

Chapter 6

General Theory of Markov Processes

Our goal in this chapter is to give a concise introduction to the main ideas of the theory of continuous time Markov processes. Markov processes form a fundamental class of stochastic processes, with many applications in real life problems outside mathematics. The reason why Markov processes are so important comes from the so-called Markov property, which enables many explicit calculations that would be intractable for more general random processes. Although the theory of Markov processes is by no means the central topic of this book, it will play a significant role in the next chapters, in particular in our discussion of stochastic differential equations. In fact the whole invention of Itô's stochastic calculus was motivated by the study of the Markov processes obtained as solutions of stochastic differential equations, which are also called diffusion processes.

This chapter is mostly independent of the previous ones, even though Brownian motion is used as a basic example, and the martingale theory developed in Chap. 3 plays an important role. After a section dealing with the general definitions and the problem of existence, we focus on the particular case of Feller processes, and in that framework we introduce the key notion of the generator. We establish regularity properties of Feller processes as consequences of the analogous results for supermartingales. We then discuss the strong Markov property, and we conclude the chapter by presenting three important classes of Markov processes.

6.1 General Definitions and the Problem of Existence

Let (E, \mathcal{E}) be a measurable space. A *Markovian transition kernel* from E into E is a mapping $Q : E \times \mathcal{E} \rightarrow [0, 1]$ satisfying the following two properties:

- (i) For every $x \in E$, the mapping $\mathcal{E} \ni A \mapsto Q(x, A)$ is a probability measure on (E, \mathcal{E}) .
- (ii) For every $A \in \mathcal{E}$, the mapping $E \ni x \mapsto Q(x, A)$ is \mathcal{E} -measurable.

In what follows we say transition kernel instead of Markovian transition kernel.

Remark In the case where E is finite or countable (and equipped with the σ -field of all subsets of E), Q is characterized by the “matrix” $(Q(x, \{y\}))_{x,y \in E}$.

If $f : E \rightarrow \mathbb{R}$ is bounded and measurable (resp. nonnegative and measurable), the function Qf defined by

$$Qf(x) = \int Q(x, dy)f(y)$$

is also bounded and measurable (resp. nonnegative and measurable) on E . Indeed, if f is an indicator function, the measurability of Qf is just property (ii) and the general case follows from standard arguments.

Definition 6.1 A collection $(Q_t)_{t \geq 0}$ of transition kernels on E is called a *transition semigroup* if the following three properties hold.

- (i) For every $x \in E$, $Q_0(x, dy) = \delta_x(dy)$.
- (ii) For every $s, t \geq 0$ and $A \in \mathcal{E}$,

$$Q_{t+s}(x, A) = \int_E Q_t(x, dy) Q_s(y, A)$$

(Chapman–Kolmogorov identity).

- (iii) For every $A \in \mathcal{E}$, the function $(t, x) \mapsto Q_t(x, A)$ is measurable with respect to the σ -field $\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{E}$.

Let $B(E)$ be the vector space of all bounded measurable real functions on E , which is equipped with the norm $\|f\| = \sup\{|f(x)| : x \in E\}$. Then the linear mapping $B(E) \ni f \mapsto Q_t f$ is a contraction of $B(E)$. From this point of view, the Chapman–Kolmogorov identity is equivalent to the relation

$$Q_{t+s} = Q_t Q_s$$

for every $s, t \geq 0$. This allows one to view $(Q_t)_{t \geq 0}$ as a semigroup of contractions of $B(E)$.

We now consider a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, \infty)}, P)$.

Definition 6.2 Let $(Q_t)_{t \geq 0}$ be a transition semigroup on E . A *Markov process* (with respect to the filtration (\mathcal{F}_t)) with transition semigroup $(Q_t)_{t \geq 0}$ is an (\mathcal{F}_t) -adapted process $(X_t)_{t \geq 0}$ with values in E such that, for every $s, t \geq 0$ and $f \in B(E)$,

$$E[f(X_{s+t}) \mid \mathcal{F}_s] = Q_t f(X_s).$$

Remark When we speak about a Markov process X without specifying the filtration, we implicitly mean that the property of the definition holds with the canonical filtration $\mathcal{F}_t^X = \sigma(X_r, 0 \leq r \leq t)$. We may also notice that, if X is a

Markov process with respect to a filtration (\mathcal{F}_t) , it is automatically also a Markov process (with the same semigroup) with respect to (\mathcal{F}_t^X) .

The definition of a Markov process can be interpreted as follows. Taking $f = \mathbf{1}_A$, we have

$$P[X_{s+t} \in A \mid \mathcal{F}_s] = Q_t(X_s, A)$$

and in particular

$$P[X_{s+t} \in A \mid X_r, 0 \leq r \leq s] = Q_t(X_s, A).$$

Hence the conditional distribution of X_{s+t} knowing the “past” $(X_r, 0 \leq r \leq s)$ before time s is given by $Q_t(X_s, \cdot)$, and this conditional distribution only depends on the “present” state X_s . This is the *Markov property* (informally, if one wants to predict the future after time s , the past up to time s does not give more information than just the present at time s).

Consequences of the definition Let $\gamma(dx)$ be the law of X_0 . Then if $0 < t_1 < t_2 < \dots < t_p$ and $A_0, A_1, \dots, A_p \in \mathcal{E}$,

$$\begin{aligned} & P(X_0 \in A_0, X_{t_1} \in A_1, X_{t_2} \in A_2, \dots, X_{t_p} \in A_p) \\ &= \int_{A_0} \gamma(dx_0) \int_{A_1} Q_{t_1}(x_0, dx_1) \int_{A_2} Q_{t_2-t_1}(x_1, dx_2) \cdots \int_{A_p} Q_{t_p-t_{p-1}}(x_{p-1}, dx_p). \end{aligned}$$

More generally, if $f_0, f_1, \dots, f_p \in B(E)$,

$$\begin{aligned} E[f_0(X_0)f_1(X_{t_1}) \cdots f_p(X_{t_p})] &= \int \gamma(dx_0) f_0(x_0) \int Q_{t_1}(x_0, dx_1) f_1(x_1) \\ &\quad \times \int Q_{t_2-t_1}(x_1, dx_2) f_2(x_2) \cdots \int Q_{t_p-t_{p-1}}(x_{p-1}, dx_p) f_p(x_p). \end{aligned}$$

This last formula is derived from the definition by induction on p . Note that, conversely, if the latter formula holds for any choice of $0 < t_1 < t_2 < \dots < t_p$ and $f_0, f_1, \dots, f_p \in B(E)$, then $(X_t)_{t \geq 0}$ is a Markov process of semigroup $(Q_t)_{t \geq 0}$, with respect to its canonical filtration $\mathcal{F}_t^X = \sigma(X_r, 0 \leq r \leq t)$ (use a monotone class argument to see that the property of the definition holds with $\mathcal{F}_t = \mathcal{F}_t^X$, see Appendix A1).

From the preceding formulas, we see that the finite-dimensional marginals of the process X are completely determined by the semigroup $(Q_t)_{t \geq 0}$ and the law of X_0 (initial distribution).

Example If $E = \mathbb{R}^d$, we can take, for every $t > 0$ and $x \in \mathbb{R}^d$,

$$Q_t(x, dy) = p_t(y - x) dy$$

where, for $z \in \mathbb{R}^d$,

$$p_t(z) = (2\pi t)^{-d/2} \exp\left(-\frac{|z|^2}{2t}\right),$$

is the density of the Gaussian vector in \mathbb{R}^d with covariance matrix $t\text{Id}$. It is straightforward to verify that this defines a transition semigroup on \mathbb{R}^d , and the associated Markov process is d -dimensional Brownian motion (it would be more accurate to say pre-Brownian motion since we have not yet said anything about sample paths). In the case $d = 1$, compare with Corollary 2.4.

We now address the problem of the existence of a Markov process with a given semigroup. To this end, we will need a general theorem of construction of random processes, namely the Kolmogorov extension theorem. We give without proof the special case of this theorem that is of interest to us (a proof in a more general setting can be found in [64, Chapter III], see also [47, Chapter VII], and [49, Section 2.2] for the special case $E = \mathbb{R}$).

Let $\Omega^* = E^{\mathbb{R}_+}$ be the space of all mappings $\omega : \mathbb{R}_+ \rightarrow E$. We equip Ω^* with the σ -field \mathcal{F}^* generated by the coordinate mappings $\omega \mapsto \omega(t)$ for $t \in \mathbb{R}_+$. Let $F(\mathbb{R}_+)$ be the collection of all finite subsets of \mathbb{R}_+ , and, for every $U \in F(\mathbb{R}_+)$, let $\pi_U : \Omega^* \rightarrow E^U$ be the mapping which associates with every $\omega : \mathbb{R}_+ \rightarrow E$ its restriction to U . If $U, V \in F(\mathbb{R}_+)$ and $U \subset V$, we similarly write $\pi_U^V : E^V \rightarrow E^U$ for the obvious restriction mapping.

We recall that a topological space is Polish if its topology is separable (there exists a dense sequence) and can be defined by a complete metric.

Theorem 6.3 *Assume that E is a Polish space equipped with its Borel σ -field \mathcal{E} . For every $U \in F(\mathbb{R}_+)$, let μ_U be a probability measure on E^U . Assume that the collection $(\mu_U, U \in F(\mathbb{R}_+))$ is consistent in the following sense: If $U \subset V$, μ_U is the image of μ_V under π_U^V . Then there exists a unique probability measure μ on $(\Omega^*, \mathcal{F}^*)$ such that $\pi_U(\mu) = \mu_U$ for every $U \in F(\mathbb{R}_+)$.*

Remark The uniqueness of μ is an immediate consequence of the monotone class lemma (cf. Appendix A1).

The Kolmogorov extension theorem allows one to construct random processes having prescribed finite-dimensional marginals. To see this, let $(X_t)_{t \geq 0}$ be the canonical process on Ω^* :

$$X_t(\omega) = \omega(t), \quad t \geq 0.$$

If μ is a probability measure on Ω^* and $U = \{t_1, \dots, t_p\} \in F(\mathbb{R}_+)$, with $t_1 < t_2 < \dots < t_p$, then $(X_{t_1}, \dots, X_{t_p})$ can be viewed as a random variable with values in E^U , provided we identify E^U with E^p via the mapping $\omega \rightarrow (\omega(t_1), \dots, \omega(t_p))$. Furthermore, the distribution of $(X_{t_1}, \dots, X_{t_p})$ under μ is $\pi_U(\mu)$. The Kolmogorov theorem can thus be rephrased by saying that given a collection $(\mu_U, U \in F(\mathbb{R}_+))$ of

finite-dimensional marginal distributions, which satisfies the consistency condition (this condition is clearly necessary for the desired conclusion), one can construct a probability measure μ on the space Ω^* , under which the finite-dimensional marginals of the canonical process X are the measures $\mu_U, U \in F(\mathbb{R}_+)$.

Corollary 6.4 *We assume that E satisfies the assumption of the previous theorem and that $(Q_t)_{t \geq 0}$ is a transition semigroup on E . Let γ be a probability measure on E . Then there exists a (unique) probability measure P on Ω^* under which the canonical process $(X_t)_{t \geq 0}$ is a Markov process with transition semigroup $(Q_t)_{t \geq 0}$ and the law of X_0 is γ .*

Proof Let $U = \{t_1, \dots, t_p\} \in F(\mathbb{R}_+)$, with $0 \leq t_1 < \dots < t_p$. We define a probability measure P^U on E^U (identified with E^p as explained above) by setting

$$\begin{aligned} & \int P^U(dx_1 \dots dx_p) 1_A(x_1, \dots, x_p) \\ &= \int \gamma(dx_0) \int Q_{t_1}(x_0, dx_1) \int Q_{t_2-t_1}(x_1, dx_2) \dots \int Q_{t_p-t_{p-1}}(x_{p-1}, dx_p) 1_A(x_1, \dots, x_p) \end{aligned}$$

for any measurable subset A of E^U .

Using the Chapman–Kolmogorov relation, one verifies that the measures P^U satisfy the consistency condition. The Kolmogorov theorem then gives the existence (and uniqueness) of a probability measure P on Ω^* whose finite-dimensional marginals are the measures $P_U, U \in F(\mathbb{R}_+)$. By a previous observation, this implies that $(X_t)_{t \geq 0}$ is under P a Markov process with semigroup $(Q_t)_{t \geq 0}$, with respect to the canonical filtration. □

For $x \in E$, let P_x be the measure given by the preceding corollary when $\gamma = \delta_x$. Then, the mapping $x \mapsto P_x$ is measurable in the sense that $x \mapsto P_x(A)$ is measurable, for every $A \in \mathcal{F}^*$. In fact, the latter property holds when A only depends on a finite number of coordinates (in that case, there is an explicit formula for $P_x(A)$) and a monotone class argument gives the general case. Moreover, for any probability measure γ on E , the measure defined by

$$P_{(\gamma)}(A) = \int \gamma(dx) P_x(A)$$

is the unique probability measure on Ω^* under which the canonical process $(X_t)_{t \geq 0}$ is a Markov process with semigroup $(Q_t)_{t \geq 0}$ and the law of X_0 is γ .

Summarizing, the preceding corollary allows one to construct (under a topological assumption on E) a Markov process $(X_t)_{t \geq 0}$ with semigroup $(Q_t)_{t \geq 0}$, which starts with a given initial distribution. More precisely, we get a measurable collection of probability measures $(P_x)_{x \in E}$ such that the Markov process X starts from x under P_x . However, a drawback of the method that we used is the fact that it does not give any information on the regularity properties of sample paths of X – at present we cannot even assert that these sample paths are measurable. We will come back

to this question later, but this will require additional assumptions on the semigroup $(Q_t)_{t \geq 0}$.

We end this section by introducing the important notion of the resolvent.

Definition 6.5 Let $\lambda > 0$. The λ -resolvent of the transition semigroup $(Q_t)_{t \geq 0}$ is the linear operator $R_\lambda : B(E) \rightarrow B(E)$ defined by

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

for $f \in B(E)$ and $x \in E$.

Remark Property (iii) of the definition of a transition semigroup is used here to get the measurability of the mapping $t \mapsto Q_t f(x)$, which is required to make sense of the definition of $R_\lambda f(x)$.

Properties of the resolvent.

- (i) $\|R_\lambda f\| \leq \frac{1}{\lambda} \|f\|$.
- (ii) If $0 \leq f \leq 1$, then $0 \leq \lambda R_\lambda f \leq 1$.
- (iii) If $\lambda, \mu > 0$, we have $R_\lambda - R_\mu + (\lambda - \mu)R_\lambda R_\mu = 0$ (resolvent equation).

Proof Properties (i) and (ii) are very easy. Let us only prove (iii). We may assume that $\lambda \neq \mu$. Then,

$$\begin{aligned} R_\lambda(R_\mu f)(x) &= \int_0^\infty e^{-\lambda s} Q_s \left(\int_0^\infty e^{-\mu t} Q_t f dt \right) (x) ds \\ &= \int_0^\infty e^{-\lambda s} \left(\int Q_s(x, dy) \int_0^\infty e^{-\mu t} Q_t f(y) dt \right) ds \\ &= \int_0^\infty e^{-\lambda s} \left(\int_0^\infty e^{-\mu t} Q_{s+t} f(y) dt \right) ds \\ &= \int_0^\infty e^{-(\lambda-\mu)s} \left(\int_0^\infty e^{-\mu(s+t)} Q_{s+t} f(y) dt \right) ds \\ &= \int_0^\infty e^{-(\lambda-\mu)s} \left(\int_s^\infty e^{-\mu r} Q_r f(y) dr \right) ds \\ &= \int_0^\infty Q_r f(y) e^{-\mu r} \left(\int_0^r e^{-(\lambda-\mu)s} ds \right) dr \\ &= \int_0^\infty Q_r f(y) \left(\frac{e^{-\mu r} - e^{-\lambda r}}{\lambda - \mu} \right) dr \end{aligned}$$

giving the desired result. □

Example In the case of real Brownian motion, one verifies that

$$R_\lambda f(x) = \int r_\lambda(y-x)f(y) dy$$

where

$$r_\lambda(y-x) = \int_0^\infty (2\pi t)^{-1/2} \exp\left(-\frac{|y-x|^2}{2t} - \lambda t\right) dt = \frac{1}{\sqrt{2\lambda}} \exp(-|y-x|\sqrt{2\lambda}).$$

A neat way of getting the last equality is to use the formula $E[e^{-\lambda T_a}] = e^{-a\sqrt{2\lambda}}$ for the Laplace transform of the hitting time $a > 0$ by a real Brownian motion started from 0 (see formula (3.7)). By differentiating with respect to λ , we get $E[T_a e^{-\lambda T_a}] = (a/\sqrt{2\lambda})e^{-a\sqrt{2\lambda}}$ and using the density of T_a (Corollary 2.22) to rewrite $E[T_a e^{-\lambda T_a}]$, we exactly find the integral that comes up in the calculation of $r_\lambda(y-x)$.

A key motivation of the introduction of the resolvent is the fact that it allows one to construct certain supermartingales associated with a Markov process.

Lemma 6.6 *Let X be a Markov process with semigroup $(Q_t)_{t \geq 0}$ with respect to the filtration (\mathcal{F}_t) . Let $h \in B(E)$ be nonnegative and let $\lambda > 0$. Then the process*

$$e^{-\lambda t} R_\lambda h(X_t)$$

is an (\mathcal{F}_t) -supermartingale.

Proof The random variables $e^{-\lambda t} R_\lambda h(X_t)$ are bounded and thus in L^1 . Then, for every $s \geq 0$,

$$Q_s R_\lambda h = \int_0^\infty e^{-\lambda t} Q_{s+t} h dt$$

and it follows that

$$e^{-\lambda s} Q_s R_\lambda h = \int_0^\infty e^{-\lambda(s+t)} Q_{s+t} h dt = \int_s^\infty e^{-\lambda t} Q_t h dt \leq R_\lambda h.$$

Hence, for every $s, t \geq 0$,

$$E[e^{-\lambda(t+s)} R_\lambda h(X_{t+s}) \mid \mathcal{F}_t] = e^{-\lambda(t+s)} Q_s R_\lambda h(X_t) \leq e^{-\lambda t} R_\lambda h(X_t),$$

giving the desired supermartingale property. \square

6.2 Feller Semigroups

From now on, we assume that E is a metrizable locally compact topological space. We also assume that E is countable at infinity, meaning that E is a countable union of compact sets. The space E is equipped with its Borel σ -field.

Under the previous assumptions, it is well known that the space E is Polish. Moreover, one can find an increasing sequence $(K_n)_{n \geq 1}$ of compact subsets of E , such that any compact set of E is contained in K_n for some n . A function $f : E \rightarrow \mathbb{R}$ tends to 0 at infinity if, for every $\varepsilon > 0$, there exists a compact subset K of E such that $|f(x)| \leq \varepsilon$ for every $x \in E \setminus K$. This is equivalent to requiring that

$$\sup_{x \in E \setminus K_n} |f(x)| \xrightarrow{n \rightarrow \infty} 0.$$

We let $C_0(E)$ stand for the set of all continuous real functions on E that tend to 0 at infinity. The space $C_0(E)$ is a Banach space for the supremum norm

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

Definition 6.7 Let $(Q_t)_{t \geq 0}$ be a transition semigroup on E . We say that $(Q_t)_{t \geq 0}$ is a *Feller semigroup* if:

- (i) $\forall f \in C_0(E), Q_t f \in C_0(E)$;
- (ii) $\forall f \in C_0(E), \|Q_t f - f\| \rightarrow 0$ as $t \rightarrow 0$.

A Markov process with values in E is a *Feller process* if its semigroup is Feller.

Remark One can prove (see for instance [70, Proposition III.2.4]) that condition (ii) can be replaced by the seemingly weaker property

$$\forall f \in C_0(E), \forall x \in E, Q_t f(x) \xrightarrow{t \rightarrow 0} f(x).$$

We will not use this, except in one particular example at the end of this chapter.

Condition (ii) implies that, for every $s \geq 0$,

$$\lim_{t \downarrow 0} \|Q_{s+t} f - Q_s f\| = \lim_{t \downarrow 0} \|Q_s(Q_t f - f)\| = 0$$

since Q_s is a contraction of $C_0(E)$. We note that the convergence is uniform when s varies over \mathbb{R}_+ , which ensures that the mapping $t \mapsto Q_t f$ is uniformly continuous from \mathbb{R}_+ into $C_0(E)$, for any fixed $f \in C_0(E)$.

In what follows, we fix a Feller semigroup $(Q_t)_{t \geq 0}$ on E . Using property (i) of the definition and the dominated convergence theorem, one easily verifies that $R_\lambda f \in C_0(E)$ for every $f \in C_0(E)$ and $\lambda > 0$.

Proposition 6.8 *Let $\lambda > 0$, and set $\mathcal{R} = \{R_\lambda f : f \in C_0(E)\}$. Then \mathcal{R} does not depend on the choice $\lambda > 0$. Furthermore, \mathcal{R} is a dense subspace of $C_0(E)$.*

Proof If $\lambda \neq \mu$, the resolvent equation gives

$$R_\lambda f = R_\mu(f + (\mu - \lambda)R_\lambda f).$$

Hence any function of the form $R_\lambda f$ with $f \in C_0(E)$ is also of the form $R_\mu g$ for some $g \in C_0(E)$. This gives the first assertion.

Clearly, \mathcal{R} is a linear subspace of $C_0(E)$. To see that it is dense, we simply note that, for every $f \in C_0(E)$,

$$\lambda R_\lambda f = \lambda \int_0^\infty e^{-\lambda t} Q_t f \, dt = \int_0^\infty e^{-t} Q_{t/\lambda} f \, dt \xrightarrow{\lambda \rightarrow \infty} f \quad \text{in } C_0(E),$$

by property (ii) of the definition of a Feller semigroup and dominated convergence. \square

Definition 6.9 We set

$$D(L) = \{f \in C_0(E) : \frac{Q_t f - f}{t} \text{ converges in } C_0(E) \text{ when } t \downarrow 0\}$$

and, for every $f \in D(L)$,

$$Lf = \lim_{t \downarrow 0} \frac{Q_t f - f}{t}.$$

Then $D(L)$ is a linear subspace of $C_0(E)$ and $L : D(L) \rightarrow C_0(E)$ is a linear operator called the *generator* of the semigroup $(Q_t)_{t \geq 0}$. The subspace $D(L)$ is called the *domain* of L .

Let us start with two simple properties of the generator.

Proposition 6.10 *Let $f \in D(L)$ and $s > 0$. Then $Q_s f \in D(L)$ and $L(Q_s f) = Q_s(Lf)$.*

Proof Writing

$$\frac{Q_t(Q_s f) - Q_s f}{t} = Q_s \left(\frac{Q_t f - f}{t} \right)$$

and using the fact that Q_s is a contraction of $C_0(E)$, we get that $t^{-1}(Q_t(Q_s f) - Q_s f)$ converges to $Q_s(Lf)$, which gives the desired result. \square

Proposition 6.11 *If $f \in D(L)$, we have, for every $t \geq 0$,*

$$Q_t f = f + \int_0^t Q_s(Lf) \, ds = f + \int_0^t L(Q_s f) \, ds.$$

Proof Let $f \in D(L)$. For every $t \geq 0$,

$$\varepsilon^{-1}(Q_{t+\varepsilon}f - Q_t f) = Q_t(\varepsilon^{-1}(Q_\varepsilon f - f)) \xrightarrow[\varepsilon \downarrow 0]{} Q_t(Lf).$$

Moreover, the preceding convergence is uniform when t varies over \mathbb{R}_+ . This implies that, for every $x \in E$, the function $t \mapsto Q_t f(x)$ is differentiable on \mathbb{R}_+ and its derivative is $Q_t(Lf)(x)$, which is a continuous function of t . The formula of the proposition follows, also using the preceding proposition. \square

The next proposition identifies the domain $D(L)$ in terms of the resolvent operators R_λ .

Proposition 6.12 *Let $\lambda > 0$.*

- (i) *For every $g \in C_0(E)$, $R_\lambda g \in D(L)$ and $(\lambda - L)R_\lambda g = g$.*
- (ii) *If $f \in D(L)$, $R_\lambda(\lambda - L)f = f$.*

Consequently, $D(L) = \mathcal{R}$ and the operators $R_\lambda : C_0(E) \rightarrow \mathcal{R}$ and $\lambda - L : D(L) \rightarrow C_0(E)$ are the inverse of each other.

Proof

- (i) If $g \in C_0(E)$, we have for every $\varepsilon > 0$,

$$\begin{aligned} \varepsilon^{-1}(Q_\varepsilon R_\lambda g - R_\lambda g) &= \varepsilon^{-1} \left(\int_0^\infty e^{-\lambda t} Q_{\varepsilon+t} g \, dt - \int_0^\infty e^{-\lambda t} Q_t g \, dt \right) \\ &= \varepsilon^{-1} \left((1 - e^{-\lambda \varepsilon}) \int_0^\infty e^{-\lambda t} Q_{\varepsilon+t} g \, dt - \int_0^\varepsilon e^{-\lambda t} Q_t g \, dt \right) \\ &\xrightarrow[\varepsilon \rightarrow 0]{} \lambda R_\lambda g - g \end{aligned}$$

using property (ii) of the definition of a Feller semigroup (and the fact that this property implies the continuity of the mapping $t \mapsto Q_t g$ from \mathbb{R}_+ into $C_0(E)$).

The preceding calculation shows that $R_\lambda g \in D(L)$ and $L(R_\lambda g) = \lambda R_\lambda g - g$.

- (ii) Let $f \in D(L)$. By Proposition 6.11, $Q_t f = f + \int_0^t Q_s(Lf) \, ds$, hence

$$\begin{aligned} \int_0^\infty e^{-\lambda t} Q_t f(x) \, dt &= \frac{f(x)}{\lambda} + \int_0^\infty e^{-\lambda t} \left(\int_0^t Q_s(Lf)(x) \, ds \right) dt \\ &= \frac{f(x)}{\lambda} + \int_0^\infty \frac{e^{-\lambda s}}{\lambda} Q_s(Lf)(x) \, ds. \end{aligned}$$

We have thus obtained the equality

$$\lambda R_\lambda f = f + R_\lambda Lf$$

giving the result in (ii).

The last assertions of the proposition follow from (i) and (ii): (i) shows that $\mathcal{R} \subset D(L)$ and (ii) gives the reverse inclusion, then the identities in (i) and (ii) show that R_λ and $\lambda - L$ are inverse of each other. \square

Corollary 6.13 *The semigroup $(Q_t)_{t \geq 0}$ is determined by the generator L (including also the domain $D(L)$).*

Proof Let f be a nonnegative function in $C_0(E)$. Then $R_\lambda f$ is the unique element of $D(L)$ such that $(\lambda - L)R_\lambda f = f$. On the other hand, knowing $R_\lambda f(x) = \int_0^\infty e^{-\lambda t} Q_t f(x) dt$ for every $\lambda > 0$ determines the continuous function $t \mapsto Q_t f(x)$. To complete the argument, note that Q_t is characterized by the values of $Q_t f$ for every nonnegative function f in $C_0(E)$. \square

Example It is easy to verify that the semigroup $(Q_t)_{t \geq 0}$ of real Brownian motion is Feller. Let us compute its generator L . We saw that, for every $\lambda > 0$ and $f \in C_0(\mathbb{R})$,

$$R_\lambda f(x) = \int \frac{1}{\sqrt{2\lambda}} \exp(-\sqrt{2\lambda}|y - x|) f(y) dy.$$

If $h \in D(L)$, we know that there exists an $f \in C_0(\mathbb{R})$ such that $h = R_\lambda f$. Taking $\lambda = \frac{1}{2}$, we have

$$h(x) = \int \exp(-|y - x|) f(y) dy.$$

By differentiating under the integral sign (we leave the justification to the reader), we get that h is differentiable on \mathbb{R} , and

$$h'(x) = \int \operatorname{sgn}(y - x) \exp(-|y - x|) f(y) dy$$

with the notation $\operatorname{sgn}(z) = \mathbf{1}_{\{z > 0\}} - \mathbf{1}_{\{z \leq 0\}}$ (the value of $\operatorname{sgn}(0)$ is unimportant). Let us also show that h' is differentiable on \mathbb{R} . Let $x_0 \in \mathbb{R}$. Then, for $x > x_0$,

$$\begin{aligned} h'(x) - h'(x_0) &= \int \left(\operatorname{sgn}(y - x) \exp(-|y - x|) - \operatorname{sgn}(y - x_0) \exp(-|y - x_0|) \right) f(y) dy \\ &= \int_{x_0}^x \left(-\exp(-|y - x|) - \exp(-|y - x_0|) \right) f(y) dy \\ &\quad + \int_{\mathbb{R} \setminus [x_0, x]} \operatorname{sgn}(y - x_0) \left(\exp(-|y - x|) - \exp(-|y - x_0|) \right) f(y) dy. \end{aligned}$$

It follows that

$$\frac{h'(x) - h'(x_0)}{x - x_0} \xrightarrow{x \downarrow x_0} -2f(x_0) + h(x_0).$$

We get the same limit when $x \uparrow x_0$, and we thus obtain that h is twice differentiable, and $h'' = -2f + h$.

On the other hand, since $h = R_{1/2}f$, Proposition 6.12 shows that

$$\left(\frac{1}{2} - L\right)h = f$$

hence $Lh = -f + \frac{1}{2}h = \frac{1}{2}h''$.

Summarizing, we have obtained that

$$D(L) \subset \{h \in C^2(\mathbb{R}) : h \text{ and } h'' \in C_0(\mathbb{R})\}$$

and that, if $h \in D(L)$, we have $Lh = \frac{1}{2}h''$.

In fact, the preceding inclusion is an equality. To see this, we may argue in the following way. If g is a twice differentiable function such that g and g'' are in $C_0(\mathbb{R})$, we set $f = \frac{1}{2}(g - g'') \in C_0(\mathbb{R})$, then $h = R_{1/2}f \in D(L)$. By the preceding argument, h is twice differentiable and $h'' = -2f + h$. It follows that $(h - g)'' = h - g$. Since the function $h - g$ belongs to $C_0(\mathbb{R})$, it must vanish identically and we get $g = h \in D(L)$.

Remark In general, it is very difficult to determine the exact domain of the generator. The following theorem often allows one to identify elements of this domain using martingales associated with the Markov process with semigroup $(Q_t)_{t \geq 0}$.

We consider again a general Feller semigroup $(Q_t)_{t \geq 0}$. We assume that on some probability space, we are given, for every $x \in E$, a process $(X_t^x)_{t \geq 0}$ which is Markov with semigroup $(Q_t)_{t \geq 0}$, with respect to a filtration $(\mathcal{F}_t)_{t \geq 0}$, and such that $P(X_0^x = x) = 1$. To make sense of the integrals that will appear below, we also assume that the sample paths of $(X_t^x)_{t \geq 0}$ are càdlàg (we will see in the next section that this assumption is not restrictive).

Theorem 6.14 *Let $h, g \in C_0(E)$. The following two conditions are equivalent:*

- (i) $h \in D(L)$ and $Lh = g$.
- (ii) For every $x \in E$, the process

$$h(X_t^x) - \int_0^t g(X_s^x) ds$$

is a martingale, with respect to the filtration (\mathcal{F}_t) .

Proof We first prove that (i) \Rightarrow (ii). Let $h \in D(L)$ and $g = Lh$. By Proposition 6.11, we have then, for every $s \geq 0$,

$$Q_s h = h + \int_0^s Q_r g dr.$$

It follows that, for $t \geq 0$ and $s \geq 0$,

$$E[h(X_{t+s}^x) \mid \mathcal{F}_t] = Q_s h(X_t^x) = h(X_t^x) + \int_0^s Q_r g(X_t^x) dr.$$

On the other hand,

$$\begin{aligned} E\left[\int_t^{t+s} g(X_r^x) dr \mid \mathcal{F}_t\right] &= \int_t^{t+s} E[g(X_r^x) \mid \mathcal{F}_t] dr = \int_t^{t+s} Q_{r-t} g(X_t^x) dr \\ &= \int_0^s Q_r g(X_t^x) dr. \end{aligned}$$

The fact that the conditional expectation and the integral can be interchanged (in the first equality of the last display) is easy to justify using the characteristic property of conditional expectations. It follows from the last two displays that

$$E\left[h(X_{t+s}^x) - \int_0^{t+s} g(X_r^x) dr \mid \mathcal{F}_t\right] = h(X_t^x) - \int_0^t g(X_r^x) dr$$

giving property (ii).

Conversely, suppose that (ii) holds. Then, for every $t \geq 0$,

$$E\left[h(X_t^x) - \int_0^t g(X_r^x) dr\right] = h(x)$$

and on the other hand, from the definition of a Markov process,

$$E\left[h(X_t^x) - \int_0^t g(X_r^x) dr\right] = Q_t h(x) - \int_0^t Q_r g(x) dr.$$

Consequently,

$$\frac{Q_t h - h}{t} = \frac{1}{t} \int_0^t Q_r g dr \xrightarrow[t \downarrow 0]{} g$$

in $C_0(E)$, by property (ii) of the definition of a Feller semigroup. We conclude that $h \in D(L)$ and $Lh = g$. \square

Example In the case of d -dimensional Brownian motion, Itô's formula shows that, if $h \in C^2(\mathbb{R}^d)$,

$$h(X_t) - \frac{1}{2} \int_0^t \Delta h(X_s) ds$$

is a continuous local martingale. This continuous local martingale is a martingale if we furthermore assume that h and Δh are in $C_0(\mathbb{R}^d)$ (hence bounded). It then follows

from Theorem 6.14 that $h \in D(L)$ and $Lh = \frac{1}{2}\Delta h$. Recall that we already obtained this result by a direct computation of L when $d = 1$ (in fact in a more precise form since here we can only assert that $D(L) \supset \{h \in C^2(\mathbb{R}^d) : h \text{ and } \Delta h \in C_0(\mathbb{R}^d)\}$, whereas equality holds if $d = 1$).

6.3 The Regularity of Sample Paths

Our aim in this section is to show that one construct Feller processes in such a way that they have càdlàg sample paths. We consider a Feller semigroup $(Q_t)_{t \geq 0}$ on a topological space E (assumed to be metrizable, locally compact and countable at infinity as above).

Theorem 6.15 *Let $(X_t)_{t \geq 0}$ be a Markov process with semigroup $(Q_t)_{t \geq 0}$, with respect to the filtration $(\mathcal{F}_t)_{t \in [0, \infty]}$. Set $\tilde{\mathcal{F}}_\infty = \mathcal{F}_\infty$ and, for every $t \geq 0$,*

$$\tilde{\mathcal{F}}_t = \mathcal{F}_{t+} \vee \sigma(\mathcal{N}),$$

where \mathcal{N} denotes the class of all \mathcal{F}_∞ -measurable sets that have zero probability.

Then, the process $(X_t)_{t \geq 0}$ has a càdlàg modification $(\tilde{X}_t)_{t \geq 0}$, which is adapted to the filtration $(\tilde{\mathcal{F}}_t)$. Moreover, $(\tilde{X}_t)_{t \geq 0}$ is a Markov process with semigroup $(Q_t)_{t \geq 0}$, with respect to the filtration $(\tilde{\mathcal{F}}_t)_{t \in [0, \infty]}$.

Remark The filtration $(\tilde{\mathcal{F}}_t)$ is right-continuous because so is the filtration (\mathcal{F}_{t+}) and the right-continuity property is preserved when adding the class of negligible sets \mathcal{N} .

Proof Let $E_\Delta = E \cup \{\Delta\}$ be the Alexandroff compactification of E , which is obtained by adding the point at infinity Δ to E (and by definition the neighborhoods of Δ are the complements of compact subsets of E). We agree that every function $f \in C_0(E)$ is extended to a continuous function on E_Δ by setting $f(\Delta) = 0$.

Write $C_0^+(E)$ for the set of all nonnegative functions in $C_0(E)$. We can find a sequence $(f_n)_{n \geq 0}$ in $C_0^+(E)$ which separates the points of E_Δ , in the sense that, for every $x, y \in E_\Delta$ with $x \neq y$, there is an integer n such that $f_n(x) \neq f_n(y)$. Then,

$$\mathcal{H} = \{R_p f_n : p \geq 1, n \geq 0\}$$

is also a countable subset of $C_0^+(E)$ which separates the points of E_Δ (use the fact that $\|pR_p f - f\| \rightarrow 0$ when $p \rightarrow \infty$).

If $h \in \mathcal{H}$, Lemma 6.6 shows that there exists an integer $p \geq 1$ such that $e^{-pt}h(X_t)$ is a supermartingale. Let D be a countable dense subset of \mathbb{R}_+ . Then Theorem 3.17 (i) shows that the limits

$$\lim_{D \ni s \downarrow t} h(X_s), \quad \lim_{D \ni s \uparrow t} h(X_s)$$

exist simultaneously for every $t \in \mathbb{R}_+$ (the second one only for $t > 0$) outside an \mathcal{F}_∞ -measurable event N_h of zero probability. Indeed, as in the proof of Theorem 3.17, we may define the complementary event N_h^c as the set of all $\omega \in \Omega$ for which the function $D \ni s \mapsto e^{-ps}h(X_s)$ makes a finite number of upcrossings along any interval $[a, b]$ ($a, b \in \mathbb{Q}, a < b$) on every finite time interval. We then set

$$N = \bigcup_{h \in \mathcal{H}} N_h$$

in such a way that we still have $N \in \mathcal{N}$. Then if $\omega \notin N$, the limits

$$\lim_{D \ni s \downarrow t} X_s(\omega), \quad \lim_{D \ni s \uparrow t} X_s(\omega)$$

exist for every $t \geq 0$ (the second one only for $t > 0$) in E_Δ . In fact, if we assume that $X_s(\omega)$ has two distinct accumulation points in E_Δ when $D \ni s \downarrow t$, we get a contradiction by considering a function $h \in \mathcal{H}$ that separates these two points. We can then set, for every $\omega \in \Omega \setminus N$ and every $t \geq 0$,

$$\tilde{X}_t(\omega) = \lim_{D \ni s \downarrow t} X_s(\omega).$$

If $\omega \in N$, we set $\tilde{X}_t(\omega) = x_0$ for every $t \geq 0$, where x_0 is a fixed point in E . Then, for every $t \geq 0$, \tilde{X}_t is an $\tilde{\mathcal{F}}_t$ -measurable random variable with values in E_Δ . Furthermore, for every $\omega \in \Omega, t \mapsto \tilde{X}_t(\omega)$, viewed as a mapping with values in E_Δ , is càdlàg by Lemma 3.16 (this lemma shows that the functions $t \mapsto h(\tilde{X}_t(\omega))$, for $h \in \mathcal{H}$, are càdlàg, and this suffices since \mathcal{H} separates points of E).

Let us now show that $P(X_t = \tilde{X}_t) = 1$, for every fixed $t \geq 0$. Let $f, g \in C_0(E)$ and let (t_n) be a sequence in D that decreases (strictly) to t . Then,

$$\begin{aligned} E[f(\tilde{X}_t)g(\tilde{X}_t)] &= \lim_{n \rightarrow \infty} E[f(X_t)g(X_{t_n})] \\ &= \lim_{n \rightarrow \infty} E[f(X_t)Q_{t_n-t}g(X_t)] \\ &= E[f(X_t)g(X_t)] \end{aligned}$$

since $Q_{t_n-t}g \rightarrow g$ by the definition of a Feller semigroup. The preceding equality entails that the two pairs (X_t, \tilde{X}_t) and (X_t, X_t) have the same distribution and thus $P(X_t = \tilde{X}_t) = 1$.

Let us then verify that $(\tilde{X}_t)_{t \geq 0}$ is a Markov process with semigroup $(Q_t)_{t \geq 0}$ with respect to the filtration $(\tilde{\mathcal{F}}_t)$. It is enough to prove that, for every $s \geq 0, t > 0$ and $A \in \tilde{\mathcal{F}}_s, f \in C_0(E)$, we have

$$E[1_A f(\tilde{X}_{s+t})] = E[1_A Q_t f(\tilde{X}_s)].$$

Since $\tilde{X}_s = X_s$ a.s. and $\tilde{X}_{s+t} = X_{s+t}$ a.s., this is equivalent to proving that

$$E[1_A f(X_{s+t})] = E[1_A Q_t f(X_s)].$$

Because A is equal a.s. to an element of \mathcal{F}_{s+} , we may assume that $A \in \mathcal{F}_{s+}$. Let (s_n) be a sequence in D that decreases to s , so that $A \in \mathcal{F}_{s_n}$ for every n . Then, as soon as $s_n \leq s + t$, we have

$$E[1_A f(X_{s+t})] = E[1_A E[f(X_{s+t}) \mid \mathcal{F}_{s_n}]] = E[1_A Q_{s+t-s_n} f(X_{s_n})].$$

But $Q_{s+t-s_n} f$ converges (uniformly) to $Q_t f$ by properties of Feller semigroups, and since $X_{s_n} = \tilde{X}_{s_n}$ a.s. we also know that X_{s_n} converges a.s. to $\tilde{X}_s = X_s$ a.s. We thus obtain the desired result by letting n tend to ∞ .

It remains to verify that the sample paths $t \mapsto \tilde{X}_t(\omega)$ are càdlàg as E -valued mappings, and not only as E_Δ -valued mappings (we already know that, for every fixed $t \geq 0$, $\tilde{X}_t(\omega) = X_t(\omega)$ a.s. is in E with probability one, but this does not imply that the sample paths, and their left-limits, remain in E). Fix a function $g \in C_0^+(E)$ such that $g(x) > 0$ for every $x \in E$. The function $h = R_1 g$ then satisfies the same property. Set, for every $t \geq 0$,

$$Y_t = e^{-t} h(\tilde{X}_t).$$

Then Lemma 6.6 shows that $(Y_t)_{t \geq 0}$ is a nonnegative supermartingale with respect to the filtration $(\tilde{\mathcal{F}}_t)$. Additionally, we know that the sample paths of $(Y_t)_{t \geq 0}$ are càdlàg (recall that $h(\Delta) = 0$ by convention).

For every integer $n \geq 1$, set

$$T_{(n)} = \inf\{t \geq 0 : Y_t < \frac{1}{n}\}.$$

Then $T_{(n)}$ is a stopping time of the filtration $(\tilde{\mathcal{F}}_t)$ (we can apply Proposition 3.9, because $T_{(n)}$ is the first hitting time of an open set by an adapted process with càdlàg sample paths, and the filtration $(\tilde{\mathcal{F}}_t)$ is right-continuous). Consequently,

$$T = \lim_{n \rightarrow \infty} \uparrow T_{(n)}$$

is a stopping time. The desired result will follow if we can verify that $P(T < \infty) = 0$. Indeed, it is clear that, for every $t \in [0, T_{(n)})$, $\tilde{X}_t \in E$ and $\tilde{X}_{t-} \in E$, and we may redefine $\tilde{X}_t(\omega) = x_0$ (for every $t \geq 0$) for all ω belonging to the event $\{T < \infty\} \in \mathcal{N}$.

To verify that $P(T < \infty) = 0$, we apply Theorem 3.25 and the subsequent remark to $Z = Y$ and $U = T_{(n)}$, $V = T + q$, where $q > 0$ is a rational number. We get

$$E[Y_{T+q} \mathbf{1}_{\{T < \infty\}}] \leq E[Y_{T_{(n)}} \mathbf{1}_{\{T_{(n)} < \infty\}}] \leq \frac{1}{n}.$$

By letting n tend to ∞ , we thus have

$$E[Y_{T+q} \mathbf{1}_{\{T < \infty\}}] = 0,$$

hence $Y_{T+q} = 0$ a.s. on $\{T < \infty\}$. By the right-continuity of sample paths of Y , we conclude that $Y_t = 0$, for every $t \in [T, \infty)$, a.s. on $\{T < \infty\}$. But we also know that, for every integer $k \geq 0$, $Y_k = e^{-k}h(\tilde{X}_k) > 0$ a.s., since $\tilde{X}_k \in E$ a.s. This suffices to get $P(T < \infty) = 0$. \square

Remark The previous proof applies with minor modifications to the different setting where we are given the process $(X_t)_{t \geq 0}$ together with a collection $(P_x)_{x \in E}$ of probability measures such that, under P_x , $(X_t)_{t \geq 0}$ is a Markov process with semigroup $(Q_t)_{t \geq 0}$, with respect to a filtration $(\mathcal{F}_t)_{t \in [0, \infty]}$, and $P_x(X_0 = x) = 1$ (in the first section above, we saw that these properties will hold for the canonical process $(X_t)_{t \geq 0}$ on the space $\Omega^* = E^{\mathbb{R}^+}$ if the measures P_x are constructed from the semigroup $(Q_t)_{t \geq 0}$ using the Kolmogorov extension theorem). In that setting, we can define the filtration $(\tilde{\mathcal{F}}_t)_{t \in [0, \infty]}$ by

$$\tilde{\mathcal{F}}_t = \mathcal{F}_{t+} \vee \sigma(\mathcal{N}'),$$

where \mathcal{N}' denotes the class of all \mathcal{F}_∞ -measurable sets that have zero P_x -probability for every $x \in E$. By the same arguments as in the preceding proof, we can then construct an $(\tilde{\mathcal{F}}_t)$ -adapted process $(\tilde{X}_t)_{t \geq 0}$ with càdlàg sample paths, such that, for every $x \in E$,

$$P_x(\tilde{X}_t = X_t) = 1, \quad \forall t \geq 0,$$

and $(\tilde{X}_t)_{t \geq 0}$ is under P_x a Markov process with semigroup $(Q_t)_{t \geq 0}$, with respect to the filtration $(\tilde{\mathcal{F}}_t)_{t \in [0, \infty]}$, such that $P_x(\tilde{X}_0 = x) = 1$.

6.4 The Strong Markov Property

In the first part of this section, we come back to the general setting of Sect. 6.1 above, where $(Q_t)_{t \geq 0}$ is a (not necessarily Feller) transition semigroup on E . We assume here that E is a metric space (equipped with its Borel σ -field), and moreover that, for every $x \in E$, one can construct a Markov process $(X_t^x)_{t \geq 0}$ with semigroup

$(Q_t)_{t \geq 0}$ such that $X_0^x = x$ a.s. and the sample paths of X are càdlàg. In the case of a Feller semigroup, the existence of such a process follows from Corollary 6.4 and Theorem 6.15.

The space of all càdlàg paths $f : \mathbb{R}_+ \rightarrow E$ is denoted by $\mathbb{D}(E)$. This space is equipped with the σ -field \mathcal{D} generated by the coordinate mappings $f \mapsto f(t)$. For every $x \in E$, we write \mathbb{P}_x for the probability measure on $\mathbb{D}(E)$ which is the law of the random path $(X_t^x)_{t \geq 0}$. Notice that \mathbb{P}_x does not depend on the choice of X^x , nor of the probability space where X^x is defined: This follows from the fact that the finite-dimensional marginals of a Markov process are determined by its semigroup and initial distribution.

We first give a version of the (simple) Markov property, which is a simple extension of the definition of a Markov process. We use the notation \mathbb{E}_x for the expectation under \mathbb{P}_x .

Theorem 6.16 (Simple Markov property) *Let $(Y_t)_{t \geq 0}$ be a Markov process with semigroup $(Q_t)_{t \geq 0}$, with respect to the filtration $(\mathcal{F}_t)_{t \geq 0}$. We assume that the sample paths of Y are càdlàg. Let $s \geq 0$ and let $\Phi : \mathbb{D}(E) \rightarrow \mathbb{R}_+$ be a measurable function. Then,*

$$E[\Phi((Y_{s+t})_{t \geq 0}) \mid \mathcal{F}_s] = \mathbb{E}_{Y_s}[\Phi].$$

Remark The right-hand side of the last display is the composition of Y_s and of the mapping $y \mapsto \mathbb{E}_y[\Phi]$. To see that the latter mapping is measurable, it is enough to consider the case where $\Phi = 1_A$, $A \in \mathcal{D}$. When A only depends on a finite number of coordinates, there is an explicit formula, and an application of the monotone class lemma completes the argument.

Proof As in the preceding remark, it suffices to consider the case where $\Phi = 1_A$ and

$$A = \{f \in \mathbb{D}(E) : f(t_1) \in B_1, \dots, f(t_p) \in B_p\},$$

where $0 \leq t_1 < t_2 < \dots < t_p$ and B_1, \dots, B_p are measurable subsets of E . In that case, we need to verify that

$$\begin{aligned} & P(Y_{s+t_1} \in B_1, \dots, Y_{s+t_p} \in B_p \mid \mathcal{F}_s) \\ &= \int_{B_1} Q_{t_1}(Y_s, dx_1) \int_{B_2} Q_{t_2-t_1}(x_1, dx_2) \cdots \int_{B_p} Q_{t_p-t_{p-1}}(x_{p-1}, dx_p). \end{aligned}$$

We in fact prove more generally that, if $\varphi_1, \dots, \varphi_p \in B(E)$,

$$\begin{aligned} & E[\varphi_1(Y_{s+t_1}) \cdots \varphi_p(Y_{s+t_p}) \mid \mathcal{F}_s] \\ &= \int Q_{t_1}(Y_s, dx_1) \varphi_1(x_1) \int Q_{t_2-t_1}(x_1, dx_2) \varphi_2(x_2) \cdots \int Q_{t_p-t_{p-1}}(x_{p-1}, dx_p) \varphi_p(x_p). \end{aligned}$$

If $p = 1$ this is the definition of a Markov process. We then argue by induction, writing

$$\begin{aligned} E[\varphi_1(Y_{s+t_1}) \cdots \varphi_p(Y_{s+t_p}) \mid \mathcal{F}_s] \\ &= E[\varphi_1(Y_{s+t_1}) \cdots \varphi_{p-1}(Y_{s+t_{p-1}}) E[\varphi_p(Y_{s+t_p}) \mid \mathcal{F}_{s+t_{p-1}}] \mid \mathcal{F}_s] \\ &= E[\varphi_1(Y_{s+t_1}) \cdots \varphi_{p-1}(Y_{s+t_{p-1}}) Q_{t_p-t_{p-1}} \varphi_p(Y_{s+t_{p-1}}) \mid \mathcal{F}_s] \end{aligned}$$

and the desired result easily follows. \square

We now turn to the strong Markov property.

Theorem 6.17 (Strong Markov property) *Retain the assumptions of the previous theorem, and suppose in addition that $(Q_t)_{t \geq 0}$ is a Feller semigroup (in particular, E is assumed to be metrizable locally compact and countable at infinity). Let T be a stopping time of the filtration (\mathcal{F}_{t+}) , and let $\Phi : \mathbb{D}(E) \rightarrow \mathbb{R}_+$ be a measurable function. Then, for every $x \in E$,*

$$E[\mathbf{1}_{\{T < \infty\}} \Phi((Y_{T+t})_{t \geq 0}) \mid \mathcal{F}_T] = \mathbf{1}_{\{T < \infty\}} \mathbb{E}_{Y_T}[\Phi].$$

Remark We allow T to be a stopping time of (\mathcal{F}_{t+}) , which is slightly more general than saying that T is a stopping time of the filtration (\mathcal{F}_t) .

Proof We first observe that the right-hand side of the last display is \mathcal{F}_T -measurable, because $\{T < \infty\} \ni \omega \mapsto Y_T(\omega)$ is \mathcal{F}_T -measurable (Theorem 3.7) and the function $y \mapsto \mathbb{E}_y[\Phi]$ is measurable. It is then enough to show that, for $A \in \mathcal{F}_T$ fixed,

$$E[\mathbf{1}_{A \cap \{T < \infty\}} \Phi((Y_{T+t})_{t \geq 0})] = E[\mathbf{1}_{A \cap \{T < \infty\}} \mathbb{E}_{Y_T}[\Phi]].$$

As above, we can restrict our attention to the case where

$$\Phi(f) = \varphi_1(f(t_1)) \cdots \varphi_p(f(t_p))$$

where $0 \leq t_1 < t_2 < \cdots < t_p$ and $\varphi_1, \dots, \varphi_p \in B(E)$. It is in fact enough to take $p = 1$: If the desired result holds in that case, we can argue by induction, writing

$$\begin{aligned} E[\mathbf{1}_{A \cap \{T < \infty\}} \varphi_1(Y_{T+t_1}) \cdots \varphi_p(Y_{T+t_p})] \\ &= E[\mathbf{1}_{A \cap \{T < \infty\}} \varphi_1(Y_{T+t_1}) \cdots \varphi_{p-1}(Y_{T+t_{p-1}}) E[\varphi_p(Y_{T+t_p}) \mid \mathcal{F}_{T+t_{p-1}}]] \\ &= E[\mathbf{1}_{A \cap \{T < \infty\}} \varphi_1(Y_{T+t_1}) \cdots \varphi_{p-1}(Y_{T+t_{p-1}}) Q_{t_p-t_{p-1}} \varphi_p(Y_{T+t_{p-1}})]. \end{aligned}$$

We thus fix $t \geq 0$ and $\varphi \in B(E)$ and we aim to prove that

$$E[\mathbf{1}_{A \cap \{T < \infty\}} \varphi(Y_{T+t})] = E[\mathbf{1}_{A \cap \{T < \infty\}} Q_t \varphi(Y_T)].$$

We may assume that $\varphi \in C_0(E)$ (a finite measure on E is determined by its values against functions of $C_0(E)$).

On the event $T < \infty$, write $[T]_n$ for the smallest real number of the form $i2^{-n}$, $i \in \mathbb{Z}_+$, which is strictly greater than T . Then,

$$\begin{aligned}
 E[\mathbf{1}_{A \cap \{T < \infty\}} \varphi(Y_{T+i})] &= \lim_{n \rightarrow \infty} E[\mathbf{1}_{A \cap \{T < \infty\}} \varphi(Y_{[T]_n+i})] \\
 &= \lim_{n \rightarrow \infty} \sum_{i=1}^{\infty} E[\mathbf{1}_{A \cap \{(i-1)2^{-n} \leq T < i2^{-n}\}} \varphi(Y_{i2^{-n}+i})] \\
 &= \lim_{n \rightarrow \infty} \sum_{i=1}^{\infty} E[\mathbf{1}_{A \cap \{(i-1)2^{-n} \leq T < i2^{-n}\}} Q_i \varphi(Y_{i2^{-n}})] \\
 &= \lim_{n \rightarrow \infty} E[\mathbf{1}_{A \cap \{T < \infty\}} Q_T \varphi(Y_{[T]_n})] \\
 &= E[\mathbf{1}_{A \cap \{T < \infty\}} Q_T \varphi(Y_T)]
 \end{aligned}$$

giving the desired result. In the first (and in the last) equality, we use the right continuity of sample paths. In the third equality, we observe that the event

$$A \cap \{(i-1)2^{-n} \leq T < i2^{-n}\}$$

belongs to $\mathcal{F}_{i2^{-n}}$ because $A \in \mathcal{F}_T$ and T is a stopping time of the filtration (\mathcal{F}_{t+}) (use Proposition 3.6). Finally, and this is the key point, in the last equality we also use the fact that $Q_t \varphi$ is continuous, since $\varphi \in C_0(E)$ and the semigroup is Feller. \square

Remark In the special case of (linear) Brownian motion, the result of Theorem 6.17 is essentially equivalent to Theorem 2.20 stated in Chap. 2. The reason why the formulation in Theorem 2.20 looks different comes from the property of stationarity and independence of the increments of Brownian motion, which of course does not subsist in our general setting. Even for Brownian motion, the formulation of Theorem 6.17 turns out to be more appropriate in a number of situations: A convincing illustration is the proof of Proposition 7.7 (ii) in the next chapter.

6.5 Three Important Classes of Feller Processes

6.5.1 Jump Processes on a Finite State Space

In this subsection, we assume that the state space E is finite (and equipped with the discrete topology). Note that any càdlàg function $f \in \mathbb{D}(E)$ must be of the following type: There exists a real $t_1 \in (0, \infty]$ such that $f(t) = f(0)$ for every $t \in [0, t_1)$, then, if $t_1 < \infty$, there exists a real $t_2 \in (t_1, \infty]$ such that $f(t) = f(t_1) \neq f(0)$ for every $t \in [t_1, t_2)$, and so on.

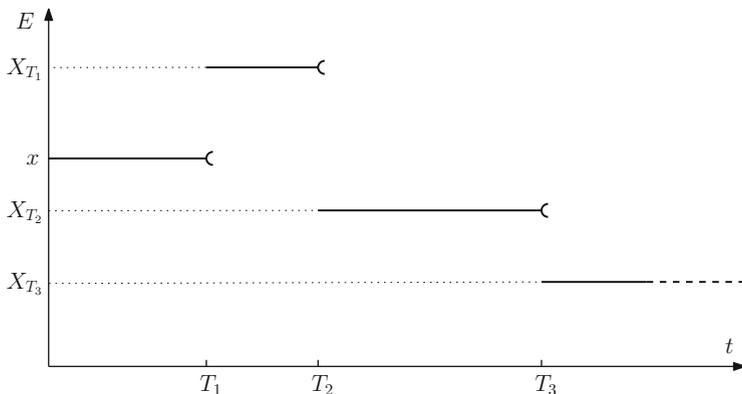


Fig. 6.1 A sample path of the jump process X under P_x

Consider a Feller semigroup $(Q_t)_{t \geq 0}$ on E . By the remark of the end of Sect. 6.3, we can construct, on a probability space Ω equipped with a right-continuous filtration $(\mathcal{F}_t)_{t \in [0, \infty]}$, a collection $(P_x)_{x \in E}$ of probability measures and a process $(X_t)_{t \geq 0}$ with càdlàg sample paths such that, under P_x , X is Markov with semigroup $(Q_t)_{t \geq 0}$ with respect to the filtration (\mathcal{F}_t) , and $P_x(X_0 = x) = 1$. As previously, E_x stands for the expectation under P_x . Since the sample paths of X are càdlàg, we know that, for every $\omega \in \Omega$, there exists a sequence

$$T_0(\omega) = 0 < T_1(\omega) \leq T_2(\omega) \leq T_3(\omega) \leq \dots \leq \infty,$$

such that $X_t(\omega) = X_0(\omega)$ for every $t \in [0, T_1(\omega))$ and, for every integer $i \geq 1$, the condition $T_i(\omega) < \infty$ implies $T_i(\omega) < T_{i+1}(\omega)$, $X_{T_i(\omega)}(\omega) \neq X_{T_{i-1}(\omega)}(\omega)$ and $X_t(\omega) = X_{T_i(\omega)}(\omega)$ for every $t \in [T_i(\omega), T_{i+1}(\omega))$. Moreover, $T_n(\omega) \uparrow \infty$ as $n \rightarrow \infty$. See Fig. 6.1.

It is not hard to verify that T_0, T_1, T_2, \dots are stopping times. For instance,

$$\{T_1 < t\} = \bigcup_{q \in [0, t) \cap \mathbb{Q}} \{X_q \neq X_0\} \in \mathcal{F}_t.$$

Recall that, for $\lambda > 0$, a positive random variable U is exponentially distributed with parameter λ if $P(U > r) = e^{-\lambda r}$ for every $r \geq 0$. In the following lemma, we make the convention that an exponential variable with parameter 0 is equal to ∞ a.s.

Lemma 6.18 *Let $x \in E$. There exists a real number $q(x) \geq 0$ such that the random variable T_1 is exponentially distributed with parameter $q(x)$ under P_x . Furthermore, if $q(x) > 0$, T_1 and X_{T_1} are independent under P_x .*

Proof Let $s, t \geq 0$. We have

$$P_x(T_1 > s + t) = E_x[\mathbf{1}_{\{T_1 > s\}} \Phi((X_{s+r})_{r \geq 0})],$$

where $\Phi(f) = \mathbf{1}_{\{f(r)=f(0), \forall r \in [0,t]\}}$ for $f \in \mathbb{D}(E)$. Using the simple Markov property (Theorem 6.16), we get

$$\begin{aligned} P_x(T_1 > s + t) &= E_x[\mathbf{1}_{\{T_1 > s\}} E_{X_s}[\Phi((X_r)_{r \geq 0})]] \\ &= E_x[\mathbf{1}_{\{T_1 > s\}} P_x(T_1 > t)] \\ &= P_x(T_1 > s) P_x(T_1 > t), \end{aligned}$$

which implies that T_1 is exponentially distributed under P_x .

Assume that $q(x) > 0$, so that $T_1 < \infty$, P_x a.s. Then, for every $t \geq 0$ and $y \in E$,

$$P_x(T_1 > t, X_{T_1} = y) = E_x[\mathbf{1}_{\{T_1 > t\}} \Psi((X_{t+r})_{r \geq 0})],$$

where for $f \in \mathbb{D}(E)$, $\Psi(f) = 0$ if f is constant, and otherwise $\Psi(f) = \mathbf{1}_{\{\gamma_1(f)=y\}}$, if $\gamma_1(f)$ is the value of f after its first jump. We thus get

$$\begin{aligned} P_x(T_1 > t, X_{T_1} = y) &= E_x[\mathbf{1}_{\{T_1 > t\}} E_{X_t}[\Psi((X_r)_{r \geq 0})]] \\ &= E_x[\mathbf{1}_{\{T_1 > t\}} P_x(X_{T_1} = y)] \\ &= P_x(T_1 > t) P_x(X_{T_1} = y), \end{aligned}$$

which gives the desired independence. \square

Points x such that $q(x) = 0$ are absorbing states for the Markov process, in the sense that $P_x(X_t = x, \forall t \geq 0) = 1$.

For every $x \in E$ such that $q(x) > 0$, and every $y \in E$, we set

$$\Pi(x, y) = P_x(X_{T_1} = y).$$

Note that $\Pi(x, \cdot)$ is a probability measure on E , and $\Pi(x, x) = 0$.

Proposition 6.19 *Let L denote the generator of $(Q_t)_{t \geq 0}$. Then $D(L) = C_0(E) = B(E)$, and, for every $\varphi \in B(E)$, for every $x \in E$:*

- if $q(x) = 0$, $L\varphi(x) = 0$;
- if $q(x) > 0$,

$$L\varphi(x) = q(x) \sum_{y \in E, y \neq x} \Pi(x, y)(\varphi(y) - \varphi(x)) = \sum_{y \in E} L(x, y) \varphi(y),$$

where

$$L(x, y) = \begin{cases} q(x)\Pi(x, y) & \text{if } y \neq x, \\ -q(x) & \text{if } y = x. \end{cases}$$

Proof Let $\varphi \in B(E)$ and $x \in E$. If $q(x) = 0$, it is trivial that $Q_t\varphi(x) = \varphi(x)$ and so

$$\lim_{t \downarrow 0} \frac{Q_t\varphi(x) - \varphi(x)}{t} = 0.$$

Suppose then that $q(x) > 0$. We first observe that

$$P_x(T_2 \leq t) = O(t^2) \tag{6.1}$$

as $t \rightarrow 0$. Indeed, using the strong Markov property at T_1 ,

$$P_x(T_2 \leq t) \leq P_x(T_1 \leq t, T_2 \leq T_1 + t) = E_x[\mathbf{1}_{\{T_1 \leq t\}} P_{X_{T_1}}(T_1 \leq t)],$$

and we can bound

$$P_{X_{T_1}}(T_1 \leq t) \leq \sup_{y \in E} P_y(T_1 \leq t) \leq t \sup_{y \in E} q(y),$$

giving the desired result since we have also $P_x(T_1 \leq t) \leq q(x)t$.

It follows from (6.1) that

$$\begin{aligned} Q_t\varphi(x) &= E_x[\varphi(X_t)] = E_x[\varphi(X_t) \mathbf{1}_{\{t < T_1\}}] + E_x[\varphi(X_{T_1}) \mathbf{1}_{\{T_1 \leq t\}}] + O(t^2) \\ &= \varphi(x) e^{-q(x)t} + (1 - e^{-q(x)t}) \sum_{y \in E, y \neq x} \Pi(x, y) \varphi(y) + O(t^2), \end{aligned}$$

using the independence of T_1 and X_{T_1} and the definition of $\Pi(x, y)$. We conclude that

$$\frac{Q_t\varphi(x) - \varphi(x)}{t} \xrightarrow{t \rightarrow 0} -q(x)\varphi(x) + q(x) \sum_{y \in E, y \neq x} \Pi(x, y) \varphi(y),$$

and this completes the proof. □

In particular, taking $\varphi(y) = \mathbf{1}_{\{y\}}$, we have if $y \neq x$,

$$L(x, y) = \frac{d}{dt} P_x(X_t = y)|_{t=0},$$

so that $L(x, y)$ can be interpreted as the instantaneous rate of transition from x to y .

The next proposition provides a complete description of the sample paths of X under P_x . For the sake of simplicity, we assume that there are no absorbing states, but the reader will easily extend the statement to the general case.

Proposition 6.20 *We assume that $q(y) > 0$ for every $y \in E$. Let $x \in E$. Then, P_x a.s., the jump times $T_1 < T_2 < T_3 < \dots$ are all finite and the sequence $X_0, X_{T_1}, X_{T_2}, \dots$ is under P_x a discrete Markov chain with transition kernel Π started from x . Furthermore, conditionally on $(X_0, X_{T_1}, X_{T_2}, \dots)$, the random variables $T_1 - T_0, T_2 - T_1, \dots$ are independent and, for every integer $i \geq 0$, the conditional distribution of $T_{i+1} - T_i$ is exponential with parameter $q(X_{T_i})$.*

Proof An application of the strong Markov property shows that all stopping times T_1, T_2, \dots are finite P_x a.s. Then, let $y, z \in E$, and $f_1, f_2 \in B(\mathbb{R}_+)$. By the strong Markov property at T_1 ,

$$\begin{aligned} & E_x[\mathbf{1}_{\{X_{T_1}=y\}}f_1(T_1)\mathbf{1}_{\{X_{T_2}=z\}}f_2(T_2-T_1)] \\ &= E_x[\mathbf{1}_{\{X_{T_1}=y\}}f_1(T_1)E_{X_{T_1}}[\mathbf{1}_{\{X_{T_1}=z\}}f_2(T_1)]] \\ &= \Pi(x, y)\Pi(y, z)\int_0^\infty ds_1 e^{-q(x)s_1}f_1(s_1)\int_0^\infty ds_2 e^{-q(y)s_2}f_2(s_2). \end{aligned}$$

Arguing by induction, we get for every $y_1, \dots, y_p \in E$ and $f_1, \dots, f_p \in B(\mathbb{R}_+)$,

$$\begin{aligned} & E_x[\mathbf{1}_{\{X_{T_1}=y_1\}}\mathbf{1}_{\{X_{T_2}=y_2\}}\cdots\mathbf{1}_{\{X_{T_p}=y_p\}}f_1(T_1)f_2(T_2-T_1)\cdots f_p(T_p-T_{p-1})] \\ &= \Pi(x, y_1)\Pi(y_1, y_2)\cdots\Pi(y_{p-1}, y_p)\prod_{i=1}^p\left(\int_0^\infty ds e^{-q(y_{i-1})s}f_i(s)\right), \end{aligned}$$

where $y_0 = x$ by convention. The various assertions of the proposition follow. \square

Jump processes play an important role in various models of applied probability, in particular in reliability and in queueing theory. In such applications, one usually starts from the transition rates of the process. It is thus important to know whether, given a collection $(q(x))_{x \in E}$ of nonnegative real numbers and, for every x such that $q(x) > 0$, a probability measure $\Pi(x, \cdot)$ on E such that $\Pi(x, x) = 0$, there exists a corresponding Feller semigroup $(Q_t)_{t \geq 0}$ and therefore an associated Markov process. The answer to this question is yes, and one can give two different arguments:

- *Probabilistic method.* Use the description of Proposition 6.20 (or its extension to the case where there are absorbing states) to construct the process $(X_t)_{t \geq 0}$ starting from any $x \in E$, and thus the semigroup $(Q_t)_{t \geq 0}$ via the formula $Q_t \varphi(x) = E_x[\varphi(X_t)]$.

- *Analytic method.* Define the generator L via the formulas of Proposition 6.19, and observe that the semigroup $(Q_t)_{t \geq 0}$, if it exists, must solve the differential equation

$$\frac{d}{dt} Q_t(x, y) = Q_t L(x, y)$$

by Proposition 6.11. This leads to

$$Q_t = \exp(tL),$$

in the sense of the exponential of matrices. Since $\lambda \text{Id} + L$ has nonnegative entries if $\lambda > 0$ is large enough, one immediately gets that Q_t has nonnegative entries. Writing $\mathbf{1}$ for the vector $(1, 1, \dots, 1)$, the property $L\mathbf{1} = 0$ gives $Q_t \mathbf{1} = \mathbf{1}$, so that $(Q_t(x, \cdot))_{x \in E}$ defines a transition kernel. Finally, the property $\exp((s+t)L) = \exp(sL)\exp(tL)$ gives the Chapman–Kolmogorov property, and we get that $(Q_t)_{t \geq 0}$ is a transition semigroup on E , whose Feller property is also immediate.

Many of the preceding results can be extended to Feller Markov processes on a *countable* state space E . Note, however, that certain difficulties arise in the question of the existence of a process with given transition rates. In fact, starting from the probabilistic description of Proposition 6.20, one needs to avoid the possibility of an accumulation of jumps in a finite time interval, which may occur if the rates $(q(y), y \in E)$ are unbounded – of course this problem does not occur when E is finite.

6.5.2 Lévy Processes

Consider a real process $(Y_t)_{t \geq 0}$ which satisfies the following three assumptions:

- $Y_0 = 0$ a.s.
- For every $0 \leq s \leq t$, the variable $Y_t - Y_s$ is independent of $(Y_r, 0 \leq r \leq s)$ and has the same law as Y_{t-s} .
- Y_t converges in probability to 0 when $t \downarrow 0$.

Two special cases are real Brownian motion (started from 0) and the process $(T_a)_{a \geq 0}$ of hitting times of a real Brownian motion (cf. Exercise 2.26).

Notice that we do not assume that sample paths of Y are càdlàg, but only the weaker regularity assumption (iii). The preceding theory will allow us to find a modification of Y with càdlàg sample paths.

For every $t \geq 0$, we denote the law of Y_t by $Q_t(0, dy)$, and, for every $x \in \mathbb{R}$, we let $Q_t(x, dy)$ be the image of $Q_t(0, dy)$ under the translation $y \mapsto x + y$.

Proposition 6.21 *The collection $(Q_t)_{t \geq 0}$ is a Feller semigroup on \mathbb{R} . Furthermore, $(Y_t)_{t \geq 0}$ is a Markov process with semigroup $(Q_t)_{t \geq 0}$.*

Proof Let us show that $(Q_t)_{t \geq 0}$ is a transition semigroup. Let $\varphi \in B(\mathbb{R})$, $s, t \geq 0$ and $x \in \mathbb{R}$. Property (ii) shows that the law of $(Y_t, Y_{t+s} - Y_t)$ is the product probability measure $Q_t(0, \cdot) \otimes Q_s(0, \cdot)$. Hence,

$$\begin{aligned} \int Q_t(x, dy) \int Q_s(y, dz) \varphi(z) &= \int Q_t(0, dy) \int Q_s(0, dz) \varphi(x + y + z) \\ &= E[\varphi(x + Y_t + (Y_{t+s} - Y_t))] \\ &= E[\varphi(x + Y_{t+s})] \\ &= \int Q_{t+s}(x, dz) \varphi(z) \end{aligned}$$

giving the Chapman–Kolmogorov relation. We should also verify the measurability of the mapping $(t, x) \mapsto Q_t(x, A)$, but this will follow from the stronger continuity properties that we will establish in order to verify the Feller property.

Let us start with the first property of the definition of a Feller semigroup. If $\varphi \in C_0(\mathbb{R})$, the mapping

$$x \mapsto Q_t \varphi(x) = E[\varphi(x + Y_t)]$$

is continuous by dominated convergence, and, again by dominated convergence, we have

$$E[\varphi(x + Y_t)] \xrightarrow{x \rightarrow \infty} 0$$

showing that $Q_t \varphi \in C_0(\mathbb{R})$. Then,

$$Q_t \varphi(x) = E[\varphi(x + Y_t)] \xrightarrow{t \rightarrow 0} \varphi(x)$$

thanks to property (iii). The uniform continuity of φ even shows that the latter convergence is uniform in x . This completes the proof of the first assertion of the proposition. To get the second one, we write, for every $s, t \geq 0$ and every $\varphi \in B(\mathbb{R})$,

$$\begin{aligned} E[\varphi(Y_{s+t}) \mid Y_r, 0 \leq r \leq s] &= E[\varphi(Y_s + (Y_{s+t} - Y_s)) \mid Y_r, 0 \leq r \leq s] \\ &= \int \varphi(Y_s + y) Q_t(0, dy) \\ &= \int \varphi(y) Q_t(Y_s, dy), \end{aligned}$$

using property (ii) and the definition of $Q_t(0, \cdot)$ in the second equality. \square

It then follows from Theorem 6.15 that there exists a modification of $(Y_t)_{t \geq 0}$ with càdlàg sample paths. Obviously this modification still satisfies (i) and (ii).

A *Lévy process* is a process satisfying properties (i) and (ii) above, and having càdlàg sample paths (which implies (iii)). We refer to [3] for a thorough account of the theory of Lévy processes.

6.5.3 Continuous-State Branching Processes

A Markov process $(X_t)_{t \geq 0}$ with values in $E = \mathbb{R}_+$ is called a *continuous-state branching process* if its semigroup $(Q_t)_{t \geq 0}$ satisfies the following property: for every $x, y \in \mathbb{R}_+$ and $t \geq 0$,

$$Q_t(x, \cdot) * Q_t(y, \cdot) = Q_t(x + y, \cdot),$$

where $\mu * \nu$ denotes the convolution of the probability measures μ and ν on \mathbb{R}_+ . Note that this implies $Q_t(0, \cdot) = \delta_0$ for every $t \geq 0$.

Exercise Verify that, if X and X' are two *independent* continuous-state branching processes with the same semigroup $(Q_t)_{t \geq 0}$, then $(X_t + X'_t)_{t \geq 0}$ is also a Markov process with semigroup $(Q_t)_{t \geq 0}$. This is the so-called branching property: compare with discrete time Galton–Watson processes.

Let us fix the semigroup $(Q_t)_{t \geq 0}$ of a continuous-state branching process, and assume that:

- (i) $Q_t(x, \{0\}) < 1$ for every $x > 0$ and $t > 0$;
- (ii) $Q_t(x, \cdot) \rightarrow \delta_x(\cdot)$ when $t \rightarrow 0$, in the sense of weak convergence of probability measures.

Proposition 6.22 *Under the preceding assumptions, the semigroup $(Q_t)_{t \geq 0}$ is Feller. Furthermore, for every $\lambda > 0$, and every $x \geq 0$,*

$$\int Q_t(x, dy) e^{-\lambda y} = e^{-x\psi_t(\lambda)}$$

where the functions $\psi_t : (0, \infty) \rightarrow (0, \infty)$ satisfy $\psi_t \circ \psi_s = \psi_{t+s}$ for every $s, t \geq 0$.

Proof Let us start with the second assertion. If $x, y > 0$, the equality $Q_t(x, \cdot) * Q_t(y, \cdot) = Q_t(x + y, \cdot)$ implies that

$$\left(\int Q_t(x, dz) e^{-\lambda z} \right) \left(\int Q_t(y, dz) e^{-\lambda z} \right) = \int Q_t(x + y, dz) e^{-\lambda z}.$$

Thus the function

$$x \mapsto -\log \left(\int Q_t(x, dz) e^{-\lambda z} \right)$$

is nondecreasing and linear on \mathbb{R}_+ , hence of the form $x\psi_t(\lambda)$ for some constant $\psi_t(\lambda) > 0$ (the case $\psi_t(\lambda) = 0$ is excluded by assumption (i)). To obtain the identity $\psi_t \circ \psi_s = \psi_{t+s}$, we write

$$\begin{aligned} \int Q_{t+s}(x, dz) e^{-\lambda z} &= \int Q_t(x, dy) \int Q_s(y, dz) e^{-\lambda z} \\ &= \int Q_t(x, dy) e^{-y\psi_s(\lambda)} \\ &= e^{-x\psi_t(\psi_s(\lambda))}. \end{aligned}$$

We still have to prove that the semigroup is Feller. For every $\lambda > 0$, set $\varphi_\lambda(x) = e^{-\lambda x}$. Then,

$$Q_t \varphi_\lambda = \varphi_{\psi_t(\lambda)} \in C_0(\mathbb{R}_+).$$

Furthermore, an application of the Stone–Weierstrass theorem shows that the vector space generated by the functions φ_λ , $\lambda > 0$, is dense in $C_0(\mathbb{R}_+)$. It easily follows that $Q_t \varphi \in C_0(\mathbb{R}_+)$ for every $\varphi \in C_0(\mathbb{R}_+)$.

Finally, if $\varphi \in C_0(\mathbb{R}_+)$, for every $x \geq 0$,

$$Q_t \varphi(x) = \int Q_t(x, dy) \varphi(y) \xrightarrow{t \rightarrow 0} \varphi(x)$$

by assumption (ii). Using a remark following the definition of Feller semigroups, this suffices to show that $\|Q_t \varphi - \varphi\| \rightarrow 0$ when $t \rightarrow 0$, which completes the proof. \square

Example For every $t > 0$ and every $x \geq 0$, define $Q_t(x, dy)$ as the law of $\mathbf{e}_1 + \mathbf{e}_2 + \cdots + \mathbf{e}_N$, where $\mathbf{e}_1, \mathbf{e}_2, \dots$ are independent random variables with exponential distribution of parameter $1/t$, and N is Poisson with parameter x/t , and is independent of the sequence (\mathbf{e}_i) . Then a simple calculation shows that

$$\int Q_t(x, dy) e^{-\lambda y} = e^{-x\psi_t(\lambda)}$$

where

$$\psi_t(\lambda) = \frac{\lambda}{1 + \lambda t}.$$

Noting that $\psi_t \circ \psi_s = \psi_{t+s}$, we obtain that $(Q_t)_{t \geq 0}$ satisfies the Chapman–Kolmogorov identity, and then that $(Q_t)_{t \geq 0}$ is the transition semigroup of a continuous-state branching process. Furthermore, $(Q_t)_{t \geq 0}$ satisfies assumptions (i) and (ii) above. In particular, the semigroup $(Q_t)_{t \geq 0}$ is Feller, and one can construct an associated Markov process $(X_t)_{t \geq 0}$ with càdlàg sample paths. One can in fact prove that the sample paths of $(X_t)_{t \geq 0}$ are continuous, and this process is called Feller’s branching diffusion, see Sect. 8.4.3 below.

Exercises

Exercise 6.23 (Reflected Brownian motion) We consider a probability space equipped with a filtration $(\mathcal{F}_t)_{t \in [0, \infty]}$. Let $a \geq 0$ and let $B = (B_t)_{t \geq 0}$ be an (\mathcal{F}_t) -Brownian motion such that $B_0 = a$. For every $t > 0$ and every $z \in \mathbb{R}$, we set

$$p_t(z) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{z^2}{2t}\right).$$

1. We set $X_t = |B_t|$ for every $t \geq 0$. Verify that, for every $s \geq 0$ and $t \geq 0$, for every bounded measurable function $f : \mathbb{R}_+ \rightarrow \mathbb{R}$,

$$E[f(X_{s+t}) | \mathcal{F}_s] = Q_t f(X_s),$$

where $Q_0 f = f$ and, for every $t > 0$, for every $x \geq 0$,

$$Q_t f(x) = \int_0^\infty (p_t(y-x) + p_t(y+x)) f(y) dy.$$

2. Infer that $(Q_t)_{t \geq 0}$ is a transition semigroup, then that $(X_t)_{t \geq 0}$ is a Markov process with values in $E = \mathbb{R}_+$, with respect to the filtration (\mathcal{F}_t) , with semigroup $(Q_t)_{t \geq 0}$.
3. Verify that $(Q_t)_{t \geq 0}$ is a Feller semigroup. We denote its generator by L .
4. Let f be a twice continuously differentiable function on \mathbb{R}_+ , such that f and f'' belong to $C_0(\mathbb{R}_+)$. Show that, if $f'(0) = 0$, f belongs to the domain of L , and $Lf = \frac{1}{2}f''$. (*Hint*: One may observe that the function $g : \mathbb{R} \rightarrow \mathbb{R}$ defined by $g(y) = f(|y|)$ is then twice continuously differentiable on \mathbb{R} .) Show that, conversely, if $f'(0) \neq 0$, f does not belong to the domain of L .

Exercise 6.24 Let $(Q_t)_{t \geq 0}$ be a transition semigroup on a measurable space E . Let π be a measurable mapping from E onto another measurable space F . We assume that, for any measurable subset A of F , for every $x, y \in E$ such that $\pi(x) = \pi(y)$, we have

$$Q_t(x, \pi^{-1}(A)) = Q_t(y, \pi^{-1}(A)),$$

for every $t > 0$. We then set, for every $z \in F$ and every measurable subset A of F , for every $t > 0$,

$$Q'_t(z, A) = Q_t(x, \pi^{-1}(A))$$

where x is an arbitrary point of E such that $\pi(x) = z$. We also set $Q'_0(z, A) = \mathbf{1}_A(z)$. We assume that the mapping $(t, z) \mapsto Q'_t(z, A)$ is measurable on $\mathbb{R}_+ \times F$, for every fixed A .

1. Verify that $(Q'_t)_{t \geq 0}$ forms a transition semigroup on F .
2. Let $(X_t)_{t \geq 0}$ be a Markov process in E with transition semigroup $(Q_t)_{t \geq 0}$ with respect to the filtration $(\mathcal{F}_t)_{t \geq 0}$. Set $Y_t = \pi(X_t)$ for every $t \geq 0$. Verify that $(Y_t)_{t \geq 0}$ is a Markov process in F with transition semigroup $(Q'_t)_{t \geq 0}$ with respect to the filtration $(\mathcal{F}_t)_{t \geq 0}$.
3. Let $(B_t)_{t \geq 0}$ be a d -dimensional Brownian motion, and set $R_t = |B_t|$ for every $t \geq 0$. Verify that $(R_t)_{t \geq 0}$ is a Markov process and give a formula for its transition semigroup. (The case $d = 1$ was treated via a different approach in Exercise 6.23.)

In the remaining exercises, we use the following notation. (E, d) is a locally compact metric space, which is countable at infinity, and $(Q_t)_{t \geq 0}$ is a Feller semigroup on E . We consider an E -valued process $(X_t)_{t \geq 0}$ with càdlàg sample paths, and a collection $(P_x)_{x \in E}$ of probability measures on E , such that, under P_x , $(X_t)_{t \geq 0}$ is a Markov process with semigroup $(Q_t)_{t \geq 0}$, with respect to the filtration (\mathcal{F}_t) , and $P_x(X_0 = x) = 1$. We write L for the generator of the semigroup $(Q_t)_{t \geq 0}$, $D(L)$ for the domain of L and R_λ for the λ -resolvent, for every $\lambda > 0$.

Exercise 6.25 (Scale function) In this exercise, we assume that $E = \mathbb{R}_+$ and that the sample paths of X are continuous. For every $x \in \mathbb{R}_+$, we set

$$T_x := \inf\{t \geq 0 : X_t = x\}$$

and

$$\varphi(x) := P_x(T_0 < \infty).$$

1. Show that, if $0 \leq x \leq y$,

$$\varphi(y) = \varphi(x)P_y(T_x < \infty).$$

2. We assume that $\varphi(x) < 1$ and $P_x(\sup_{t \geq 0} X_t = +\infty) = 1$, for every $x > 0$. Show that, if $0 < x \leq y$,

$$P_x(T_0 < T_y) = \frac{\varphi(x) - \varphi(y)}{1 - \varphi(y)}.$$

Exercise 6.26 (Feynman–Kac formula) Let v be a nonnegative function in $C_0(E)$. For every $x \in E$ and every $t \geq 0$, we set, for every $\varphi \in B(E)$,

$$Q_t^* \varphi(x) = E_x \left[\varphi(X_t) \exp \left(- \int_0^t v(X_s) ds \right) \right].$$

1. Show that, for every $\varphi \in B(E)$ and $s, t \geq 0$, $Q_{t+s}^* \varphi = Q_t^*(Q_s^* \varphi)$.
2. After observing that

$$1 - \exp \left(- \int_0^t v(X_s) ds \right) = \int_0^t v(X_s) \exp \left(- \int_s^t v(X_r) dr \right) ds$$

show that, for every $\varphi \in B(E)$,

$$Q_t \varphi - Q_t^* \varphi = \int_0^t Q_s(v Q_{t-s}^* \varphi) ds.$$

3. Assume that $\varphi \in D(L)$. Show that

$$\frac{d}{dt} Q_t^* \varphi|_{t=0} = L\varphi - v\varphi.$$

Exercise 6.27 (Quasi left-continuity) Throughout the exercise we fix the starting point $x \in E$. For every $t > 0$, we write $X_{t-}(\omega)$ for the left-limit of the sample path $s \mapsto X_s(\omega)$ at t .

Let $(T_n)_{n \geq 1}$ be a strictly increasing sequence of stopping times, and $T = \lim \uparrow T_n$. We assume that there exists a constant $C < \infty$ such that $T \leq C$. The goal of the exercise is to verify that $X_{T-} = X_T, P_x$ a.s.

1. Let $f \in D(L)$ and $h = Lf$. Show that, for every $n \geq 1$,

$$E_x[f(X_T) | \mathcal{F}_{T_n}] = f(X_{T_n}) + E_x \left[\int_{T_n}^T h(X_s) ds \mid \mathcal{F}_{T_n} \right].$$

2. We recall from the theory of discrete time martingales that

$$E_x[f(X_T) | \mathcal{F}_{T_n}] \xrightarrow[n \rightarrow \infty]{\text{a.s., } L^1} E_x[f(X_T) | \widetilde{\mathcal{F}}_T]$$

where

$$\widetilde{\mathcal{F}}_T = \bigvee_{n=1}^{\infty} \mathcal{F}_{T_n}.$$

Infer from question (1) that

$$E_x[f(X_T) \mid \tilde{\mathcal{F}}_T] = f(X_{T-}).$$

3. Show that the conclusion of question (2) remains valid if we only assume that $f \in C_0(E)$, and infer that, for every choice of $f, g \in C_0(E)$,

$$E_x[f(X_T)g(X_{T-})] = E_x[f(X_{T-})g(X_{T-})].$$

Conclude that $X_{T-} = X_T, P_x$ a.s.

Exercise 6.28 (Killing operation) In this exercise, we assume that X has continuous sample paths. Let A be a compact subset of E and

$$T_A = \inf\{t \geq 0 : X_t \in A\}.$$

1. We set, for every $t \geq 0$ and every bounded measurable function φ on E ,

$$Q_t^* \varphi(x) = E_x[\varphi(X_t) \mathbf{1}_{\{t < T_A\}}], \quad \forall x \in E.$$

Verify that $Q_{t+s}^* \varphi = Q_t^*(Q_s^* \varphi)$, for every $s, t > 0$.

2. We set $\bar{E} = (E \setminus A) \cup \{\Delta\}$, where Δ is a point added to $E \setminus A$ as an isolated point. For every bounded measurable function φ on \bar{E} and every $t \geq 0$, we set

$$\bar{Q}_t \varphi(x) = E_x[\varphi(X_t) \mathbf{1}_{\{t < T_A\}}] + P_x[T_A \leq t] \varphi(\Delta), \quad \text{if } x \in E \setminus A$$

and $\bar{Q}_t \varphi(\Delta) = \varphi(\Delta)$. Verify that $(\bar{Q}_t)_{t \geq 0}$ is a transition semigroup on \bar{E} . (The proof of the measurability of the mapping $(t, x) \mapsto \bar{Q}_t \varphi(x)$ will be omitted.)

3. Show that, under the probability measure P_x , the process \bar{X} defined by

$$\bar{X}_t = \begin{cases} X_t & \text{if } t < T_A \\ \Delta & \text{if } t \geq T_A \end{cases}$$

is a Markov process with semigroup $(\bar{Q}_t)_{t \geq 0}$, with respect to the canonical filtration of X .

4. We take it for granted that the semigroup $(\bar{Q}_t)_{t \geq 0}$ is Feller, and we denote its generator by \bar{L} . Let $f \in D(L)$ such that f and Lf vanish on an open set containing A . Write \bar{f} for the restriction of f to $E \setminus A$, and consider \bar{f} as a function on \bar{E} by setting $\bar{f}(\Delta) = 0$. Show that $\bar{f} \in D(\bar{L})$ and $\bar{L}\bar{f}(x) = Lf(x)$ for every $x \in E \setminus A$.

Exercise 6.29 (Dynkin's formula)

1. Let $g \in C_0(E)$ and $x \in E$, and let T be a stopping time. Justify the equality

$$E_x \left[\mathbf{1}_{\{T < \infty\}} e^{-\lambda T} \int_0^\infty e^{-\lambda t} g(X_{T+t}) dt \right] = E_x \left[\mathbf{1}_{\{T < \infty\}} e^{-\lambda T} R_\lambda g(X_T) \right].$$

2. Infer that

$$R_\lambda g(x) = E_x \left[\int_0^T e^{-\lambda t} g(X_t) dt \right] + E_x [\mathbf{1}_{\{T < \infty\}} e^{-\lambda T} R_\lambda g(X_T)].$$

3. Show that, if $f \in D(L)$,

$$f(x) = E_x \left[\int_0^T e^{-\lambda t} (\lambda f - Lf)(X_t) dt \right] + E_x [\mathbf{1}_{\{T < \infty\}} e^{-\lambda T} f(X_T)].$$

4. Assuming that $E_x[T] < \infty$, infer from the previous question that

$$E_x \left[\int_0^T Lf(X_t) dt \right] = E_x[f(X_T)] - f(x).$$

How could this formula have been established more directly?

5. For every $\varepsilon > 0$, we set $T_{\varepsilon,x} = \inf\{t \geq 0 : d(x, X_t) > \varepsilon\}$. Assume that $E_x[T_{\varepsilon,x}] < \infty$, for every sufficiently small ε . Show that (still under the assumption $f \in D(L)$) one has

$$Lf(x) = \lim_{\varepsilon \downarrow 0} \frac{E_x[f(X_{T_{\varepsilon,x}})] - f(x)}{E_x[T_{\varepsilon,x}]}.$$

6. Show that the assumption $E_x[T_{\varepsilon,x}] < \infty$ for every sufficiently small ε holds if the point x is not absorbing, that is, if there exists a $t > 0$ such that $Q_t(x, \{x\}) < 1$. (*Hint:* Observe that there exists a nonnegative function $h \in C_0(E)$ which vanishes on a ball centered at x and is such that $Q_t h(x) > 0$. Infer that one can choose $\alpha > 0$ and $\eta \in (0, 1)$ such that $P_x(T_{\alpha,x} > nt) \leq (1 - \eta)^n$ for every integer $n \geq 1$.)

Notes and Comments

The theory of Markov processes is a very important area of probability theory. Markov processes have a long history that would be too long to summarize here. Dynkin and Feller played a major role in the development of the theory (see in particular Dynkin's books [20, 21]). We limited our treatment to the minimal material needed for our later applications to stochastic differential equations. Our treatment of Feller processes is inspired by the corresponding chapters in [70] and [71]. We chose to focus on Feller semigroups because this special case allows an easy presentation of key notions such as the generator, and at the same time it includes the main examples we consider in this book. The reader interested in the more general theory of Markov processes may have a look at the classical books of Blumenthal and Gettoor [5], Meyer [59] and Sharpe [73]. The idea of characterizing a Markov process by a collection of associated martingales (in the

spirit of Theorem 6.14) has led to the theory of martingale problems, for which we refer the reader to the classical book of Stroock and Varadhan [77]. Martingale problems are also discussed in the book [24] of Ethier and Kurtz, which focuses on problems of characterization and convergence of Markov processes, with many examples and applications. Markov processes with a countable state space are treated, along with other topics, in the more recent book [76] of Stroock. We refer to the monograph [3] of Bertoin for a modern presentation of the theory of Lévy processes.