}\}$$

for some $n \in \bar{\mathbb{N}} = \{0, 1, \ldots\} \cup \{\infty\}$, $t \in [0, \infty]$, which we can rewrite as

$$\xi(t) = \pi(t - \tau_{N(t)}, \xi_{N(t)}), \quad t \in [0, \infty].$$

Theorem 2.4 *The minimal process $\xi(t)$, $t \geq 0$, as defined in (2.24) is a strong Markov process.*

We outline the main steps of the proof. It is similar to the proof given in [28, 29] for processes with Euclidean state space. Let $\mathscr{F}_t$ be the filtration generated by $\xi(t)$. Then each τ_n is a stopping time. Observe that for any $t \geq 0$

$$\mathscr{F}_t = \sigma(\mathbf{1}_B(\xi_k)\mathbf{1}_{[0,r]}(\tau_k) \colon k \in \bar{\mathbb{N}},\ B \in \Sigma_\Delta,\ 0 \leq r \leq t).$$

It is easy to see that $\mathscr{F}_{\tau_n} = \sigma(\tau_k, \xi_k : k \leq n)$, $n \in \bar{\mathbb{N}}$, and that

$$\mathscr{F}_s \cap \{\tau_n \leq s < \tau_{n+1}\} = \mathscr{F}_{\tau_n} \cap \{\tau_n \leq s < \tau_{n+1}\}, \quad n \in \bar{\mathbb{N}}.$$

Note that for $q, r \geq 0$ we have

$$\Phi_x^{\leftarrow}(q) > r \quad \Longleftrightarrow \quad \Phi_x(r) > q.$$

Thus we obtain

$$\begin{aligned}\mathbb{P}(\tau_{n+1} > r|\mathscr{F}_{\tau_n}) = \mathbb{P}(\sigma_{n+1} > r - \tau_n|\mathscr{F}_{\tau_n}) &= \mathbb{P}(\Phi_{\xi_n}^{\leftarrow}(\vartheta_{2n+1}) > r - \tau_n|\mathscr{F}_{\tau_n}) \\ &= \Phi_{\xi_n}(r - \tau_n)\mathbf{1}_{\{\tau_n \leq r\}} + \mathbf{1}_{\{\tau_n > r\}},\end{aligned}$$

which implies that

$$\mathbb{P}(\tau_{n+1} > t + s|\mathscr{F}_s) = \Phi_{\xi(s)}(t) \quad \text{on } \{\tau_n \leq s < \tau_{n+1}\} \tag{2.27}$$

and leads to the weak Markov property. To show the strong Markov property, we take a stopping time τ. We have

$$\mathscr{F}_\tau \cap \{\tau_n \leq \tau < \tau_{n+1}\} = \mathscr{F}_{\tau_n} \cap \{\tau_n \leq \tau < \tau_{n+1}\}, \quad n \in \bar{\mathbb{N}},$$

and for each n there exists $\mathscr{F}_{\tau_n}$-measurable random variable ζ_n such that

$$\tau\mathbf{1}_{\{\tau < \tau_{n+1}\}} = \zeta_n\mathbf{1}_{\{\tau < \tau_{n+1}\}},$$

so that (2.27) remains valid for $s = \tau$.

Let $\mathbb{P}_x$ be the distribution of the process $\xi(t)$ starting at x. The transition probability function is given by

$$P(t, x, B) = \mathbb{P}_x(\xi(t) \in B) = \mathbb{P}_x(\xi(t) \in B,\ t < \tau_\infty) + \mathbb{P}_x(\xi(t) \in B,\ t \geq \tau_\infty).$$

For $x = \Delta$ we have $\xi(t) = \Delta$, thus $P(t, \Delta, B) = \delta_\Delta(B)$ for all $t \geq 0$. If $x \in X$ and $\Delta \notin B$, then $\mathbb{P}_x(\xi(t) \in B,\ t \geq \tau_\infty) = 0$ for all t. Thus for any $x \in X$ and $B \in \Sigma$ we have

$$P(t, x, B) = \mathbb{P}_x(\xi(t) \in B,\ t < \tau_\infty) = \sum_{n=0}^{\infty} \mathbb{P}_x(\xi(t) \in B,\ \tau_n \leq t < \tau_{n+1}). \quad (2.28)$$

Note that if

$$\mathbb{E}_x(N(t)) = \mathbb{E}_x\Big(\sum_n \mathbf{1}_{(0,t]}(\tau_n)\Big) < \infty \quad \text{for all } t > 0,\ x \in X, \quad (2.29)$$

then the process ξ is non-explosive, i.e. $\tau_\infty = \infty$ a.s. In that case we have $P(t, x, X) = 1$ for all $t > 0$ and $x \in X$.

Remark 2.2 If $P(x, \{x\}) \neq 0$ for some $x \in X$ then we can extend the state space X to $\widehat{X} = X \times \{0, 1\}$ and define a transition probability $\widehat{P}$ by

$$\widehat{P}((x, i), B \times \{1 - i\}) = P(x, B), \quad \widehat{P}((x, i), B \times \{i\}) = 0, \quad (x, i) \in (X \cup \Gamma) \times \{0, 1\}.$$

It follows from (2.22) that $\widehat{\kappa}\colon \widehat{X} \times [0, 1] \to \widehat{X}$ given by $\widehat{\kappa}((x, i), r) = (\kappa(x, r), 1 - i)$ satisfies

$$\widehat{P}((x, i), B \times \{j\}) = \mathrm{Leb}\{r \in [0, 1]\colon \widehat{\kappa}((x, i), r) \in B \times \{j\}\}.$$

We define the semiflow $\widehat{\pi}$ and the survival function $\widehat{\Phi}$ by

$$\widehat{\pi}(t, x, i) = (\pi(t, x), i) \quad \text{and} \quad \widehat{\Phi}_{(x,i)}(t) = \Phi_x(t), \quad (x, i) \in \widehat{X},\ t \geq 0.$$

Using the characteristics $(\widehat{\pi}, \widehat{\Phi}, \widehat{P})$ we construct the process $\widehat{\xi}(t) = (\xi(t), i(t))$, $t \geq 0$, on $\widehat{X}_\Delta = X_\Delta \times \{0, 1\}$ as in (2.24). It is strong Markov by Theorem 2.4. Its restriction ξ to the state space X remains a strong Markov process.

2.3.6 Transition Functions and Generators of PDMPs

We consider the minimal PDMP ξ with characteristics (π, Φ, P) and jump times (τ_n) as given in Sect. 2.3.5. Let $\mathbb{P}_x$ be the distribution of the process $\xi(t)$ starting at x. For any non-negative measurable functions h defined on $X_\Delta \times [0, \infty]$, we have

$$\mathbb{E}_x[h(\xi(\tau_1), \tau_1)] = \int_{X_\Delta \times [0,\infty]} h(y, s) P(\pi(s^-, x), dy) \Phi_x(ds).$$

We define the transition kernel

$$K(x, B \times J) = \mathbb{E}_x[\mathbf{1}_B(\xi(\tau_1))\mathbf{1}_J(\tau_1)], \quad x \in X_\Delta,$$

for $B \in \Sigma_\Delta$, $J \in \mathscr{B}([0,\infty])$. The strong Markov property of the process $\xi(t)$ at τ_n implies that the sequence $(\xi(\tau_n), \tau_n)$, $n \ge 0$, is a Markov chain on $X_\Delta \times [0,\infty]$ such that for $B \in \Sigma_\Delta$ and $J \in \mathscr{B}([0,\infty])$

$$\mathbb{P}(\xi(\tau_{n+1}) \in B, \tau_{n+1} - \tau_n \in J | \mathscr{F}_{\tau_n}) = K(\xi(\tau_n), B \times J).$$

We have the iterative formula

$$K^n(x, B \times J) = \mathbb{P}_x(\xi(\tau_n) \in B, \tau_n \in J) = \int_{X_\Delta \times [0,\infty]} K^{n-1}(y, B \times (J - s))K(x, dy, ds)$$

for $n \ge 1$, $K^1 = K$, and $K^0(x, B \times J) = \mathbf{1}_B(x)\delta_0(J)$. Note that

$$\mathbb{P}_x(\xi(t) \in B, t < \tau_1) = \mathbf{1}_B(\pi(t,x))\mathbb{P}_x(\tau_1 > t) = \mathbf{1}_B(\pi(t,x))\Phi_x(t).$$

Since $\xi(t) = \pi(t - \tau_n, \xi(\tau_n))$ on $\{\tau_n \le t < \tau_{n+1}\}$, it follows form (2.28) that

$$P(t, x, B) = \sum_{n=0}^{\infty} \int_{X \times [0,t]} \mathbf{1}_B(\pi(t-s, y))\Phi_y(t-s)K^n(x, dy, ds) \tag{2.30}$$

for all $x \in X$, $B \in \Sigma$, $t > 0$.

We now show that the transition function P of the process ξ satisfies the following *Kolmogorov equation*

$$P(t, x, B) = \mathbf{1}_B(\pi(t,x))\Phi_x(t) + \int_0^t \int_X P(t-s, y, B)K(x, dy, ds) \tag{2.31}$$

for $x \in X$, $t > 0$, $B \in \Sigma$. To this end define for each $n \ge 0$ and $t \ge 0$

$$P_n(t, x, B) = \mathbb{P}_x(\xi(t) \in B,\ t < \tau_{n+1}), \quad x \in X,\ B \in \Sigma. \tag{2.32}$$

It follows from the monotone convergence theorem and (2.28) that

$$P_n(t, x, B) = \mathbb{E}_x(\mathbf{1}_B(\xi(t))\mathbf{1}_{\{t<\tau_{n+1}\}}) \uparrow \mathbb{E}_x(\mathbf{1}_B(\xi(t))\mathbf{1}_{\{t<\tau_\infty\}}) = P(t, x, B).$$

For any non-negative measurable function g we have

$$\mathbb{E}_x g(\xi(t))\mathbf{1}_{\{t<\tau_{n+1}\}} = \mathbb{E}_x g(\xi(t))\mathbf{1}_{\{t<\tau_1\}} + \mathbb{E}_x g(\xi(t))\mathbf{1}_{\{\tau_1 \le t<\tau_{n+1}\}}$$

and the strong Markov property implies that

$$\mathbb{E}_x \mathbf{1}_B(\xi(t))\mathbf{1}_{\{\tau_1 \le t<\tau_{n+1}\}} = \int_{X \times [0,t]} P_{n-1}(t-s, y, B)K(x, dy, ds).$$

Hence, the monotone convergence theorem completes the proof of (2.31).

Let $M(X)_+$ (respectively $B(X)_+$) be the space of all non-negative (bounded) measurable functions on X. We define

$$T_n(t)g(x) = \int_X g(y)P_n(t, x, dy) \tag{2.33}$$

for $t \geq 0, x \in X, g \in M(X)_+, n \geq 0$, where P_n is as in (2.32). Let T be the transition operator corresponding to the jump distribution P. It is defined for $g \in M(X)_+$ by

$$Tg(x) = \int_X g(y)P(x, dy), \quad x \in X.$$

Then, for $t \geq 0$, we have

$$\begin{aligned}\int_{X\times[0,t]} T_{n-1}(t-s)g(y)K(x, dy, ds) &= \int_0^t \int_X T_{n-1}(t-s)g(y)P(\pi(s^-, x), dy)\Phi_x(ds)\\ &= \int_0^t T(T_{n-1}(t-s)g)(\pi(s^-, x))\Phi_x(ds).\end{aligned}$$

We now suppose that the semiflow is continuous in X, i.e. $\pi(s^-, x) = \pi(s, x)$ for all $s < t_*(x), x \in X$. We consider Φ as in (1.38) defined with the help of a jump rate function φ. Then

$$\int_0^t T(T_{n-1}(t-s)g)(\pi(s, x))\Phi_x(ds) = \int_0^t T(T_{n-1}(t-s)g)(\pi(s, x))\varphi(\pi(s, x))\Phi_x(s)\, ds$$

and we obtain

$$T_n(t)\mathbf{1}_B(x) = T_0(t)\mathbf{1}_B(x) + \int_0^t T_0(s)(\varphi T(T_{n-1}(t-s)\mathbf{1}_B))(x)\, ds \tag{2.34}$$

for all $x \in X, t \geq 0$, and $n \geq 1$.

We conclude this section with a description of the extended generator $(\widetilde{L}, \mathscr{D}(\widetilde{L}))$ in the case when the active boundary might be non-empty, the survival function is as in (1.41), and the minimal process ξ satisfies (2.29). In particular, if the jump rate function φ in condition (1.41) is bounded then (2.29) holds. Given the active boundary Γ defined in (1.40) we write that $g \in M_\Gamma(X)$ if $g\colon X \to \mathbb{R}$ is measurable and the function $t \mapsto g(\pi(t, x))$ has a finite limit as $t \to t_*(x)$ for $x \in X$ with finite $t_*(x)$, where $t_*(x)$ is the exit time from X as defined in (1.39). If $g \in M_\Gamma(X)$ has the following properties

(1) for each $x \in X$ the function $t \mapsto g(\pi(t, x))$ is absolutely continuous on $(0, t_*(x))$,
(2) for each $x \in \Gamma$ we have

$$g(x) = \int_X g(y)P(x, dy),$$

(3) for each $t \geq 0$, $x \in X$,

$$\mathbb{E}_x\left(\sum_{\tau_n \leq t} |g(\xi(\tau_n)) - g(\xi(\tau_n^-))|\right) < \infty,$$

then $g \in \mathscr{D}(\widetilde{L})$ and

$$\widetilde{L}g(x) = \widetilde{L}_0 g(x) + \varphi(x)\int_X (g(y) - g(x))P(x, dy), \quad x \in X, \tag{2.35}$$

with $\widetilde{L}_0 g$ defined by

$$g(\pi(t, x)) - g(x) = \int_0^t \widetilde{L}_0 g(\pi(s, x))\, ds, \quad t < t_*(x),\ x \in X.$$

A more general condition instead of (3) characterizes all elements of the domain of the extended generator as defined and showed in [28, 29]. Note that if g is bounded and condition (2.29) holds, then g satisfies condition (3).

http://www.springer.com/978-3-319-61293-5

Piecewise Deterministic Processes in Biological Models
Rudnicki, R.; Tyran-Kamińska, M.
2017, IX, 169 p. 10 illus., Softcover
ISBN: 978-3-319-61293-5