Chapter IX

Markov Processes

A stochastic process is said to have the Markov property if, at every instant, given the past until that instant, the conditional probability law governing its future depends only on its present state. This property is the probabilistic generalization of the classical notion that, if the present state of a physical system is described in sufficient detail, the system's future evolution would be determined by the present state, without regard to how the system arrived at that state.

The definitions of "time" and "state" depend on the application at hand and on the demands of mathematical tractability. Otherwise, if such practical considerations are ignored, every stochastic process can be made Markovian by enhancing its state space sufficiently.

The theory of Markov processes is the most extensively developed part of probability theory. It covers, in particular, Poisson processes, Brownian motions, and all other Lévy processes. Our aim is to introduce the basic concepts and illustrate them with a few examples and counter-examples. No attempt is made at completeness.

Section 1 is on the Markov property in general. There are examples of Markov chains (discrete-time), of Markov processes (continuous-time), and of anomalous processes lacking the strong Markov property.

Sections 2 and 3 are on two important classes of processes: Itô diffusions and jump-diffusions. They are introduced as solutions to stochastic integral equations. Markov and strong Markov properties are proved directly, generators and resolvents are calculated, and forward and backward equations of Kolmogorov are derived. A quick introduction to stochastic differential equations is given as an appendix in Section 7 for the needs of these sections. These sections can be omitted if the interest is on the general theory.

Markov processes are re-introduced in Section 4 within a modern axiomatic setting. Their Markov property is discussed once more, Blumenthal's zero-one law is proved, the states are classified as holding versus instantaneous, and the behavior at holding states is clarified.

E. Çınlar, *Probability and Stochastics*, Graduate Texts in Mathematics 261, DOI 10.1007/978-0-387-87859-1_9,

Section 5 continues the axiomatic treatment by introducing Hunt processes and Feller processes. The meaning of quasi-left-continuity is explained, the total unpredictability of jump times is given, and the effects of strong Markov property are illustrated.

Section 6 is on resolvents and excessive functions, the connections between them, and their relationships to martingales. It is almost independent of the earlier sections and can be read after Section 2 if desired.

1 Markov Property

Throughout this section, $\mathbb{T}$ is a subset of $\mathbb{R}$; its elements are called times; it will mostly be $\mathbb{R}_+$ and sometimes $\mathbb{N}$. Throughout, $(\Omega,\ \mathcal{H},\ \mathbb{P})$ is a probability space, and $\mathcal{F} = (\mathcal{F}_t)_{t\in\mathbb{T}}$ is a filtration over it.

Let $X = (X_t)_{t\in\mathbb{T}}$ be a stochastic process with some state space $(E,\ \mathcal{E})$ and adapted to the filtration $\mathcal{F}$. We let $\mathcal{G}^o = (\mathcal{G}^o_t)_{t\in\mathbb{T}}$ be the filtration generated by it and put $\mathcal{G}^t_\infty = \sigma\{X_u :\ u \ge t,\ u \in \mathbb{T}\}$, its future after time t.

1.1 Definition. *The process X is said to be Markovian relative to $\mathcal{F}$ if, for every time t, the past $\mathcal{F}_t$ and the future $\mathcal{G}^t_\infty$ are conditionally independent given the present state X_t.*

If X is Markovian relative to $\mathcal{F}$, then it is such relative to $\mathcal{G}^o$ as well, because $\mathcal{G}^o_t \subset \mathcal{F}_t$ by the adaptedness of X to $\mathcal{F}$. It is said to be Markovian, without mentioning a filtration, if it is such relative to $\mathcal{G}^o$.

A similar notion, the strong Markov property, is said to hold if the fixed times t in the preceding definition can be replaced by stopping times. Most Markovian processes are strong Markov, but there are exceptions (see the Examples 1.28 and 1.29).

Characterization

The next theorem uses the definition of conditional independence and properties of repeated conditioning. We use the common shorthand for conditional expectations.

1.2 Theorem. *The following are equivalent:*

a) The process X is Markovian relative to $\mathcal{F}$.

b) For every time t and time $u > t$ and function f in $\mathcal{E}_+$,

1.3 $$\mathbb{E}\,(f\circ X_u\,|\,\mathcal{F}_t) = \mathbb{E}\,(f\circ X_u\,|\,X_t)\,.$$

c) For every time t and positive variable V in $\mathcal{G}^t_\infty$,

1.4 $$\mathbb{E}\,(V|\mathcal{F}_t) = \mathbb{E}\,(V|X_t)\,.$$

d) For every time t and positive variable V in $\mathcal{G}^t_\infty$,

1.5 $$\mathbb{E}\,(V|\mathcal{F}_t) \in \sigma X_t.$$

1.6 Remark. i) The statement (d) is the closest to the intuitive meaning of the Markov property: estimate of a variable determined by the future is a deterministic function of the present state only (regardless of all the past information) – recall that σX_t is the σ-algebra generated by X_t.

ii) The collection of all f for which 1.3 holds is a monotone class. Thus, the theorem remains true, when, in the statement (b), the condition 1.3 holds for every f in $\mathcal{E}_b$ (bounded $\mathcal{E}$-measurable), or every f in $\mathcal{E}_{b+}$, or every indicator $f = 1_A$ with A in $\mathcal{E}$, or every indicator $f = 1_A$ with A in some p-system generating $\mathcal{E}$.

iii) Similarly, by monotone class arguments again, the theorem remains true if 1.4 (or, equivalently, 1.5) is required only for V having the form

1.7 $$V_n = f_1 \circ X_{u_1} \cdots f_n \circ X_{u_n}$$

with some integer $n \geq 1$, some times $t \leq u_1 < u_2 < \ldots < u_n$, and some functions $f_1, \ldots, f_n$ in $\mathcal{E}_+$. Moreover, the functions f_i can be restricted further as in the preceding remark.

Proof. $(a) \Leftrightarrow (c)$ by the definition of conditional independence; $(c) \Rightarrow (b)$ trivially; and we shall show that $(b) \Rightarrow (d) \Rightarrow (c)$. The last implication is easy: assuming 1.5,

$$\mathbb{E}\,(V|\mathcal{F}_t) = \mathbb{E}\,(\mathbb{E}\,(V|\mathcal{F}_t)\,|X_t) = \mathbb{E}\,(V|X_t)\,.$$

To prove that $(b) \Rightarrow (d)$, assume (b). By Remark 1.6 iii, it is enough to show 1.5 for V having the form 1.7. We do this by induction. For $n = 1$, we have 1.5 from (b). Assume that 1.5 holds for every V_n having the form 1.7. Note that

$$\mathbb{E}\,(V_{n+1}|\mathcal{F}_{u_n}) = V_n\,\mathbb{E}\left(f_{n+1} \circ X_{u_{n+1}}|\mathcal{F}_{u_n}\right) = V_n \cdot g \circ X_{u_n} = \hat{V}_n$$

for some g in $\mathcal{E}_+$ in view of (b). Since $\hat{V}_n$ has the form 1.7 with gf_n replacing f_n, the induction hypothesis applies to $\hat{V}_n$ to yield

$$\mathbb{E}\,(V_{n+1}|\mathcal{F}_t) = \mathbb{E}\left(\hat{V}_n|\mathcal{F}_t\right) \in \sigma X_t.$$

Thus, 1.5 holds for V_{n+1} as well. □

Transition functions

Recall that a Markov kernel on $(E, \mathcal{E})$ is a transition probability kernel from $(E, \mathcal{E})$ into $(E, \mathcal{E})$; see I.6.5 *et seq.* Let $(P_{t,u})$ be a family of Markov kernels on $(E, \mathcal{E})$ indexed by pairs of times $t \le u$. It is said to be a *Markovian transition function* on $(E, \mathcal{E})$ if

1.8 $$P_{s,t}\, P_{t,u} = P_{s,u}, \qquad 0 \le s < t \le u.$$

The preceding is called the *Chapman-Kolmogorov equation.*

The Markovian process X is said to *admit* $(P_{t,u})$ as a transition function if

1.9 $$\mathbb{E}\,(f \circ X_u | X_t) = (P_{t,u} f) \circ X_t, \qquad t < u, f \in \mathcal{E}_+.$$

Obviously, it is sufficient to check 1.9 for f that are indicators. This provides the intuitive meaning for the kernels:

1.10 $$P_{t,u}\,(x, A) = \mathbb{P}\,\{X_u \in A | X_t = x\}\,.$$

1.11 Remark. There are Markov processes that have no transition functions. Here is an example. Suppose that $\mathbb{T} = \mathbb{R}_+$, and $E = \mathbb{R}_+ \times \Omega$, and $\mathcal{E}$ consists of subsets A of E such that the section $\{\omega \in \Omega : (t, \omega) \in A\}$ belongs to $\mathcal{F}_t$ for every t. Suppose, further, that $X_t(\omega) = (t, \omega)$ for t in $\mathbb{R}_+$ and ω in Ω. Then, $X = (X_t)_{t\in\mathbb{R}_+}$ is a stochastic process with state space $(E, \mathcal{E})$ and is adapted to $\mathcal{F}$. Note that the σ-algebra generated by X_t is exactly $\mathcal{F}_t$, and, hence, the condition 1.3 holds automatically. This Markovian process has no transition function. It is also devoid of interest, since there is nothing further to be said about it.

1.12 Remark. The preceding example illustrates that every process can be made Markovian, but at the cost of mathematical tractability. Begin with a process X^0 with some state space $(D, \mathcal{D})$. Let $\mathcal{F}$ be the filtration generated by it. Define $(E, \mathcal{E})$ and X as in the preceding remark. Now, the "state" of X at time t is the whole history of X^0 until t. By this device, X^0 is converted to the Markovian process X.

Time-homogeneity

Suppose that X is Markovian and admits $(P_{t,u})$ as its transition function. It is said to be *time-homogeneous* if, for every time t and time $u > t$, the dependence of $P_{t,u}$ on the pair (t,u) is through $u - t$ only, that is, if

1.13 $$P_{t,u} = P_{u-t}$$

for some Markov kernel P_{u-t}. Theoretically, there is no loss of generality in assuming time-homogeneity: if X is not, then it can be studied through $\hat{X} = (t, X_t)$, and $\hat{X}$ is Markovian and time-homogeneous. Note that this trick makes time a part of the state description. See Exercise 1.40.

Chains and Processes

Suppose that X is Markovian and time-homogeneous. We call it a *Markov chain* if $\mathbb{T} = \mathbb{N}$, and *Markov process* if $\mathbb{T} = \mathbb{R}_+$.

Suppose that X is a Markov chain. Then, $Q = P_{t,t+1}$ is free of t, and the Chapman-Kolmogorov equation 1.8 yields

1.14 $$P_{t,u} = Q^n, \qquad t \in \mathbb{N},\ u - t = n \in \mathbb{N}.$$

This is expressed by saying that X is a Markov chain with *transition kernel* Q.

Suppose that X is a Markov process. Then, the Markov kernels $P_t,\ t \in \mathbb{R}_+$, must satisfy the semigroup property

1.15 $$P_t\ P_u = P_{t+u}, \qquad t, u \in \mathbb{R}_+,$$

this being the Chapman-Kolmogorov equation in view of 1.13. Then, it is usual to call (P_t) a *transition semigroup* and to say that X is a Markov process with *transition function* (P_t).

For a chain, since the time-set has only one limit point, the analysis required is more straight forward and has more to do with limits in distribution of X_n as $n \to \infty$. For a process, the mathematical treatment has greater ties to classical analysis and semigroups and partial differential equations. We shall concentrate on processes; an analogous program for chains can be carried out without difficulty. However, as a way of displaying the Markov property in its most direct form, we give examples of chains next.

Markov chains

Every Markov chain encountered in applications is constructed from a sequence of independent and identically distributed random variables through a deterministic transformation. In fact, if the state space $(E, \mathcal{E})$ is standard, we may construct every Markov chain in this fashion; see Exercise 1.38 for an illustration with $E = \mathbb{R}$. Interestingly, this form shows that every Markov chain (and, by extension, every Markov process) is a Lévy chain (or Lévy process) in an abstract sense.

Let $(E, \mathcal{E})$ and $(D, \mathcal{D})$ be measurable spaces. Let $\varphi :\ E \times D \mapsto E$ be measurable with respect to $\mathcal{E} \otimes \mathcal{D}$ and $\mathcal{E}$. Let X_0 be a random variable taking values in $(E, \mathcal{E})$ and, independent of it, let $(Z_n)_{n\in\mathbb{N}}$ be an independency of identically distributed variables taking values in $(D, \mathcal{D})$. Define

1.16 $$X_{n+1} = \varphi\left(X_n, Z_{n+1}\right), \qquad n \in \mathbb{N}.$$

Together with X_0, this defines a Markov chain $X = (X_n)_{n\in\mathbb{N}}$ with state space $(E, \mathcal{E})$ and transition kernel

1.17 $$Q(x, A) = \mathbb{P}\left\{\varphi(x, Z_0) \in A\right\}, \qquad x \in E,\ A \in \mathcal{E}.$$

The formula 1.16 encapsulates the essence of Markov chains: the next state X_{n+1} is a deterministic function of the present state X_n and the next random influence Z_{n+1}. The deterministic function φ remains the same over all time;

this is time-homogeneity. In this context, φ is called the structure function and the Z_n are the driving variables. Here are some examples and implications of this construction.

1.18 *Random walks.* Suppose that $E = D = \mathbb{R}^d$ with the attendant Borel σ-algebras $\mathcal{E}$ and $\mathcal{D}$. Take $\varphi(x,\ z) = x + z$. The resulting Markov chain is called a random walk on $\mathbb{R}^d$.

1.19 *Gauss-Markov chains.* Let $E = D = \mathbb{R}^d$ again. Suppose that the Z_n have the d-dimensional standard Gaussian distribution. Take

$$\varphi(x,z) = Ax + Bz, \qquad x,z \in \mathbb{R}^d,$$

where A and B are some $d \times d$ matrices. The resulting chain X is called a Gauss–Markov chain. If X_0 is fixed or has some Gaussian distribution, then the chain (X_n) is Gaussian; this can be seen by noting that

$$X_n = A^n X_o + A^{n-1} B\ Z_1 + \cdots + AB\ Z_{n-1} + B\ Z_n.$$

1.20 *Products of random matrices.* Suppose that $E = D = \mathbb{R}^{d\times d}$, the space of $d \times d$ matrices with real entries. Then, the Z_n are independent and identically distributed random matrices. Take $\varphi(x,z) = zx$, the matrix x multiplied on the left by the matrix z. The resulting chain is given by $X_n = Z_n \cdots Z_1 X_0$; it is a "left random walk" on the set of $d \times d$ matrices. Similarly, taking $\varphi(x,z) = xz$ yields a "right random walk".

1.21 *Continuation.* Suppose that $E = \mathbb{R}^d$ and $D = \mathbb{R}^{d\times d}$; and take $\varphi(x,\ z) = zx$, the product of the matrix z and the column vector x, Then, the chain (X_n) becomes the orbit of the random point X_0 under successive applications of the random linear transformations represented by the matrices $Z_1,\ Z_2, \ldots$.

1.22 *Random dynamical systems.* This is to give a different interpretation to 1.16. Leave $(E, \mathcal{E})$ arbitrary. Define, for each n, a random transformation Φ_n from E into E by letting

$$\Phi_n^\omega(x) = \varphi\left(x, Z_n(\omega)\right), \qquad \omega \in \Omega, x \in E.$$

Then, $\Phi_1, \Phi_2, \ldots$ are independent and identically distributed random transformations from $(E, \mathcal{E})$ into $(E, \mathcal{E})$, and $X_{n+1} = \Phi_{n+1}\ (X_n)$. So, the chain X is obtained by successive applications of independent and identically distributed random transformations.

1.23 *Continuation.* Let $\varphi_{m,n}^\omega$ be the composition of the transformations $\Phi_{m+1}^\omega, \ldots, \Phi_n^\omega$, that is, define

$$\varphi_{m,n}^\omega = \begin{cases} \text{identity} & \text{if } m = n, \\ (\Phi_n^\omega) \circ \cdots \circ \left(\Phi_{m+1}^\omega\right) & \text{if } m < n, \end{cases}$$

for $0 \le m \le n < \infty$. For each ω, the family $\left\{\varphi_{m,n}^\omega : 0 \le m \le n < \infty\right\}$ is a flow, that is, the flow equation

$$\varphi_{m,n}^\omega \left(\varphi_{k,m}^\omega(x)\right) = \varphi_{k,n}^\omega(x), \qquad 0 \le k \le m \le n,$$

is satisfied. And, the Markov chain X is the path of X_0 under the action of the random flow $\varphi = (\varphi_{m,n})_{0\le m\le n<\infty}$.

Regarding $\varphi_{m,n}$ as the "increment" of φ over the interval $(m,\ n]$, we see that φ has stationary and independent increments. Thus, the Markov chain X is, in this abstract sense, a discrete-time Lévy process in the space of transformations.

Examples of Markov processes

Brownian motion is a Markov process. A number of Markov processes related to it were given in Chapter VIII on Brownian motion; see Examples VIII.1.19, VIII.1.21, VIII.1.22. The following examples are to forge some connections, and also give some pathological (see 1.28) and fascinating (see 1.29) cases where the strong Markov property fails.

1.24 *Lévy processes.* Suppose that $E = \mathbb{R}^d$ and $\mathcal{E} = \mathcal{B}(\mathbb{R}^d)$, and assume that $X_t = X_0 + Y_t,\ t \in \mathbb{R}_+$, where $Y = (Y_t)$ is a Lévy process independent of X_0. Let π_t be the distribution of Y_t, and recall that $A - x = \{y - x :\ y \in A\}$. Then, X is a Markov process with transition function

1.25 $$P_t(x, A) = \pi_t(A - x), \qquad x \in E,\ A \in \mathcal{E},\ t \in \mathbb{R}_+.$$

In other words, X is both time-homogeneous and spatially homogeneous. Conversely, if X is such, that is, if X is a Markov process whose transition semigroup (P_t) has the form 1.25, then $X = X_0 + Y$ for some Lévy process Y.

1.26 *Markov chains subordinated to Poisson.* Let $(Y_n)_{n\in\mathbb{N}}$ be a Markov chain with state space $(E,\ \mathcal{E})$ and transition kernel Q. Let (N_t) be a Poisson process, with rate c, independent of the chain (Y_n). Suppose that

$$X_t = Y_{N_t}, \qquad t \in \mathbb{R}_+.$$

Then, X is a Markov process with state space (E, $\mathcal{E}$) and transition function (P_t), where

1.27 $$P_t(x, A) = \sum_{n=0}^{\infty} \frac{e^{-ct}(ct)^n}{n!} Q^n(x, A).$$

1.28 *Delayed uniform motion.* The state space is $E = \mathbb{R}_+$. The process depicts the motion of a particle that is at the origin initially, stays there an exponentially distributed amount T of time, and then moves upward at unit speed:

$$X_t = (t - T)^+, \qquad t \in \mathbb{R}_+.$$

This X is a Markov process. Its transition function (P_t) is easy to compute by using the working formula

$$P_t f(x) = \mathbb{E}\,(f \circ X_{s+t} | X_s = x) :$$

If $x > 0$, then $X_{s+t} = x + t$. If $x = 0$, the particle's sojourn at 0 has not ended yet, that is, $T > s$. By the exponential nature of T, then, the remaining sojourn time $T - s$ has the same exponential distribution as T itself. Letting c be the parameter of that exponential distribution, we get

$$P_t f(x) = \begin{cases} f(x+t) & if\ x > 0, \\ e^{-ct} f(0) + \int_0^t du\ ce^{-cu}\ f(t-u) & if\ x = 0. \end{cases}$$

Suppose now that the filtration $\mathcal{F}$ is taken to be $\left(\mathcal{G}^o_{t+}\right)$. Then, T is a stopping time of $\mathcal{F}$, and $X_T = X_0 = 0$. If X were strong Markov, the future after T would have the same law as the future at $t = 0$. But it is not so; future at $t = 0$ starts with a sojourn of some length at 0, whereas the future at T is that of immediate motion. So, this process is *not* strong Markov.

Intuitive notion of the Markov property is that the present state determines the law of the future; and this is tacitly extended to cases where "the present time" is allowed to be a stopping time. The present example is cautionary. At the same time, it displays the reason for the failure of the strong Markov property: the state 0 is allowed to play two different roles: as a point of sojourn, and as a launching pad for the motion. If we re-define the process as

$$\hat{X}_t(\omega) = \begin{cases} -1 & \text{if } t < T(\omega), \\ t - T(\omega) & \text{if } t \geq T(\omega), \end{cases}$$

then we have a strong Markov process $\hat{X}$ with state space $\{-1\} \cup \mathbb{R}_+$.

1.29 *Lévy's increasing continuous process.* This is an example, due to Lévy, of another process whose Markov property does not extend to stopping times. Moreover, it illustrates the importance of choosing the correct state space and the correct construction of the process. As a by-product, it shows the advantages of concentrating on the dynamics of the random motion, instead of the analytic machinery of transition functions and the like.

The canonical process has state space $\mathbb{R}_+$. Started at 0, its paths are increasing continuous with limit $+\infty$ as time goes to $+\infty$. Every rational number in $\mathbb{R}_+$ is a holding point, that is, the process has an exponential sojourn there before resuming its upward creep. The process spends no time in the set of irrationals. By re-labeling the states, we shall get a bizarre Markov process with state space $\bar{\mathbb{N}} = \{0, 1, \ldots, +\infty\}$.

Let $\mathbb{Q}_+$ denote the set of rational numbers in $\mathbb{R}_+$. Over the probability space $(\Omega, \mathcal{H}, \mathbb{P})$, we suppose that $\{Z_q : q \in \mathbb{Q}_+\}$ is an independency of $\mathbb{R}_+$-valued exponential random variables with Z_q having the mean $m(q)$, where

$$\sum_{q \in \mathbb{Q}_+ \cap [0,1)} m(q) = 1,$$

and, for each integer $n \geq 1$, we have $m(q) = m(q - n)$ for q in $[n, n+1)$. We are thinking of a particle that moves upward in $\mathbb{R}_+$, having a sojourn of length Z_q at each rational q, and spending no time elsewhere. Thus, the cumulative time it spends in the set $[0,x]$ is

1.30 $$S_x = \sum_{q \in \mathbb{Q}_+ \cap [0,\, x]} Z_q, \qquad x \in \mathbb{R}_+,$$

and, therefore, the particle's position at time t is

1.31 $$X_t = \inf \{x \in \mathbb{R}_+ : S_x > t\}, \qquad t \in \mathbb{R}_+.$$

In view of the way the means $m(q)$ are chosen, for almost every ω, we have $S_x(\omega) < \infty$ for all x in $\mathbb{R}_+$, but with limit $+\infty$ as $x \to \infty$. Clearly, $x \mapsto S_x(\omega)$ is right-continuous and strictly increasing, which implies that $t \mapsto X_t(\omega)$ is continuous increasing. Moreover, since the path $S(\omega)$ is of the pure-jump type,

1.32 $$\text{Leb}\,\{t \in \mathbb{R}_+ : X_t(\omega) \notin \mathbb{Q}_+\} = 0.$$

The process $(S_x)_{x \in \mathbb{R}_+}$ is a pure-jump process with independent (but non-stationary) increments; see Exercise VI.4.24. It jumps at every rational q by the exponential amount Z_q. Thus, $X = (X_t)_{t \in \mathbb{R}_+}$ is a Markov process (time-homogeneous) with a transition function (P_t) that can be specified, see Exercise 1.39; it is Markov relative to $\left(\mathcal{G}^o_{t+}\right)$ as well.

Heuristically, given that $X_t = x$ and $x \in \mathbb{Q}_+$, then, the particle will stay at x a further amount of time that is exponential with mean $m(x)$ and then start its upward motion. This is for fixed time t. But, when t is replaced by the random time T at which the particle departs the fixed rational point x, the future looks different. Thus, X lacks the strong Markov property.

Next, we take advantage of 1.32 to define a Markov process with a discrete state space. Let b be a bijection from $\mathbb{Q}_+$ onto $\mathbb{N}$; this is just a re-labeling of the rationals by integers. Define

1.33 $$Y_t(\omega) = \begin{cases} b \circ X_t(\omega) & \textit{if } X_t(\omega) \in \mathbb{Q}_+, \\ +\infty & \text{otherwise.} \end{cases}$$

Then, $Y = (Y_t)$ is a Markov process with state space $\bar{\mathbb{N}}$. Its paths are difficult to describe directly: if $Y_t = i$, then the particle stays there an exponential time, but there is no "next" integer state to go. The state $+\infty$ is "fictitious"; the total amount of time spent there by Y is zero by 1.32. The paths have discontinuities of the second kind. We may define

$$\mathrm{Q}_t(i, A) = \mathbb{P}\,\{Y_{s+t} \in A | Y_s = i\}, \qquad t \in \mathbb{R}_+, i \in \mathbb{N},\ A \subset \mathbb{N},$$

to obtain a Markov transition semigroup, that is, $Q_t Q_u = Q_{t+u}$ and $Q_t(i, \mathbb{N}) = 1$ for each i. For this reason, Y is said to have $\mathbb{N}$ as its *minimal state space*. This process is a good example of inadequacy of transition functions (and generators to come) as the base to build the theory on. Despite this sentiment, we continue with ...

Probability Laws

We return to the general case with index set $\mathbb{T}$ and suppose that X is Markovian and admits $(P_{t,u})$ as its transition function. Suppose, further, that $\mathbb{T} \subset \mathbb{R}_+$ and $0 \in \mathbb{T}$. Let μ_0 be the distribution of X_0. Then, for times $0 = t_0 < t_1 < \cdots < t_n$,

$$\mathbb{P}\{X_{t_0} \in dx_0, X_{t_1} \in dx_1, X_{t_2} \in dx_2, \ldots, X_{t_n} \in dx_n\}$$

1.34 $$= \mu_{t_0}(dx_0) P_{t_0,t_1}(x_0, dx_1) P_{t_1,t_2}(x_1, dx_2) \cdots P_{t_{n-1},t_n}(x_{n-1}, dx_n)\,.$$

This follows from repeated applications of the Markov property. It shows, as well, that the probability law of X is determined by the initial distribution μ_0 and the transition function $(P_{t,u})$. Modern theory treats $(P_{t,u})$ as fixed, but μ_0 as a variable; it is usual to write $\mathbb{P}^\mu$ for $\mathbb{P}$ when $\mu_0 = \mu$, and $\mathbb{P}^x$ when $\mu_0 = \delta_x$, Dirac at x.

Existence and construction

Let μ be a probability measure and $(P_{t,u})$ a Markov transition function, both on some measurable space $(E, \mathcal{E})$. If $\mathbb{T} = \mathbb{N}$, then Theorem IV.4.7 shows the existence of a probability space $(\Omega, \mathcal{H}, \mathbb{P}^\mu)$ and a process $X = (X_t)_{t\in\mathbb{T}}$ such that X is Markovian with initial distribution μ and transition function $(P_{t,u})$, that is, such that 1.34 holds. If $\mathbb{T} = \mathbb{R}_+$, the same existence result follows from the Kolmogorov extension theorem, IV.4.18, under a slight condition on $(E, \mathcal{E})$. We refer to Chapter IV, Sections 4 and 5, for the details as well as for a discussion of some special cases and issues regarding "time" and "space".

In practice, however, one rarely has $(P_{t,u})$ specified from the start. Instead, X is constructed from well-known objects, and $(P_{t,u})$ is defined implicitly from X. For instance, the example 1.28 is constructed from one exponential variable, and the example 1.29 from a countable independency of exponentials. As we know, Wiener processes and Poisson random measures on $\mathbb{R}^2$ can be constructed from a countable independency of uniform variables, and Lévy processes are constructed from Wiener processes and Poisson random measures. Similarly, most Markov processes are constructed from a countable independency of uniform variables via Wiener processes and Poisson random measures; the constructions are sometimes direct, and often by means of stochastic integral equations. Sections 2 and 3 illustrate the method.

Exercises and complements

1.35 *Processes with discrete state spaces.* Suppose that X is a Markov process (time-set $\mathbb{R}_+$, time-homogeneous) with state space $(E, \mathcal{E})$ and transition function (P_t). Suppose that $(E, \mathcal{E})$ is discrete, that is, E is countable and $\mathcal{E} = 2^E$, the discrete σ-algebra on E. Then, each P_t has the form

$$P_t(x, A) = \sum_{y \in A} p_t(x, y), \qquad x \in E,\ A \in \mathcal{E},$$

and we may regard $y \mapsto p_t(x, y)$ as the density of the measure $A \mapsto P_t(x, A)$ with respect to the counting measure on $(E, \mathcal{E})$. Of course, then, we may identify the kernel P_t with the matrix whose entries are the probabilities $p_t(x, y)$. We shall do this without further comment.

1.36 *Continuation.* Suppose that E consists of two elements, a and b. Let

$$P_t = \begin{bmatrix} q + pe^{-ct} & p - pe^{-ct} \\ q - qe^{-ct} & p + qe^{-ct} \end{bmatrix}, \qquad t \ge 0,$$

where p and q are numbers in [0,1] with $p + q = 1$, and c is a number in $\mathbb{R}_+$. Show that the matrices P_t satisfy $P_0 = I$ and $P_t P_u = P_{t+u}$. When E consists of two states, this is the most general form of a transition function (P_t). The case $c = 0$ is degenerate (what happens then?). Describe the paths in the cases $p = 0$ or $q = 0$.

1.37 *Subordination of Markov to Lévy.* Let X be a Markov process with state space $(E, \mathcal{E})$ and transition function (P_t). Let $S = (S_t)$ be an increasing Lévy process independent of X, and with distribution π_t for S_t. Define

$$\hat{X}_t = X_{S_t}, \qquad t \in \mathbb{R}_+.$$

Show that $\hat{X}$ is again a Markov process with state space $(E, \mathcal{E})$. Compute its transition function $(\hat{P}_t)$ in terms of (P_t) and (π_t).

1.38 *Markov chains.* Let Q be a Markov kernel on $(\mathbb{R}, \mathcal{B}_\mathbb{R})$. For each x in $\mathbb{R}$, define

$$\varphi(x, u) = \inf \{ y \in \mathbb{R} : Q(x, (-\infty, y]) > u \}, \qquad u \in (0, 1).$$

Then, $u \mapsto \varphi(x, u)$ is increasing and right-continuous.

a) Show that $x \mapsto \varphi(x, u)$ is Borel measurable for each u. Conclude that φ is a Borel function on $\mathbb{R} \times (0, 1)$.

b) Let (Z_n) be an independency of uniform variables taking values in (0,1). Suppose X_0 is independent of (Z_n), and define (X_n) by 1.16. Show that (X_n) is a Markov chain with transition kernel Q.

1.39 *Lévy's example.* Let X be as in Example 1.29. Let (P_t) be its transition function. Show that, for real numbers $0 \le x \le y$,

$$P_t(x, (y, \infty)) = \mathbb{P} \left\{ \sum_{x \le q \le y} Z_q < t \right\}$$

where the sum is over the rationals q in the interval $[x,y]$. Show that

$$\int_0^\infty dt\ e^{-pt}P_t\left(x,(y,\infty)\right) = \frac{1}{p}\prod_{x\le q\le y}\frac{1}{1+m(q)p}, \quad p\in\mathbb{R}_+.$$

This specifies (P_t), at least in principle.

1.40 *Time-homogeneity.* Suppose that X is Markovian with state space $(E,\mathcal{E})$ and admits $(P_{t,u})$ as its transition function (we do not assume time-homogeneity). Define

$$\hat{X}_t = (t, X_t)\,, \qquad t\in\mathbb{R}_+.$$

Then, $\hat{X}$ is Markovian with state space $(\hat{E},\hat{\mathcal{E}}) = (\mathbb{R}_+\times E,\ \mathcal{B}_{\mathbb{R}_+}\otimes\mathcal{E})$. Show that it is time-homogeneous. Show that its transition function $(\hat{P}_t)$ is given by, for positive f in $\hat{\mathcal{E}}$,

$$\hat{P}_t f\left(\hat{x}\right) = \int_E P_{s,s+t}(x,dy)f(s+t,y), \qquad \hat{x}=(s,x)\in\hat{E}.$$

1.41 *Processes with independent increments.* Let X be a process in $\mathbb{R}^d$ having independent increments, but without stationarity of increments. Then, X is Markovian, but without time-homogeneity; we have

$$P_{t,u}(x,A) = \mathbb{P}\left\{X_u\in A|X_t=x\right\} = \mathbb{P}\left\{X_u - X_t\in A-x\right\} = \pi_{t,u}\left(A-x\right).$$

Define $\hat{X}_t = (t,\ X_t)$ as in the preceding example. Then, $\hat{X}$ is a time-homogeneous Markov process. Compute its transition function $(\hat{P}_t)$ in terms of $\pi_{s,t}$. Note that $\hat{X}$ has independent increments, but still without the stationarity of increments.

1.42 *Expanded filtrations.* Suppose that X is Markovian relative to the filtration $\mathcal{F}$. Let $\mathcal{H}_0$ be a sub-σ-algebra of $\mathcal{H}$ that is independent of $\mathcal{F}_\infty$. Put $\hat{\mathcal{F}}_t = \mathcal{H}_0\vee\mathcal{F}_t$ for each time t. Then, X is Markovian relative to $\hat{\mathcal{F}}$ as well. Show. This is helpful when X is being studied in the presence of other processes that are independent of X.

1.43 *Entrance laws.* Suppose that X is Markovian with time-set $\mathbb{T}$ and transition function $(P_{t,u})$. Suppose that $\mathbb{T}$ does not have an initial element; $\mathbb{T}=(0,\infty)$ or $\mathbb{T}=(-\infty,+\infty)$ for instance. Let μ_t be the distribution of X_t. Then, the formula 1.34 holds for times $t_0<t_1<\ldots<t_n$ in $\mathbb{T}$. Note that, necessarily,

1.44 $$\int_E \mu_t(dx)P_{t,u}\left(x,A\right) = \mu_u\left(A\right), \qquad t<u, A\in\mathcal{E}.$$

In general, if a family (μ_t) of probability measures satisfies 1.44 for some transition function $(P_{t,u})$, then (μ_t) is said to be an *entrance law* for $(P_{t,u})$. If $\mathbb{T}$ has an initial element t_0, then $\mu_t = \mu_{t_0}\ P_{t_0,t},\ t\ge t_0$, defines an entrance law (μ_t) from the initial law μ_{t_0}.

2 Itô Diffusions

Itô diffusions are continuous strong Markov processes satisfying certain stochastic differential equations. They are generalizations of Brownian motions in the following way.

Over some probability space, let X be a Brownian motion on $\mathbb{R}$. It has the form $X_t = X_0 + at + b\,W_t$, where W is a Wiener process, and a and b constants. The dynamics of the motion is expressed better in the classical fashion:

$$dX_t = a\;dt + b\;dW_t,$$

that is, velocity is equal to a constant a perturbed by some "noise." We notice that X will remain Markovian in the more general case where a is replaced with $a(X_t)$, some function of the current position X_t, and the noise multiplier b is replaced with $b(X_t)$, some function of X_t. The result is

2.1 $$dX_t = a \circ X_t\;dt + b \circ X_t\;dW_t$$

or, equivalently, in the formal language of integrals,

2.2 $$X_t = X_0 + \int_0^t a \circ X_s\;ds + \int_0^t b \circ X_s\;dW_s.$$

But there arises a problem: the second integral does not have a conventional meaning, because the paths $t \mapsto W_t$ have infinite total variation over every interval $[s,\;s+u]$ with $u > 0$. Fortunately, it is possible to give a meaning to such integrals, called stochastic integrals of Itô to distinguish them from the ordinary ones.

This section can be read without previous exposure to stochastic calculus if one is willing to take some results on faith. Nevertheless, we put a summary of stochastic integration, as an appendix, in Section 7.

Stochastic base

The motion of interest will be a continuous process with state space $(E,\;\mathcal{E})$, where $E = \mathbb{R}^d$ for some fixed dimension $d \geq 1$ and $\mathcal{E} = \mathcal{B}(\mathbb{R}^d)$. The process will be the solution of a stochastic differential equation driven by a multi-dimensional Wiener process.

Throughout this section, $(\Omega, \mathcal{H}, \mathbb{P})$ is a complete probability space, $\mathcal{F} = (\mathcal{F}_t)_{t\in\mathbb{R}_+}$ is an augmented right-continuous filtration, and $W = (W^1, \ldots, W^m)$ is an m-dimensional Wiener process adapted to $\mathcal{F}$; the integer m will remain fixed. In addition, X_0 is an E-valued random variable in $\mathcal{F}_0$ and is, thus, independent of W. We let $(x, H) \mapsto \mathbb{P}^x(H)$ be a regular version of the conditional probabilities

2.3 $$\mathbb{P}^x(H) = \mathbb{P}\left(H | X_0 = x\right).$$

Equation of motion

The deterministic data are some vector fields $u_0, \dots, u_m$ on $E = \mathbb{R}^d$, that is, each u_n is a mapping from E into E. Throughout, we assume that the following condition of *Lipschitz continuity* holds; here $|x|$ is the length of the vector x for each x in E.

2.4 CONDITION. There is a constant c in $\mathbb{R}_+$ such that

$$|u_n(x) - u_n(y)| \le c\,|x - y|\,, \qquad x, y \in E,\ 0 \le n \le m. \qquad \square$$

This condition ensures that the following equation of motion makes sense and has a unique solution (see Theorem 2.13 below):

2.5
$$X_t = X_0 + \int_0^t u_0 \circ X_s\; ds + \sum_{n=1}^m \int_0^t u_n \circ X_s \quad dW_s^n;$$

here, the integrals involving the W^n are to be understood as Itô integrals of stochastic calculus (see Section 7, Appendix).

Somewhat more explicitly, writing X_t^i for the i-component of X_t, and $u_n^i(x)$ for the i-component of the vector $u_n(x)$, the stochastic integral equation 2.5 becomes

2.6
$$X_t^i = X_0^i + \int_0^t u_0^i \circ X_s\; ds + \sum_{n=1}^m \int_0^t u_n^i \circ X_s\; dW_s^n, \quad 1 \le i \le d.$$

Again equivalently, 2.5 can be written as a stochastic differential equation:

2.7
$$dX_t = u_0 \circ X_t\; dt + \sum_{n=1}^m u_n \circ X_t\; dW_t^n.$$

The looks of the preceding line can be simplified: put $a(x) = u_0(x)$ and let $b(x)$ be the $d \times m$ matrix whose (i,n)-entry is $u_n^i(x)$; then 2.7 becomes

2.8
$$dX_t = a \circ X_t\; dt + b \circ X_t\; dW_t,$$

which looks exactly like 2.1, and 2.5 gets to look like 2.2. But, 2.5 and 2.7 are better at conveying the role of each W^n: the effect of W^n is carried to the motion by the vector field u_n; this issue becomes important when we consider a cloud of particles whose motions satisfy the same differential equation 2.7.

Examples

2.9 *Geometric Brownian motion.* With $d = m = 1$, and b and c constants in $\mathbb{R}$, consider the geometric Brownian motion

$$X_t = X_0 \exp\left(bW_t + ct\right), \qquad t \in \mathbb{R}_+.$$

Using Itô's formula (Theorem 7.20), we see that X is the solution to

$$dX_t = aX_t\ dt + bX_t\ dW_t,$$

where $a = c + \frac{1}{2}\,b^2$. In particular, when $a = 0$, we obtain the exponential martingale $X_t = X_o \exp(bW_t - \frac{1}{2}b^2 t)$ as the solution to $dX_t = bX_t\ dW_t$.

2.10 *Ornstein–Uhlenbeck process.* In 2.8, suppose that $a(x) = Ax$ and $b(x) = B$, where A and B are matrices of dimensions $d \times d$ and $d \times m$ respectively; we get

$$dX_t = AX_t\ dt + B\ dW_t,$$

which is also called the *Langevin equation.* The solution is

$$X_t = e^{tA}X_0 + \int_0^t e^{(t-s)A}B\ dW_s,$$

where $e^{tA} = \sum_{k=0}^{\infty} \left(t^k/k!\right) A^k$. In the particular case where $d = m = 1$, the matrices reduce to real numbers; and assuming that A is a negative constant, say $A = -c$ and $B = b$, we obtain

$$X_t = e^{-ct}X_0 + b\int_0^t e^{-c(t-s)}\ dW_s, \quad t \in \mathbb{R}_+.$$

This is the one-dimensional velocity process in the model of Ornstein and Uhlenbeck for the physical Brownian motion; see Exercise 2.60 also.

2.11 *Brownian motion on the unit circle.* This is the motion X, on the unit circle in $\mathbb{R}^2$, whose components are

$$X_t^1 = \cos W_t, \quad X_t^2 = \sin W_t,$$

where W is a Wiener process; one can think of it as the complex-valued motion $\exp(iW_t)$. Using Itô's formula, Theorem 7.20, we see that X satisfies (here $d = 2$ and $m = 1$)

$$dX_t = a \circ X_t\ dt + b \circ X_t\ dW_t$$

where $a(x) = -\frac{1}{2}(x_1, x_2)$ and $b(x) = (-x_2, x_1)$ for $x = (x_1, x_2)$.

2.12 *Correlated Brownian motions.* With $d = 1$ and $m = 2$, consider the equation 2.7 with $u_0 = 0,\ u_1 = \sin,\ u_2 = \cos$, that is, consider

$$dX_t = (\sin X_t)\, dW_t^1 + (\cos X_t)\, dW_t^2.$$

This process X is a continuous martingale and its quadratic variation has the differential (see Example 7.5, Theorem 7.15, and Lemma 7.22 for these)

$$(dX_t)^2 = (\sin X_t)^2\ dt + (\cos X_t)^2\ dt = dt.$$

It follows from Theorem 7.24 that $X - X_0$ is a Wiener process. For studying X, then, writing $X = X_0 + \hat{W}$ would be simpler. But this simple description is inadequate for describing two motions under the same regime. For instance, in addition to X with $X_0 = x$, let Y satisfy the same equation with $Y_0 = y$, that is, with the same W^1 and W^2 as for X,

$$dY_t = (\sin Y_t)\ dW_t^1 + (\cos Y_t)\, dW_t^2.$$

Then, X and Y are standard Brownian motions, but they depend on each other. Their correlation structure is specified by the cross variation process $\langle X, Y\rangle$, which is given by (in differential form)

$$dX_t\ dY_t = (\sin X_t)(\sin Y_t)\ dt + (\cos X_t)(\cos Y_t)\, dt = \cos(X_t - Y_t)\ dt.$$

Existence and uniqueness

Consider the stochastic integral equation 2.5 under Condition 2.4 on the vector fields u_n. As with deterministic differential equations, Lipschitz continuity 2.4 ensures the existence of a unique solution (in the sense to be explained shortly). The method of solution is also the same as in the deterministic case, namely, Pickard's method of successive approximations. The result is listed next; its proof is delayed to 2.52.

2.13 THEOREM. *The equation 2.5 has a pathwise unique solution X; the process X is continuous.*

2.14 REMARK. The proof 2.52 will also show that X is a *strong solution* in the following sense, thus explaining *pathwise uniqueness*: There exists a unique mapping

$$\varphi : E \times C(\mathbb{R}_+ \mapsto \mathbb{R}^m) \mapsto C(\mathbb{R}_+ \mapsto E)$$

such that, for almost every ω, the paths $X(\omega):\ t \mapsto X_t(\omega)$ and $W(\omega):\ t \mapsto W_t(\omega) = \big(W_t^1(\omega), \ldots,\ W_t^m(\omega)\big)$ satisfy

$$X(\omega) = \varphi\left(X_0(\omega), W(\omega)\right).$$

Markov property

The next theorem shows that X is a (time-homogeneous) Markov process with state space E and transition function (P_t), where

2.15 $$P_t f(x) = \mathbb{E}^x\ f \circ X_t, \quad x \in E,\ f \in \mathcal{E}_+,\ t \in \mathbb{R}_+.$$

2.16 THEOREM. *For each t in $\mathbb{R}_+$, the process $\hat{X} = (X_{t+u})_{u\in\mathbb{R}_+}$ is conditionally independent of $\mathcal{F}_t$ given X_t; moreover, given that $X_t = y$, the conditional law of $\hat{X}$ is the same as the law of X under $\mathbb{P}^y$.*

2.17 REMARK. The claim of the theorem is that, for every integer $k \geq 1$ and positive Borel function f on E^k,

$$\mathbb{E}^x\left(f\left(X_{t+u_1}, \ldots, X_{t+u_k}\right) | \mathcal{F}_t\right) = \mathbb{E}^{X_t} f\left(X_{u_1}, \ldots, X_{u_k}\right),$$

where the right side stands for $g \circ X_t$ with $g(y) = \mathbb{E}^y f(X_{u_1}, \ldots, X_{u_k})$. Of course, as in Theorem 1.2, this is also equivalent to

$$\mathbb{E}^x \left(f \circ X_{t+u} | \mathcal{F}_t \right) = P_u f \circ X_t, \quad x \in E,\ f \in \mathcal{E}_+,\ t, u \in \mathbb{R}_+.$$

Proof. Fix t and let $\hat{W} = (W_{t+u} - W_t)_{u \in \mathbb{R}_+}$. Note that, in the notation system of 2.8,

$$\begin{aligned} \hat{X}_u &= X_t + \int_t^{t+u} a \circ X_s \, ds + \int_t^{t+u} b \circ X_s \, dW_s \\ &= \hat{X}_0 + \int_0^u a \circ \hat{X}_s \, ds + \int_0^u b \circ \hat{X}_s \, d\hat{W}_s. \end{aligned}$$

Thus, with φ defined as in Remark 2.14,

$$X = \varphi\left(X_0, W\right), \qquad \hat{X} = \varphi\left(\hat{X}_0, \hat{W}\right).$$

By the Lévy nature of W, the process $\hat{W}$ is independent of $\mathcal{F}_t$ and is again a Wiener process just as W. Thus, $\hat{X}$ is conditionally independent of $\mathcal{F}_t$ given $\hat{X}_0 = X_t$. Moreover, given that $X_t = \hat{X}_o = y$, the conditional law of $\hat{X}$ is the law of $\varphi(y, \hat{W})$, which is in turn the same as the law of $\varphi(y, W)$, namely, the law of X given $X_0 = y$. □

Strong Markov property

The preceding theorem remains true when the deterministic time t is replaced with a stopping time T, provided that we make provisions for the possibility that T might take the value $+\infty$. To that end we introduce the following.

2.18 CONVENTION. *Let ∂ be a point outside E; put $\bar{E} = E \cup \{\partial\}$, and let $\bar{\mathcal{E}}$ be the σ-algebra on $\bar{E}$ generated by $\mathcal{E}$. We define $X_\infty(\omega) = \partial$ for all ω. Every function $f : E \mapsto \mathbb{R}$ is extended onto $\bar{E}$ by setting $f(\partial) = 0$. If the original f is in $\mathcal{E}_+$, for instance, then the extended function is in $\bar{\mathcal{E}}_+$, but we still write $f \in \mathcal{E}_+$. The convention applies to the function $P_t f$ as well: $P_t f(\partial) = 0$. Finally, $\mathcal{F}_\infty = \lim \mathcal{F}_t = \vee_t \mathcal{F}_t$ as usual.*

2.19 THEOREM. *The process X is strong Markov: For every stopping time T of $\mathcal{F}$, the variable X_T is $\mathcal{F}_T$ – measurable, and the process $\hat{X} = (X_{T+u})_{u \in \mathbb{R}_+}$ is conditionally independent of $\mathcal{F}_T$ given X_T; moreover, for y in E, on the event $\{X_T = y\}$, the conditional law of $\hat{X}$ given X_T is the same as the law of X under $\mathbb{P}^y$.*

Proof. Since X is continuous and adapted to $\mathcal{F}$, the random variable X_T is measurable with respect to $\mathcal{F}_T$ and $\bar{\mathcal{E}}$. The rest of the proof follows that of the last theorem: replace t by T throughout to handle the conditional expectations on the event $\{T < \infty\}$. On the event $\{T = \infty\}$, we have $X_{T+u} = \partial$ for all u, and the claim holds trivially in view of the conventions 2.18. □

Generator

We introduce a differential operator which will describe the differential structure of the transition function (P_t). First, some notation: We put

2.20 $$\mathcal{C} = C(E\mapsto\mathbb{R}), \mathcal{C}_K = C_K\,(E\mapsto\mathbb{R})\,,\ \mathcal{C}^2 = C^2(E\mapsto\mathbb{R}),\ \mathcal{C}^2_K = \mathcal{C}^2 \cap \mathcal{C}_K;$$

Thus, $\mathcal{C}$ is the set of all continuous functions $f:\ E \mapsto \mathbb{R}$, and $\mathcal{C}_K$ is the set of such f with compact support, and $\mathcal{C}^2$ is the set of such f that are twice differentiable with continuous derivatives of first and second order. For f in $\mathcal{C}^2$, we write $\partial_i f$ for the partial derivative with respect to the i^{th} argument, and $\partial_{ij} f$ for the second order partial derivative with respect to the i^{th} and j^{th} arguments; the classical notations are $\partial_i = \frac{\partial}{\partial x_i}$ and $\partial_{ij} = \frac{\partial^2}{\partial x_i \partial x_j}$. When $d = 1$, these become f', the derivative, and f'', the second derivative. With these notations, we introduce the operator G on $\mathcal{C}^2$ by

2.21 $$Gf(x) = \sum_{i=1}^{d} u_0^i(x)\partial_i f(x) + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{n=1}^{m} u_n^i(x)u_n^j(x)\,\partial_{ij} f(x), \quad x \in E.$$

When $d = 1$ (and more generally with proper interpretation, in the notation system of 2.8) this becomes

$$Gf(x) = a(x)f'(x) + {}^1\!/_2\, b(x)^2\, f''\,(x)\,.$$

2.22 Example. *Brownian motion.* Suppose that X is a standard Brownian motion in $\mathbb{R}^d$, that is, take $m = d$ and put $u_0(x) = 0$ and let $u_n^i(x)$ be free of x and equal to 1 or 0 according as $i = n$ or $i \neq n$. Then

$$Gf = \frac{1}{2}\sum_{i=1}^{d} \partial_{ii} f, \qquad f \in \mathcal{C}^2;$$

thus, $Gf = {}^1\!/_2\Delta f$, where Δ is the Laplacian operator of classical analysis.

Itô's formula

The equation 2.5 of motion shows that X is a semimartingale. Applying to it Itô's formula, Theorem 7.20, yields the following.

2.23 Theorem. *For every f in $\mathcal{C}^2_K$,*

2.24 $$M_t = f \circ X_t - f \circ X_0 - \int_0^t ds\ Gf \circ X_s, \qquad t \in \mathbb{R}_+,$$

is a martingale; it is given by

$$M_t = \sum_{n=1}^{m}\sum_{i=1}^{d}\int_0^t \left(u_n^i \circ X_s\right)\left(\partial_i f \circ X_s\right)\, dW_s^n.$$

Proof. We use Itô's formula, Theorem 7.20:

$$d\,(f\circ X_t) = \sum_{i=1}^{d} (\partial_i f\circ X_t)\,dX_t^i + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}(\partial_{ij} f\circ X_t)\,dX_t^i\,dX_t^j.$$

In view of 2.6 for X_t^i, this yields the claim once we note that

$$dX_t^i\,dX_t^j = \sum_{n=1}^{m}\left(u_n^i\circ X_t\right)\left(u_n^j\circ X_t\right)\,dt$$

in view of the rules of Theorem 7.19 and Lemma 7.22. □

2.25 COROLLARY. *Let $f\in\mathcal{C}_K^2$. Then, $Gf\in\mathcal{C}_K$, and*

$$\mathbb{E}^x\ f\circ X_t = f(x) + \mathbb{E}^x\int_0^t ds\ Gf\circ X_s,\qquad x\in E,\ t\in\mathbb{R}_+.$$

Proof. The vector fields u_n are continuous by Condition 2.4. Thus, for f in $\mathcal{C}_K^2$, the formula 2.21 shows that Gf is continuous and has compact support. The claimed formula is now immediate from the preceding theorem, since $\mathbb{E}^x M_t = 0$. □

Moreover, in the preceding corollary, since $Gf\in\mathcal{C}_K$ and thus is bounded, we may change the order of integration and expectation (by Fubini's theorem). Recalling 2.15, then, we obtain the following.

2.26 COROLLARY. *Let $f\in\mathcal{C}_K^2$. Then, $Gf\in\mathcal{C}_K$ and*

$$P_t f(x) = f(x) + \int_0^t ds\ P_s\ Gf(x),\qquad x\in E,\ t\in\mathbb{R}_+.$$

Dynkin's formula

This is essentially Corollary 2.25, but with a stopping time replacing the deterministic time.

2.27 THEOREM. *Let $f\in\mathcal{C}_K^2$. Let T be an $\mathcal{F}$-stopping time. For fixed x in E, suppose that $\mathbb{E}^x T<\infty$; then,*

$$\mathbb{E}^x\ f\circ X_T = f(x) + \mathbb{E}^x\int_0^T ds\ Gf\circ X_s.$$

Proof. Let f, T, x be as described. By Theorem 2.23, the proof is reduced to showing that $\mathbb{E}^x M_T = 0$ for the martingale M there. Since M is a sum of finitely many martingales, it is enough to show that $\mathbb{E}^x\hat{M}_T = 0$ for one of the terms there, say, for the martingale

$$\hat{M}_t = \int_0^t g\circ X_s\ d\hat{W}_s,$$

where, for fixed i and n, we put $g = u_n^i\partial_i f$ and $\hat{W} = W^n$.

Since f has compact support, and since $\partial_i f$ and u_n^i are continuous, the function g is continuous and bounded, say by c. Thus, applying 7.6 with F there taken as the bounded left-continuous process $s \mapsto (g \circ X_s) 1_{\{s \le T\}}$,

$$\mathbb{E}^x \left(\hat{M}_{T\wedge t}\right)^2 = E^x \left(\int_0^t g \circ X_s \; 1_{\{s\le T\}} \, d\hat{W}_s\right)^2$$
$$= \mathbb{E}^x \int_0^t (g\circ X_s)^2 \; 1_{\{s\le T\}} \, ds \le \mathbb{E}^x c^2 T < \infty$$

by the assumption that $\mathbb{E}^x T < \infty$. So, on $(\Omega, \mathcal{H}, \mathbb{P}^x)$, the martingale $(\hat{M}_{T\wedge t})_{t\in\mathbb{R}_+}$ is L^2-bounded and, therefore, is uniformly integrable. By Theorem V.5.14, then, $\hat{\text{M}}$ is a Doob martingale on $[0,T]$, which implies that $\mathbb{E}^x \hat{M}_T = \mathbb{E}^x \hat{M}_0 = 0$ as needed. □

Infinitesimal generator

This is an extension of the operator G defined by 2.21. We keep the same notation, but we define it anew.

Let $\mathcal{D}_G$ be the collection of functions $f : E \mapsto \mathbb{R}$ for which the limit

2.28 $$Gf(x) = \lim_{t\downarrow 0} \frac{1}{t}\left[P_t f(x) - f(x)\right]$$

exists for every x in E. Then, G is called the *infinitesimal generator* of X, and $\mathcal{D}_G$ is called its *domain.*

2.29 LEMMA. *Let $f \in \mathcal{C}_K^2$. Then, $f \in \mathcal{D}_G$, and the limit in 2.28 is given by 2.21.*

Proof. Let $f \in \mathcal{C}_K^2$, define *Gf* by 2.21. By Corollary 2.26, then, *Gf* is continuous and bounded, which implies that $P_s Gf(x) = \mathbb{E}^x\, Gf \circ X_s$ goes to *Gf(x)* as $s \to 0$; this is by the bounded convergence theorem and the continuity of X. Thus, from the formula of 2.26, *Gf(x)* is equal to the limit on the right side of 2.28. □

Forward and backward equations

2.30 THEOREM. *Let $f \in \mathcal{C}_K^2$. Then, $f \in \mathcal{D}_G$, Gf is given by 2.21, and*

2.31 $$\frac{d}{dt} P_t f(x) = P_t G f(x), \qquad x \in E,\; t \in \mathbb{R}_+.$$

Moreover, for f in $\mathcal{C}_K^2$ again, $P_t f \in \mathcal{D}_G$ and, with G as in 2.28,

2.32 $$\frac{d}{dt} P_t f(x) = G P_t f(x), \qquad x \in E,\; t \in \mathbb{R}_+.$$

2.33 Remark. The equation 2.31 is called *Kolmogorov's forward* equation, because G is in front of P_t. By the same logic, 2.32 is called *Kolmogorov's backward equation*. Writing $u(t,x)$ for $P_t f(x)$ for fixed f, the backward equation can be re-written as

$$\frac{d}{dt}\, u = Gu, \qquad u(0,x) = f(x),$$

with the understanding that G applies to the spatial variable, that is, to $x \mapsto u(t,x)$. This sets up a correspondence between diffusions and partial differential equations, since functions in $\mathcal{D}_G$ can be approximated by sequences of functions in $\mathcal{C}_K^2$.

Proof. The first statement is mostly in Lemma 2.29 and Corollary 2.26: Let $f \in \mathcal{C}_K^2$. Then, Gf is given by 2.21, belongs to $\mathcal{D}_G$, and $s \mapsto Gf \circ X_s$ is continuous and bounded. Thus, by the bounded convergence theorem, $s \mapsto P_s Gf(x) = \mathbb{E}^x\, Gf \circ X_s$ is continuous and bounded. Hence, in the equation for $P_t f$ given in Corollary 2.26, the integral on the right side defines a differentiable function in t; and, taking derivatives on both sides yields 2.31.

For f in $\mathcal{C}_K^2$, we have just shown that $t \mapsto P_t f(x)$ is differentiable. Thus, since $P_s P_t = P_{t+s}$, the limit

$$\begin{aligned} GP_t f(x) &= \lim_{s\to 0} \frac{1}{s}\,[P_s P_t f(x) - P_t f(x)] \\ &= \lim_{s\to 0} \frac{1}{s}\,[P_{t+s} f(x) - P_t f(x)] = \frac{d}{dt}\, P_t f(x) \end{aligned}$$

exists, that is, $P_t f \in \mathcal{D}_G$ and 2.32 holds. □

Potentials, resolvent

Let $f \in \mathcal{E}_b$ and $p > 0$. By the continuity of X, the mapping $(t,\omega) \mapsto X_t(\omega)$ is measurable relative to $\mathcal{B}_{\mathbb{R}_+} \otimes \mathcal{H}$ and $\mathcal{B}(\mathbb{R}^d)$. Thus, the following defines a function in $\mathcal{E}_b$:

2.34 $$U_p f(x) = \mathbb{E}^x \int_0^\infty dt\; e^{-pt}\; f \circ X_t = \int_0^\infty dt\; e^{-pt}\; P_t f(x), \quad x \in E.$$

The function $U_p f$ is called the p-potential of f, and U_p is called the p-potential operator, and the family $(U_p)_{p>0}$ is called the *resolvent* of (P_t) or of X. Of course, 2.34 makes sense for f in $\mathcal{E}_+$ and $p \ge 0$ as well. The next theorem relates the resolvent to the infinitesimal generator: the operators U_p and $p - G$ are inverses of each other.

2.35 Theorem. *For $p > 0$ and $f \in \mathcal{C}_K^2$,*

$$U_p(p - G)f = (p - G)\, U_p f = f.$$

Proof. Fix p and f such. From Corollary 2.26, then,

$$pU_pf = \int_0^\infty dt\ pe^{-pt}\ f + \int_0^\infty dt\ pe^{-pt} \int_0^t ds\ P_s\ Gf = f + U_pGf$$

by a change in the order of integration over s and t. Thus, $U_p(pf - Gf) = f$. For the other claim, we start by noting that, since $P_sP_t = P_{s+t}$ and f is bounded,

$$P_sU_pf = \int_0^\infty dt\ e^{-pt}\ P_{s+t}\ f = e^{ps}\ U_pf - e^{ps} \int_0^s dt\ e^{-pt}\ P_tf.$$

In the rightmost integral, the integrand goes to f as $t \to 0$. Thus,

$$\begin{aligned} GU_pf &= \lim_{s\to 0} \frac{1}{s}\left(P_sU_pf - U_pf\right) \\ &= \lim_{s\to 0} \frac{e^{ps}-1}{s} U_pf - \lim_{s\to 0} \frac{1}{s}\ e^{ps} \int_0^s dt\ e^{-pt}\ P_tf = pU_pf - f. \end{aligned}$$

Thus, $(p - G)U_pf = f$ as well. □

Interpretations

Fix f in $\mathcal{E}_b$ and $p > 0$. Let T_p be a random variable having the exponential distribution with parameter p. Suppose that T_p is independent of X. Since $\mathbb{P}\{T_p > t\} = e^{-pt}$, we may express 2.34 as

2.36 $$U_pf(x) = \mathbb{E}^x \int_0^\infty dt\ 1_{\{T_p>t\}} f \circ X_t = \mathbb{E}^x \int_0^{T_p} dt\ f \circ X_t,$$

which is the expected earnings during $(0, T_p)$ if the rate of earnings is $f(y)$ per unit of time spent at y. A related interpretation is that

2.37 $$pU_pf(x) = \mathbb{E}^x \int_0^\infty dt\ pe^{-pt}\ f \circ X_t = \mathbb{E}^x\ f \circ X_{T_p},$$

and, equivalently, writing $P_{T_p}f(x)$ for $g \circ T_p$ with $g(t) = P_tf(x)$,

2.38 $$pU_pf(x) = \mathbb{E}\ P_{T_p}\ f(x).$$

These show, in particular, that pU_p is a Markov kernel on $(E, \mathcal{E})$. Moreover, noting that $T = pT_p$ is standard exponential, and since $\frac{1}{p}T \to 0$ as $p \to \infty$, it follows from 2.38 that

2.39 $$\lim_{p\to\infty} pU_p\ f(x) = f(x)$$

provided that $\lim_{t\to 0} P_tf(x) = f(x)$, for instance, if f is continuous in addition to being bounded.

Resolvent equation

2.40 THEOREM. *For $p > 0$ and $q > 0$, we have $U_p U_q = U_q\, U_p$, and*

2.41 $$U_p + p\, U_p\, U_q = U_q + q\, U_q\, U_p.$$

Proof. Let $f \in \mathcal{E}_b$. Let T_p and T_q be independent of each other and of X, both exponential variables, with respective parameters p and q. Since $P_s P_t = P_t P_s$, it follows from 2.38 that

$$pq\, U_p U_q f(x) = \mathbb{E}\, P_{T_p}\, P_{T_q}\, f(x) = \mathbb{E}\, P_{T_q} P_{T_p}\, f(x) = qp\, U_q U_p f(x),$$

that is, $U_p\, U_q = U_q\, U_p$. To show the resolvent equation 2.41, we start with the ordinary Markov property:

$$\begin{aligned}\mathbb{E}^x \int_0^{s+t} du\; f\circ X_u &= \mathbb{E}^x \int_0^s du\; f\circ X_u + \mathbb{E}^x \int_0^t du\; f\circ X_{s+u} \\ &= \mathbb{E}^x \int_0^s du\; f\circ X_u + \int_E P_s(x,dy)\; \mathbb{E}^y \int_0^t du\; f\circ X_u.\end{aligned}$$

Since T_p and T_q are independent of X, we may replace s with T_p, and t with T_q. Then, using the interpretations 2.36 and 2.38, we get

$$\mathbb{E}^x \int_0^{T_p+T_q} du\; f\circ X_u = U_p f(x) + p\, U_p U_q f(x).$$

This proves 2.41 since $T_p + T_q = T_q + T_p$. □

Killing the diffusion

This is to describe an operation that yields an absorbing Markov process that coincides with X over an initial interval of time. Here X is the diffusion (described by Theorem 2.5 and examined above) with state space $E = \mathbb{R}^d$.

Let k be a positive Borel function on E. Let T be independent of the process X and have the standard exponential distribution (with mean 1). Define, for t in $\mathbb{R}_+$ and ω in Ω,

2.42 $$\hat{X}_t(\omega) = \begin{cases} X_t(\omega) & \text{if } T(\omega) > \int_0^t ds\; k\circ X_s(\omega), \\ \partial & \text{otherwise,}\end{cases}$$

where ∂ is a point outside E. This defines a stochastic process $\hat{X}$ with state space $\bar{E} = E \cup \{\partial\}$. We think of ∂ as the cemetery; it is a trap, and

2.43 $$\zeta = \inf\left\{t \in \mathbb{R}_+ : \hat{X}_t = \partial\right\}$$

is the time X is killed. It follows from 2.42 and the assumptions on T that, with $\exp_{-x} = e^{-x}$ and $\mathcal{G}^0_\infty = \sigma\{X_s = s \in \mathbb{R}_+\}$,

2.44 $$\mathbb{P}\left\{\zeta > t \,|\, \mathcal{G}^0_\infty\right\} = \exp_- \int_0^t ds\; k\circ X_s.$$

Thus, in the language of Chapter VI, the particle X is killed at the time ζ of first arrival in a conditionally Poisson process with random intensity process $k \circ X$. It is common to refer to $\hat{X}$ as the process obtained from X by killing X at the rate $k(x)$ when at x.

The process $\hat{X}$ is Markov with state space $(\bar{E}, \bar{\mathcal{E}})$; its living space is $(E, \mathcal{E})$. We adopt the conventions 2.18 regarding the trap ∂; recall that every f in $\mathcal{E}_+$ is extended onto $\bar{E}$ by setting $f(\partial) = 0$. Thus, the transition function of $\hat{X}$ is determined by

2.45
$$\begin{aligned}\hat{P}_t f\,(x) &= \mathbb{E}^x\ f \circ \hat{X}_t \\ &= \mathbb{E}^x\ f \circ X_t\ 1_{\{\zeta > t\}} = \mathbb{E}^x\,(f \circ X_t)\left(\exp_- \int_0^t ds\ k \circ X_s\right),\end{aligned}$$

with f in $\mathcal{E}_+$ and t in $\mathbb{R}_+$ and x in E. The Markov property of $\hat{X}$ implies that $(\hat{P}_t)$ is a transition semigroup. Each $\hat{P}_t$ is a sub-Markov kernel on $(E, \mathcal{E})$; the defect $1 - \hat{P}_t(x,\ E)$ being $\mathbb{P}^x\{\zeta \le t\}$. The following relates $(\hat{P}_t)$ to (P_t).

2.46 Proposition. *Let $t \in \mathbb{R}_+,\ x \in E,\ f \in \mathcal{E}_+$. Then,*

2.47
$$P_t f(x) = \hat{P}_t f(x) + \int_0^t ds\ \int_E \hat{P}_s\,(x, dy)\ \ k(y)\ P_{t-s} f(y).$$

Proof. We condition on whether killing occurs before or after time t, and we use the Markov property of X:

$$\begin{aligned}P_t f(x) &= \mathbb{E}^x\ f \circ X_t\ 1_{\{\zeta > t\}} + \mathbb{E}^x\ f \circ X_t\ 1_{\{\zeta \le t\}} \\ &= \mathbb{E}^x\ f \circ \hat{X}_t + \int_{[0,t]\times E} \mathbb{P}^x\,\{\zeta \in ds, X_\zeta \in dy\}\, P_{t-s}\ f(y).\end{aligned}$$

This yields the claim via 2.45 and the observation that

$$\mathbb{P}^x\,\{\zeta \in ds,\ X_\zeta \in dy\} = ds\ \hat{P}_s(x, dy)\ k(y). \qquad \square$$

Let $(\hat{U}_p)$ denote the resolvent of the semigroup $(\hat{P}_t)$, and recall the resolvent (U_p) of (P_t). Taking Laplace transforms on both sides of 2.47 we get

2.48
$$U_p f(x) = \hat{U}_p f(x) + \int_E \hat{U}_p(x, dy)\ k(y)\ U_p f(x).$$

We use this to obtain the generator $\hat{G}$ corresponding to $(\hat{P}_t)$ from the generator G of (P_t); see 2.21 and 2.28; in particular, $\hat{G}$ is defined by 2.28 from $(\hat{P}_t)$.

Let $f \in \mathcal{C}_K^2$. Recall Theorem 2.35 to the effect that $f = U_p(p-G)f$. Thus, in view of 2.48,

$$f = \hat{U}_p(p - G)\ f + \hat{U}_p(kf) = \hat{U}_p(p - \hat{G})f$$

where

2.49 $$\hat{G}f(x) = Gf(x) - k(x)f(x), \qquad f \in \mathcal{C}_K^2,\ x \in E.$$

In words, every f in $\mathcal{C}_K^2$ is in the domain of $\hat{G}$, and $\hat{G}$ is related to G and k through 2.49. Considering the relationship 2.28 for $\hat{G}$ and $(\hat{P}_t)$, and considering the formula 2.45 for $\hat{P}_t$, we obtain the following. This is known as *Feynman-Kac formula.*

2.50 PROPOSITION. *Let $f \in \mathcal{C}_K^2$ and put*

$$u(t,x) = \mathbb{E}^x (f \circ X_t) \left(\exp_- \int_0^t ds\ k \circ X_s \right), \qquad t \in \mathbb{R}_+, x \in E.$$

Then, u satisfies the partial differential equation

$$\frac{\partial}{\partial t}\, u = Gu - ku, \qquad u(0,\cdot) = f.$$

Proof of existence and uniqueness

We start the proof of Theorem 2.13 with a lemma on some approximations. We omit time subscripts that are variables of integration. Condition 2.4 is in force throughout.

2.51 LEMMA. *Let Y and Z be continuous processes with state space $E = \mathbb{R}^d$, put*

$$A_t = \int_0^t (u_0 \circ Y - u_0 \circ Z)\ ds,\ \ M_t = \sum_{n=1}^m \int_0^t (u_n \circ Y - u_n \circ Z)\ dW^n.$$

Then,

$$\mathbb{E} \sup_{s \le t} |A_s + M_s|^2 \le (2t + 8m)\, c^2 \int_0^t \mathbb{E}\, |Y - Z|^2\ ds.$$

Proof. By ordinary considerations, using the Lipschitz condition 2.4,

$$\mathbb{E}\, |A_t|^2 \le t \int_0^t |u_0 \circ Y - u_0 \circ Z|^2\ ds \le tc^2 \int_0^t \mathbb{E}\, |Y - Z|^2\ ds.$$

Applying the rule 7.6 to each component of the d-dimensional martingale M, recalling that the W^n are independent, we get

$$\mathbb{E}\, |M_t|^2 = \sum_{n=1}^m \int_0^t |u_n \circ Y - u_n \circ Z|^2\ ds \le mc^2 \int_0^t \mathbb{E}\, |Y - Z|^2\ ds,$$

where the last step used the Lipschitz condition. From these, we obtain

$$\mathbb{E} \sup_{s \le t} |A_s|^2 \le tc^2 \int_0^t \mathbb{E}\, |Y - Z|^2\ ds,$$

$$\mathbb{E} \sup_{s \le t} |M_s|^2 \le 4\ \mathbb{E}\, |M_t|^2 \le 4\ mc^2 \int_0^t \mathbb{E}\, |Y - Z|^2\ ds,$$

where we used the Doob-Kolmogorov inequality in the last line. These two last expressions yield the lemma. □

2.52 *Proof of Theorem 2.13*. Consider the equation 2.5 with $X_0 = x$ for some fixed x in E. We define a sequence of continuous processes $X^{(k)}$ by setting $X_t^{(0)} = x$ for all t, and by letting

2.53 $$X_t^{(k+1)} = x + \int_0^t u_0 \circ X_s^{(k)}\, ds + \sum_{n=1}^{m} \int_0^t u_n \circ X_s^{(k)}\, dW_s^n$$

for $k \geq 0$. Then, $X^{(k+1)} - X^{(k)} = A + M$, where A and M are as in the preceding lemma with $Y = X^{(k)}$ and $Z = X^{(k-1)}$. Fix τ in $\mathbb{R}_+$ and put $\alpha = (2\tau + 8\,m)c^2$. It follows from the lemma that

2.54 $$\mathbb{E} \sup_{s\leq t} \left| X_s^{(k+1)} - X_s^{(k)} \right|^2 \leq \alpha \int_0^t \mathbb{E} \left| X_s^{(k)} - X_s^{(k-1)} \right|^2\, ds$$

for every $t \leq \tau$ and $k \geq 1$. Whereas, by the lemma again, this time with $Y = X^{(1)}$ and $Z = X^{(0)} = x$,

2.55 $$\mathbb{E} \left| X_t^{(1)} - X_t^{(0)} \right|^2 = \mathbb{E} \left| u_0(x)t + \sum_{n=1}^{m} u_n(x) W_t^n \right|^2 = |u_0(x)|^2\ t^2 + \sum_{n=1}^{m} |u_n(x)|^2\, t \leq \beta$$

where $\beta = (\tau^2 + m\tau)\ c^2(1 + |x|)^2$ in view of Condition 2.4.

We put the bound 2.55 into 2.54 with $k = 1$, put the resulting inequality back into 2.54 with $k = 2$, and continue recursively. We get

$$\mathbb{E} \sup_{s\leq\tau} \left| X_s^{(k+1)} - X_s^{(k)} \right|^2 \leq \beta\ \alpha^k \tau^k / k!,$$

which, via Markov's inequality, yields

2.56 $$\mathbb{P} \left\{ \sup_{s\leq\tau} \left| X_s^{(k+1)} - X_s^{(k)} \right|^2 > \frac{1}{2^k} \right\} \leq \beta\, (4\alpha\tau)^k / k!.$$

The right side is summable over k. By the Borel–Cantelli lemma, then, there is an almost sure event Ω_τ such that, for every ω in Ω_τ, the sequence $\left(X_t^{(k)}(\omega) \right)_{k\in\mathbb{N}}$ is convergent in $E = \mathbb{R}^d$ uniformly for t in $[0, \tau]$. We define $X_t(\omega)$ to be the limit for ω in Ω_τ and put $X_t(\omega) = x$ for all other ω.

It follows from the uniformity of convergence and the continuity of $X^{(k)}$ that X is continuous on $[0, \tau]$. It follows from 2.53 that X satisfies the equation 2.5 for $t \leq \tau$. And, τ is arbitrary.

There remains to show the uniqueness. Let X and $\hat{X}$ be solutions to 2.5 with $X_0 = \hat{X}_0 = x$. Then, $X - \hat{X} = A + M$ in the notation of Lemma 2.51 with $Y = X$ and $Z = \hat{X}$. Thus, for fixed τ in $\mathbb{R}_+$, we have

$$\mathbb{E} \sup_{s \le t} \left| X_s - \hat{X}_s \right|^2 \le (2\tau + 8m)\, c^2 \int_0^t \mathbb{E} \left| X_s - \hat{X}_s \right|^2 \, ds$$

for all $t \le \tau$. It now follows from Gronwall's inequality (see Exercise 2.70) that the left side vanishes. Thus, almost surely, $X_t = \hat{X}_t$ for all $t \le \tau$; and τ is arbitrary. □

Dependence on the initial position

Let $X_t(\omega, x)$ denote the position $X_t(\omega)$ when $X_0 = x$. The next proposition shows that the dependence of X_t on x is continuous in the L^2-space of $(\Omega, \mathcal{H}, \mathbb{P})$ and, hence, in probability.

2.57 PROPOSITION. *For each t in $\mathbb{R}_+$,*

$$\lim_{x \to y} \mathbb{E} \left| X_t(x) - X_t(y) \right|^2 = 0.$$

Proof. Fix x and y in E. Note that $X_t(x) - X_t(y) = x - y + A_t + M_t$ in the notation of Lemma 2.51 with $Y_t = X_t(x)$ and $Z_t = X_t(y)$. Thus with fixed $\tau < \infty$ and $\alpha = (2\tau + 8m)c^2$, we have

$$\mathbb{E} \left| X_t(x) - X_t(y) \right|^2 \le 2 \left| x - y \right|^2 + 2\alpha \int_0^t \mathbb{E} \left| X_s(x) - X_s(y) \right|^2 \, ds$$

for all $t \le \tau$. Via Gronwall's inequality (see 2.70), this implies that

$$\mathbb{E} \left| X_t(x) - X_t(y) \right|^2 \le 2 \left| x - y \right|^2 \, e^{2\alpha t}, \qquad 0 \le t \le \tau.$$

The claim is immediate since τ is arbitrary. □

The preceding proposition implies that $X_t(x) \to X_t(y)$ in probability as $x \to y$ in $E = \mathbb{R}^d$. Thus, for $f : E \mapsto \mathbb{R}$ bounded and continuous, as $x \to y$,

2.58 $$P_t f(x) = \mathbb{E}\, f \circ X_t(x) \to \mathbb{E}\, f \circ X_t(y) = P_t f(y)$$

as in Theorem III.1.6. In other words, if f is bounded continuous, then so is $P_t f$ for each t. This is called the *Feller property* for (P_t); it will show that Itô diffusions form a subset of Hunt processes to be introduced in Section 5.

Exercises and complements

2.59 *Differential operator.* Specify the operator G defined by 2.21 for

a) the geometric Brownian motion of Example 2.9,

b) Ornstein-Uhlenbeck process of 2.10,

c) Brownian motion on the unit circle, Example 2.11.

2.60 *Ornstein-Uhlenbeck model.* Let V be the Ornstein–Uhlenbeck velocity process for the physical Brownian motion on $\mathbb{R}$; it satisfies

$$dV_t = -cV_t\ dt + b\ dW_t,$$

where $c > 0$ and $b > 0$ are constants, and W is Wiener. Then, the particle position process X satisfies $dX_t = V_t\ dt$. Write the equation of motion for the $\mathbb{R}^2$-valued motion (V_t, X_t). What is the corresponding generator G on $C_K^2(\mathbb{R}^2 \mapsto \mathbb{R})$? show that V and X are Gaussian processes assuming that $V_0 = v$ and $X_0 = x$ are fixed.
Hint: Write the solution for V, and use integration by parts to express V as an ordinary integral of W.

2.61 *Graphs.* Let X be an Itô diffusion satisfying 2.5. Put $Y_t = (t, X_t)$. Write Itô's formula for $f \circ Y_t$ with f in $C_K^2(\mathbb{R} \times \mathbb{R}^d \mapsto \mathbb{R})$. Show that Y is an Itô diffusion that satisfies

$$dY_t = \sum_{n=0}^{m} v_n \circ Y_t\ d\ Z_t^n,$$

where $Z_t^o = t$ and $Z_t^n = W_t^n$ for $n \geq 1$ and the vector fields $v_0, \ldots, v_m$ on $\mathbb{R}^{d+1}$ chosen appropriately. Specify the v_n.

2.62 *Applications to Brownian motion.* Here and in Exercises 2.63–2.66 below, X is a standard Brownian motion in $E = \mathbb{R}^d$ as in Example 2.22. For Borel subsets D of E define τ_D to be the time of exit from D, that is,

$$\tau_D = \inf\{t \in \mathbb{R}_+ :\ X_t \notin D\}.$$

Recall that, when $d = 1$ and $D = (-r,\ r)$, we have $\mathbb{E}^0\tau_D = r^2$. Show that, in general, $\mathbb{E}^x\tau_D < \infty$ for x in D, for D bounded.
Hint: If D is bounded, it is contained in an open ball of some radius $r < \infty$ centered at x, and that ball is contained in the cylinder $C = (x_1 - r,\ x_1 + r) \times \mathbb{R}^{d-1}$ if $x = (x_1, \ldots, x_d)$. Then, $\tau_D \leq \tau_C$, and $\mathbb{E}^x\tau_C = r^2$.

2.63 *Continuation.* Let D be a ball of radius r centered at the origin. Show that, for x in D,

$$\mathbb{E}^x\ \tau_D = \frac{r^2 - |x|^2}{d}$$

Hint: Use Dynkin's formula, Theorem 2.27, with f in $\mathcal{C}_K^2$ chosen so that $f(x) = |x|^2$ for x in D.

2.64 *Hitting of spheres.* For $r \geq 0$, let T_r be the time that Brownian motion X hits the sphere of radius r centered at the origin of $\mathbb{R}^d$. For $0 < q < |x| < r$, consider the probability

$$\alpha = \mathbb{P}^x\{T_q < T_r\},$$

that is, the probability that X exits $D = \{x \in E :\ q < |x| < r\}$ by touching the inner sphere.

a) For $d = 1$, X is a Doob martingale on $[0, \tau_D]$; use this to show that $\alpha = (r - |x|)/(r - q)$.

b) Let $d = 2$. Show that $\alpha = (\log r - \log|x|)/(\log r - \log q)$.

Hint: Let $f \in \mathcal{C}_K^2$ such that $f(x) = \log|x|$ for x in D. Use Dynkin's formula for such f and stopping time τ_D.

c) Let $d = 3$. Show that

$$\alpha = \left(r^{2-d} - |x|^{2-d}\right) \Big/ \left(r^{2-d} - q^{2-d}\right).$$

Hint: use Dynkin's formula with f in $\mathcal{C}_K^2$ such that $f|x| = |x|^{2-d}$ for x in D.

2.65 *Recurrence properties.* For $d = 1$, the Brownian motion X will hit every point y repeatedly without end; see Chapter VIII.

a) Let $d = 2$. Let $r \to \infty$ in 2.64b to show that

$$\mathbb{P}^x\{T_q < \infty\} = 1, \qquad 0 < q < |x|,$$

however small the disk of radius q is. Show, however, that

$$\mathbb{P}^x\{T_0 < \infty\} = 0, \qquad |x| > 0.$$

b) Let $d = 3$ and $0 < q < |x|$. Show that

$$\mathbb{P}^x\{T_q < \infty\} = (q/|x|)^{d-2}.$$

In summary, standard Brownian motion is "point recurrent" for $d = 1$, fails to be point recurrent but is "disk recurrent" for $d = 2$, and is "transient" for $d \geq 3$.

2.66 *Bessel processes with index $d \geq 2$.* Let $X = X_0 + W$, a standard Brownian motion in $\mathbb{R}^d$. Define $R = |X|$. Started at $x \neq 0$, the process X never visits the point 0; see 2.65a. Thus, the true state space for R is $(0, \infty)$. Since $d \geq 2$, the function $f : x \mapsto |x|$ is twice differentiable everywhere except the origin.

a) Use Itô's formula on $R = f \circ X$ to show that

$$dR_t = \frac{d-1}{2\ R_t}\ dt + \sum_{i=1}^{d} \frac{1}{R_t} X_t^i\ dW_t^i = \frac{d-1}{2R_t}\ dt + d\hat{W}_t$$

with an obvious definition for $\hat{W}$.

b) Show that $\hat{W}$ is a continuous local martingale with $\hat{W}_0 = 0$. Show, using 7.24, that $\hat{W}$ is a Wiener process (one-dimensional).

2.67 *Bessel with index 2.* Let $d = 2$ in the preceding exercise, and let $R_0 = r > 0$ be fixed. Define

$$Y_t = \log\ R_t, \qquad t \in \mathbb{R}_+.$$

a) Show that Y is a continuous local martingale with $Y_0 = \log r$.

b) Let τ be the time of exit for R from the interval $(p,\ q)$, where $0 < p < r < q$. Show that $\tau < \infty$ almost surely and that Y is bounded on $[0, \tau]$. Show that, as in 2.64b,

$$\mathbb{P}\left\{R_\tau = p\right\} = \frac{\log\ q - \log\ r}{\log\ q - \log\ p}.$$

2.68 *Continuation.* a) Show that $C = \langle Y, Y\rangle$ is given by

$$C_t = \int_0^t ds\ e^{-2Y_s}.$$

Use Theorem 7.27 to conclude that

$$Y_t = \log r + \tilde{W}_{C_t}$$

for some Wiener process $\tilde{W}$.

b) Solve the last equation for Y by expressing C_t in terms of $\tilde{W}$.

c) Conclude that the Bessel process R is a time-changed geometric Brownian motion: $R = Z_C$, where

$$Z_s = re^{\tilde{W}_s}, \qquad S_u = \int_0^u ds\,(Z_s)^2, \qquad C_t = \inf\left\{u > t:\ S_u > t\right\}.$$

2.69 *Bessel with index $d \geq 3$.* Take $d = 3$ in 2.66 and fix $R_0 = r > 0$.

a) Show that $Y = R^{2-d}$ is a local martingale.

b) Use Theorem 7.27 to show that $Y = Y_0 + \tilde{W}_C$, where $\tilde{W}$ is a Wiener process and $C = \langle Y, Y\rangle$. Thus,

$$Y_t = Z_{C_t},$$

where $Z_u = r^{2-d} + \tilde{W}_u,\ u \geq 0$, a Brownian motion.

c) Show that C is the functional inverse of S, where

$$S_u = (d-2)^{-2}\int_0^u (Z_s)^{(2-2d)/(d-2)}\ ds.$$

Conclude that R is a deterministic function of a random time-changed Brownian motion:

$$R_t = \left(Z_{C_t}\right)^{-1/(d-2)}.$$

2.70 *Gronwall's inequality.* Let f and g be positive continuous functions on $\mathbb{R}_+$. Suppose that, for some c in $\mathbb{R}$,

$$f(t) \leq g(t) + c\int_0^t f(s)ds.$$

show that, then,

$$f(t) \le g(t) + c\int_0^t e^{c(t-s)}\, g(s)\, ds.$$

Hint: First show that

$$e^{-ct}\int_0^t f(s)\, ds \le \int_0^t e^{-cs}\, g(s)\, ds.$$

3 Jump-Diffusions

Jump-diffusions are processes that are Itô diffusions between the jumps. The jump times form a point process, and the diffusions and jumps interact. The treatment uses notions from Itô diffusions and Poisson random measures.

The motion of interest is a right-continuous, piecewise continuous process with state space $E = \mathbb{R}^d$ and the attendant σ-algebra $\mathcal{E} = \mathcal{B}(\mathbb{R}^d)$. Throughout, $(\Omega, \mathcal{H}, \mathbb{P})$ is a complete probability space, and $\mathcal{F}$ is an augmented right-continuous filtration. Adapted to $\mathcal{F}$, and independent of each other, W is an m-dimensional Wiener process and M is a standard Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}_+$ (with mean $Leb \times Leb$). In addition, X_0 is an E-valued random variable in $\mathcal{F}_0$; it will often be treated as a deterministic parameter. As before, $(x, H) \mapsto \mathbb{P}^x(H)$ is a regular version of the conditional probabilities 2.3.

The motion

The deterministic data are some vector fields $u_0, \ldots, u_m$ on E and a Borel function $j : E \times \mathbb{R}_+ \mapsto E$. The vector fields are as in the preceding section; the function j rules the jump sizes. The motion X of interest satisfies the following stochastic integral equation:

$$X_t = X_0 + \int_0^t a \circ X_s\, ds + \int_0^t b \circ X_s\, dW_s + \int_{[0,t]\times\mathbb{R}_+} M\,(ds, dv)\, j(X_{s-}, v).$$

3.1

Here, $a = u_0$ and b is the $d \times m$ matrix whose columns are $u_1, \ldots, u_m$; see 2.5–2.8 for various equivalent ways of expressing the first two integrals. Unless stated otherwise, the next condition is in force throughout.

3.2 Condition. a) Lipschitz condition 2.4 holds. b) There is a constant c in $\mathbb{R}_+$ such that $j(x, v) = 0$ for $v > c$ for all x in E.

This condition is sufficient to ensure the existence and uniqueness of a piecewise continuous solution to 3.1. The condition on j makes the last integral in 3.1 to be effectively over $[0, t] \times [0, c]$, which means that the jump times of X form a subset of the arrival times in a Poisson process with rate c. Between two successive jumps, the motion is an Itô diffusion satisfying

3.3 $$d\bar{X}_t = a \circ \bar{X}_t\, dt + b \circ \bar{X}_t\, dW_t.$$

In particular, the initial segment (before the first jump) of X coincides with the Itô diffusion $\bar{X}$ satisfying 3.3 with the initial condition $\bar{X}_0 = X_0$.

Obviously, if $j = 0$, then $X = \bar{X}$; this was the subject of the preceding section. At the other extreme is the case $u_0 = \cdots = u_m = 0$, in which case X is piecewise constant, that is, each path $X(\omega)$ is a step function; the reader is invited to take this to be the case on a first reading; we shall treat this special case toward the end of this section. In the middle, there is the case where $u_1 = \cdots = u_m = 0$, in which case X is piecewise deterministic.

Construction of X

The next theorem describes X under the standing condition 3.2. The proof is constructive and is helpful for visualizing the paths.

3.4 Theorem. *The equation 3.1 has a pathwise unique solution X that is piecewise continuous, right-continuous, and locally bounded.*

Proof. For fixed x in E and s in $\mathbb{R}_+$, let $(t, \omega) \mapsto \bar{X}_{s,t}(\omega, x)$ be the process that is the solution $\bar{X}_t$ to 3.3 with $t \geq s$ and $\bar{X}_s = x$. Under the Lipschitz condition 2.4, Theorem 2.13 applies, and $t \mapsto \bar{X}_{s,t}(\omega, x)$ is pathwise unique and continuous for almost every ω. Under condition 3.2b, the last integral in 3.1 is over $[0, t] \times [0, c]$ effectively. Since M is standard Poisson, its atoms over $\mathbb{R}_+ \times [0, c]$ can be labeled $(S_n,\ V_n)$ so that, for almost every ω,

3.5 $$0 < S_1(\omega) < S_2(\omega) < \cdots, \ \lim S_n(\omega) = +\infty.$$

By eliminating from Ω a negligible event, we assume that these properties (on $\bar{X}$ and M) hold for every ω.

Fix ω, put $S_0(\omega) = 0$, and suppose that $X_t(\omega)$ is specified for all $t \leq s$, where $s = S_n(\omega)$ for some $n \geq 0$. We proceed to specify it for t in $(s,u]$, where we put $u = S_{n+1}(\omega)$. Since M_ω has no atoms in $(s, u) \times [0, c]$, the equation 3.1 implies that

$$X_t(\omega) = \bar{X}_{s,t}(\omega, X_s(\omega)), \qquad s \leq t < u.$$

Since $t \mapsto \bar{X}_{s,t}(\omega, x)$ is continuous and bounded on the interval $[s,u]$, we have

$$X_{u-}(\omega) = \lim_{t \nearrow u} X_t(\omega) = \bar{X}_{s,u}(\omega, X_s(\omega)),$$

which point is in E. Now, 3.1 implies that

$$X_u(\omega) = X_{u-}(\omega) + j\left(X_{u-}(\omega), V_{n+1}(\omega)\right).$$

This completes the specification of $X_t(\omega)$ for $t \leq S_{n+1}(\omega)$, and therefore for all t in $\mathbb{R}_+$, by recursion, in view of 3.5. The other claims of the theorem are immediate from this construction. □

In the proceding proof, the times S_n are the arrival times of the Poisson process $t \mapsto M([0, t] \times [0, c])$. The proof shows that the path $X(\omega)$ is continuous except possibly at times $S_n(\omega)$. Generally, not every $S_n(\omega)$ is a jump time for $X(\omega)$. See Exercise 3.88 for an example where X has at most finitely many jumps, and, in fact, there is a strictly positive probability that it has no jumps.

Markov and strong Markov properties

The process X has both. The proofs follow the same lines as those for diffusions: uniqueness of solutions to 3.1, and the Markov and strong Markov properties for W and M.

3.8 THEOREM. *For each time t, the process $\hat{X} = (X_{t+u})_{u\in\mathbb{R}_+}$ is conditionally independent of $\mathcal{F}_t$ given X_t; given that $X_t = y$, the conditional low of $\hat{X}$ is the same as the law of X under $\mathbb{P}^y$.*

Proof. Analogous to the proof of Theorem 2.16, we have

$$\hat{X}_u = \hat{X}_0 + \int_0^u a \circ \hat{X}_s \, ds + \int_0^u b \circ \hat{X}_s \, d\hat{W}_s + \int_{[0,u]\times\mathbb{R}_+} \hat{M}\,(ds, dv)\, j\left(\hat{X}_{s-}, v\right),$$

where $\hat{W} = (W_{t+u} - W_t)_{u\in\mathbb{R}_+}$ and $\hat{M} = \{M(B_t) : B \in \mathcal{B}(\mathbb{R}_+ \times \mathbb{R}_+)\}$ with $B_t = \{(t+u, z) : (u, z) \in B\}$. Note that $\hat{W}$ and $\hat{M}$ are independent of $\mathcal{F}_t$ and of each other, $\hat{W}$ is Wiener just as W, and $\hat{M}$ is Poisson just as M. Finally, the uniqueness shown in Theorem 3.4 implies that $\hat{X}$ is obtained from $(\hat{X}_0, \hat{W}, \hat{M})$ by the same mechanism as X is obtained from (X_0, W, M). Hence, the claim. □

The strong Markov property is shown next; the wording repeats Theorem 2.19; the conventions 2.18 are in force regarding the end of time.

3.9 THEOREM. *The process X is strong Markov: For every $\mathcal{F}$-stopping time T, the variable X_T is $\mathcal{F}_T$-measurable, and $\hat{X} = (X_{T+u})_{u\in\mathbb{R}_+}$ is conditionally independent of $\mathcal{F}_T$ given X_T; moreover, for y in E, on the event $\{X_T = y\}$, the conditional law of $\hat{X}$ given X_T is the same as the law of X under $\mathbb{P}^y$.*

Proof. The measurability claimed for X_T is via Theorem V.1.14 and the right-continuity of X, and the adaptedness to $\mathcal{F}$. On $\{T = +\infty\}$, the claims regarding the conditional law given $\mathcal{F}_T$ are immediate from the conventions 2.18. On $\{T < \infty\}$, the claims are proved as in the preceding proof: replace t with T, recall that $\hat{W}$ is again Wiener by the strong Markov property for W (see Theorem VII.3.10), and use the next lemma.

3.10 LEMMA. *Let T be an $\mathcal{F}$-stopping time. For ω in $\{T < \infty\}$, define $\hat{M}(\omega, B) = M(\omega, B_{T(\omega)})$, where $B_t = (t, 0) + B$, $B \in \mathcal{B}(\mathbb{R}_+ \times \mathbb{R}_+)$. Then, on $\{T < \infty\}$, the conditional law of $\hat{M}$ given $\mathcal{F}_T$ is the law of the standard Poisson random measure M.*

Proof. Define $h(v) = v^{-2}$ for $v > 0$. Note that the lebesgue integral of $h \wedge 1$ over $\mathbb{R}_+$ is equal to 2, a finite number. It follows, since M is Poisson, adapted to $\mathcal{F}$, and with mean Leb $\times$ Leb on $\mathbb{R}_+ \times \mathbb{R}_+$, that

$$Z_t = \int_{[0,t]\times(0,\infty)} M\,(ds,\ dv)\ \ h(v)$$

defines an increasing pure-jump Lévy process adapted to $\mathcal{F}$; in fact, Z is stable with index 1/2. By Theorem VII.3.10, then, Z is strong Markov: on the event $\{T < \infty\}$, the conditional law of $\hat{Z} = (Z_{T+u} - Z_T)_{u\in\mathbb{R}_+}$ given $\mathcal{F}_T$ is the same as the law of Z.

Since h is a homeomorphism of $(0, \infty)$ onto $(0, \infty)$, the measure M_ω and the path $Z(\omega)$ determine each other for almost every ω. Similarly, on $\{T < \infty\}$, the measure $\hat{M}_\omega$ and the path $\hat{Z}(\omega)$ determine each other. Obviously, $\hat{M}$ bears the same relationship to $\hat{Z}$, as M does to Z. Thus, the claim follows from the strong Markov property of Z. □

Lévy kernel for jumps

This is a transition kernel from $(E, \mathcal{E})$ into $(E, \mathcal{E})$. It gives the rates and effects of jumps. It is defined by, for x in E and B in $\mathcal{E}$,

3.11 $$L\,(x, B) = \text{Leb}\,\{v \ge 0 :\ j(x, v) \ne 0,\ x + j(x, v) \in B\}\,.$$

Note that $L(x, \{x\}) = 0$. In a sense to be made precise by the next theorem, $L(x, B)$ is the rate of jumps from x into B per unit of time spent at x. If X were a Lévy process with a Lévy measure λ for its jumps, then $j(x, v)$ would be free of x, and $L(x, B)$ would be equal to $\lambda(B - x)$. Hence, the term "Lévy kernel" for L.

Heuristically, then, the "rate" of jumps when at x is

3.12 $$k(x) = L(x, E) = \text{Leb}\,\{v \ge 0 :\ j(x, v) \ne 0\}\,,\ x \in E.$$

Clearly, k is a positive Borel function on E. The general theory allows $k(x) = +\infty$ for some or for all x; Condition 3.2b implies that k is bounded by the constant c of 3.2b, and then, L is a bounded kernel. The next theorem does not assume 3.2b.

3.13 Theorem. *Let f be a positive Borel function on $E \times E$. Let F be a positive left-continuous process adapted to $\mathcal{F}$. Then, for every x in E,*

3.14 $$\mathbb{E}^x \sum_{s\in\mathbb{R}_+} F_s\ f \circ (X_{s-}, X_s)\ \ 1_{\{X_{s-}\ne X_s\}}$$
$$= \mathbb{E}^x \int_{\mathbb{R}_+} ds\ F_s \int_E L(X_s, dy)\ f \circ (X_s, y).$$

3.15 REMARKS. a) The proof below will show that this theorem holds without condition 3.2; all that is needed is that X be right-continuous and have left-limits in E, and that X satisfy the equation 3.1.

b) On the left side of 3.14, the sum is over the countably many times s of jumps. On the right side, we may replace X_s by X_{s-} since the integration over s will wash away the difference; the result with X_{s-} is closer to intuition.

c) Take F to be the indicator of $[0,\ t]$ for some fixed t, and let f be the indicator of a Borel rectangle $A \times B$. The sum on the left side of 3.14 becomes the number $N_t(A \times B)$ of jumps, during $[0,\ t]$, from somewhere in A to somewhere in B. Thus,

3.16 $$\mathbb{E}^x N_t(A \times B) = \mathbb{E}^x \int_0^t ds\,(1_A \circ X_{s-})\,L\,(X_{s-}, B)\,.$$

This is the precise meaning of the heuristic phrase that $L(y, B)$ is the rate of jumps from y into B. The phrase is short for the statement that $s \mapsto (1_A \circ X_{s-})L(X_{s-}, B)$ is the random intensity for the point process $s \mapsto N_s(A \times B)$ in the sense employed in Chapter IV, Section 6.

c) In particular, for the total number $N_t(E \times E)$ of jumps X makes during $[0,\ t]$, we see from 3.16 that

3.17 $$\mathbb{E}^x\ N_t(E \times E) = \mathbb{E}^x \int_0^t ds\ k \circ X_{s-} = \mathbb{E}^x \int_0^t ds\ k \circ X_s.$$

Proof. Since X satisfies 3.1, the sum on the left side of 3.14 is equal to

$$\int_{\mathbb{R}_+ \times \mathbb{R}_+} M(ds, dv)\ F_s\ \hat{f}\,(X_{s-}, X_{s-} + j(X_{s-}, v))$$

where $\hat{f}(y,\ z) = f(y,\ z)$ for $y \neq z$, and $\hat{f}(y,\ z) = 0$ for $y = z$. Here, the integrand is a process $(\omega, s, v) \mapsto G(\omega, s, v)$ that satisfies the predictability conditions of Theorem VI.6.2 on Poisson integrals: F is left-continuous, $s \mapsto X_{s-}$ is left-continuous, both are adapted to $\mathcal{F}$, and $\hat{f}$ and j are Borel. It follows from that theorem that the left side of 3.14 is equal to

$$\begin{aligned}\mathbb{E}^x &\int_{\mathbb{R}_+ \times \mathbb{R}_+} ds\ dv\ F_s\ \hat{f}\,(X_{s-}, X_{s-} + j\,(X_{s-}, v)) \\ &= \mathbb{E}^x \int_{\mathbb{R}_+} ds\ F_s \int_E L\,(X_{s-}, dy)\,f(X_{s-}, y),\end{aligned}$$

where we used the definition 3.11 of L to evaluate the integral over v. The last expression is equal to the right side of 3.14 since X_{s-} differs from X_s for only countably many s. □

Under Condition 3.2b, the kernel L is bounded, as $k(x) = L(x,\ E) \leq c$. We repeat this, and add a consequence:

3.18 Corollary. *The kernel L is bounded. For every bounded Borel function f on $E \times E$,*

$$m_t = \sum_{s \le t} f \circ (X_{s-}, X_s)\, 1_{\{X_{s-} \ne X_s\}} - \int_0^t ds \int_E L\,(X_s, dy)\ \ f(X_s, y),\ t \ge 0,$$

is an $\mathcal{F}$-martingale with $m_0 = 0$.

Proof. Fix $0 \le t < u$, fix f bounded and positive. For an event H in $\mathcal{F}_t$, put $F_s = 1_H\ 1_{(t,u]}(s)$. Then, F is left-continuous and adapted, and it follows from the preceding theorem that

$$\mathbb{E}^x\ 1_H \sum_{t<s\le u} f\,(X_{s-}, X_s)\, 1_{\{X_{s-} \ne X_s\}} = \mathbb{E}^x 1_H \int_t^u ds \int_E L(X_s, dy)\, f(X_s, y).$$

Since f is bounded, and L is a bounded kernel, the right side is real-valued; passing it to the left, we see that

$$\mathbb{E}^x\ 1_H \cdot (m_u - m_t) = 0.$$

That is, (m_t) is a martingale when f is bounded positive Borel. For f bounded Borel, the same conclusion holds obviously. □

Generator

Recall the Itô diffusion $\bar{X}$, which is the solution to 3.3 with $\bar{X}_0 = X_0$. Let $\bar{G}$ be its generator, that is, $\bar{G}f(x)$ is given by the right side of 2.21 for f in $\mathcal{C}_K^2 = C_K^2(E \mapsto \mathbb{R})$. We introduce (condition 3.2 is in force)

3.19 $$Gf(x) = \bar{G}f(x) + \int_E L\,(x, dy)\,[f(y) - f(x)]\,, \quad f \in \mathcal{C}_K^2.$$

This integro-differential operator is the *generator* for X:

3.20 Theorem. *For every f in $\mathcal{C}_K^2$,*

$$M_t = f \circ X_t - f \circ X_0 - \int_0^t ds\ Gf \circ X_s, \qquad t \in \mathbb{R}_+,$$

is an $\mathcal{F}$-martingale.

Proof. a) Fix f in $\mathcal{C}_K^2$. Put $T_0 = 0$, and let $T_1,\ T_2, \ldots$ be the successive jump times of X, defined recursively via $T_{n+1} = \inf\{t > T_n : X_{t-} \ne X_t\}$ with $n \ge 0$. On the event $\{T_n \le t < T_{n+1}\}$ consider the telescoping sum

3.21 $$\begin{aligned} f \circ X_t - f \circ X_0 &= \sum_{i=1}^n \left(f \circ X_{T_i} - f \circ X_{T_{i-}}\right) \\ &\quad + \sum_{i=1}^n \left(f \circ X_{T_{i-}} - f \circ X_{T_{i-1}}\right) + f \circ X_t - f \circ X_{T_n} \\ &= A + B + C. \end{aligned}$$

b) The term A is a sum over the jump times during $[0,t]$. By Corollary 3.18,

3.22 $$A = \sum_{s \le t} (f \circ X_s - f \circ X_{s-}) = m_t + \int_0^t ds \int_E L(X_s, dy)\,[f(y) - f(X_s)]\,.$$

c) We now prepare to evaluate $B + C$. Let S and T be stopping times, chosen so that, on the event $\{T < \infty\}$, we have $S < T$ and X continuous over the interval $(S,\ T)$. Thus, on $\{T < \infty\}$, the process X coincides over $(S,\ T)$ with some diffusion satisfying 3.3; and, since that diffusion has the generator $\bar{G}$,

3.23 $$f \circ X_{T-} - f \circ X_S = \bar{m}_T - \bar{m}_S + \int_S^T ds\ \bar{G} f \circ X_s,$$

by Theorem 2.23, with $\bar{m}$ as the martingale on the right side of 2.24.

d) Apply 3.23 repeatedly, with $S = T_{i-1}$ and $T = T_i$ for $i = 1, \dots, n$, and with $S = T_n$ and $T = t$. We see that, on the event $\{T_n \le t < T_{n+1}\}$

3.24 $$B + C = \bar{m}_t + \int_0^t ds\ \bar{G} f \circ X_s.$$

Finally, put 3.22 and 3.24 into 3.21, recall the definition 3.19 of G, and put $M = m + \bar{m}$. The result is the claim, since the union of the events $\{T_n \le t < T_{n+1}\}$ is the event $\{\lim\ T_n = +\infty\}$, and the latter event is almost sure in view of Condition 3.2b. □

Transition function, forward equation

The transition function (P_t) for X is defined, as usual, by

3.25 $$P_t f(x) = \mathbb{E}^x f \circ X_t, \qquad x \in E,\ f \in \mathcal{E}_+.$$

It follows from the preceding theorem that, for f in $\mathcal{C}_K^2$,

3.26 $$P_t f(x) = f(x) + \mathbb{E}^x \int_0^t ds\ G f \circ X_s = f(x) + \int_0^t ds\ P_s G f(x),$$

where the interchange of expectation and integration is justified by noting that Gf is bounded: For f in $\mathcal{C}_K^2$, Corollary 2.25 shows that $\bar{G}f$ is bounded continuous, and L is a bounded kernel under the standing condition 3.2.

The equation 3.26 is the integrated form of Kolmogorov's forward equation; see Corollary 2.26 and Theorem 2.30 for the diffusion case. Indeed, a formal differentiation of 3.26 yields the counterpart of the differential equations 2.31. The differentiability here, however, requires some continuity for the jump function j (in addition to Lipschitz continuity for the velocity fields u_n).

3.27 THEOREM. *Suppose that $x \mapsto j(x,v)$ is continuous for every v in $\mathbb{R}_+$. Then, for $f \in \mathcal{C}_K^2$, Gf is bounded and continuous, and*

3.28
$$\frac{d}{dt} P_t f = P_t G f.$$

Proof. Fix f in $\mathcal{C}_K^2$. As mentioned in Corollary 2.25, then, $\bar{G} f \in \mathcal{C}_K$ and thus bounded continuous. On the other hand, by the definition of L, and by 3.2b,

3.29
$$\int_E L(x,dy)\,[f(y)-f(x)] = \int_0^c dv \quad [f(x+j(x,v))-f(x)].$$

Since f is bounded continuous, and $x \mapsto j(x,v)$ is continuous by assumption, the integral on the right side yields a bounded continuous function (via the bounded convergence theorem). Adding 3.29 to $\bar{G}f$, we see that Gf is bounded continuous.

Consequently, by the right-continuity of X,

$$\lim_{s\to 0} P_{t+s} G f(x) = \lim_{s\to 0} \mathbb{E}^x \; Gf \circ X_{t+s} = \mathbb{E}^x \; Gf \circ X_t = P_t G f(x).$$

Hence, with the help of 3.26, we get

$$\lim_{u\to 0} \frac{1}{u} \left[P_{t+u} f(x) - P_t f(x)\right] = \lim_{u\to 0} \frac{1}{u} \int_0^u ds \; P_{t+s} G f(x) = P_t G f(x). \qquad \square$$

The first jump time

We return to the master equation 3.1. Define R to be the time of first jump:

3.30
$$R = \inf\{t > 0 : X_{t-} \neq X_t\}.$$

We show next that R is the lifetime of the diffusion $\bar{X}$ killed at the rate $k(x)$ when at x; recall $\bar{X}$ of 3.3 with $\bar{X}_0 = X_0$, and recall the notation $\exp_- x$ for e^{-x}.

3.31 LEMMA. $\mathbb{P}\{R > t | \bar{X}\} = \exp_- \int_0^t ds \; k \circ \bar{X}_s, \; t \in \mathbb{R}_+$.

Proof. Pick an outcome ω. Note that $R(\omega) > t$ if and only if $X_s(\omega) = \bar{X}_s(\omega)$ for all $s \le t$, which is in turn equivalent to having

$$M(\omega, D_\omega) = 0 \quad \text{for} \quad D_\omega = \left\{(s,v) \in \mathbb{R}_+ \times \mathbb{R}_+ : s \le t, j\left(\bar{X}_{s-}(\omega), v\right) \neq 0\right\}.$$

The diffusion $\bar{X}$ is determined by W, and M is independent of W. Thus, since M is Poisson with mean $\mu = \text{Leb} \times \text{Leb}$,

$$\mathbb{P}\left\{R > t | \bar{X}\right\} = \mathbb{P}\left\{M(D) = 0 | \bar{X}\right\} = e^{-\mu(D)}.$$

Finally, it follows from the definition 3.12 of k that

$$\mu(D_\omega) = \int_0^t ds \; k \circ \bar{X}_s(\omega). \qquad \square$$

3.32 REMARK. The preceding lemma is without conditions on the jump function. As a result, all values in $[0, \infty]$ are possible for R. When Condition 3.2b is in force, k is bounded and, thus, $R > 0$ almost surely. But R can take $+\infty$ as a value, that is, it is possible that there are no jumps; see Exercise 3.88 for an example.

3.33 PROPOSITION. *Let $x \in E$, $f \in \mathcal{E}_+$, $t \in \mathbb{R}_+$. Then,*

$$\mathbb{E}^x f \circ X_t \, 1_{\{R>t\}} = \mathbb{E}^x \; f \circ \bar{X}_t \; \exp_- \int_0^t ds \; k \circ \bar{X}_s,$$

where $\bar{X}$ is the diffusion that is the solution to 3.3 with $\bar{X}_0 = X_0 = x$.

Proof. On the left side, we may replace X_t with $\bar{X}_t$, since they are the same on the event $\{R > t\}$. Now, conditioning on $\bar{X}$ and applying the last lemma yield the claim. □

The preceding proposition establishes a connection to the Feynman-Kac formula discussed earlier. Define

3.34 $$\hat{P}_t f(x) = \mathbb{E}^x \; f \circ \bar{X}_t \; \exp_- \int_0^t ds \; k \circ \bar{X}_s, \quad x \in E, \; f \in \mathcal{E}_+, \; t \in \mathbb{R}_+,$$

which is the right side of the formula in the preceding proposition. Then, $(\hat{P}_t)$ is the sub-Markov transition semigroup of the Markov process $\hat{X}$ obtained from the diffusion $\bar{X}$ by killing the latter at the rate $k(x)$ when at x. See 2.42–2.50 for these matters, the semigroup, computational results for it, and the associated resolvant and generator. At this point, we regard $(\hat{P}_t)$ as known.

Regeneration at R

Heuristically, R is the killing time of $\bar{X}$. On the event that the killing succeeds, it occurs at the location $\bar{X}_R = X_{R-}$, and a new diffusion is born at the point X_R. We think of R as the time of first *regeneration* for X.

3.35 THEOREM. *For every x in E,*

$$\mathbb{P}^x\{R \in ds, \; X_{R-} \in dy, \; X_R \in dz\} = ds \; \hat{P}_s\,(x, dy)\, L(y, dz), \quad s \in \mathbb{R}_+, \; y \in E, \; z \in E.$$

Proof. The claim is equivalent to the more precise claim that

3.36 $$\mathbb{E}^x \; g \circ R \; f \circ (X_{R-}, X_R) = \int_{\mathbb{R}_+} ds \; g(s) \int_E \hat{P}_s\,(x, dy) \int_E L(y, dz) \; f(y, z)$$

for f positive Borel on $E \times E$ and g positive continuous and with compact support in $\mathbb{R}_+$. Fix f and g such. Let $F_s = g(s) \; 1_{\{s \le R\}}$ in Theorem 3.13. On the left side of 3.14, the sum consists of a single term, namely, $g \circ R$ $f \circ (X_{R-}, X_R)$; this is because the only jump time s in $[0, R(\omega)]$ is at $s = R(\omega)$

if $R(\omega) < \infty$, and $g \circ R(\omega) = 0$ if $R(\omega) = +\infty$ since g has compact support. Hence, 3.14 becomes

$$\mathbb{E}^x\ g \circ R\ f \circ (X_{R-}, X_R) = \mathbb{E}^x \int_{\mathbb{R}+} ds\ g(s)\ 1_{\{s \le R\}} \int_E L\,(X_s, dz)\ \ f(X_s, z).$$

On the right side, we may replace $\{s \le R\}$ with $\{R > s\}$ without altering the integral. The result is the right side of 3.36 in view of Proposition 3.33 and the definition 3.34. □

3.37 Remark. With k as defined by 3.12, we let K be a Markov kernel on $(E, \mathcal{E})$ that satisfies

$$L(x, B) = k(x)\ K\,(x, B)\,, \qquad x \in E,\ B \in \mathcal{E}.$$

If $k(x) > 0$, then this defines $K(x, \cdot)$ uniquely; when $k(x) = 0$, it matters little how $K(x, \cdot)$ is defined, we choose $K(x, \{x\}) = 1$ in that case, note that $K(x, \{x\}) = 0$ if $k(x) > 0$. Replacing $L(y, dz)$ with $k(y)\ K(y,\ dz)$ yields the following heuristic explanation of the preceding theorem: Suppose that the particle is started at x. It survives until t and moves as a diffusion to arrive at dy; this has probability $\hat{P}_t(x, dy)$. Then, it gets killed during dt; this has probability $k(y)\ dt$. Finally, it is reborn in dz; this has probability $K(y, dz)$.

The process at its jumps

This is to describe X at its successive jumps. We do it under the standing condition 3.2, although much of this requires less.

Put $T_0 = 0$ and let $T_1,\ T_2, \ldots$ be the successive jump times, that is,

3.38 $$T_{n+1} = \inf\{t > T_n : X_{t-} \ne X_t\}\,, \qquad n \in \mathbb{N}.$$

we have $T_1 = R$ as in 3.30. Condition 3.2b implies that $R > 0$ almost surely, which implies, through the strong Markov property at T_n, that $T_{n+1} > T_n$ almost surely on $\{T_n < \infty\}$. Moreover, as the construction in Theorem 3.4 makes clear,

3.39 $$\lim\ T_n = +\infty \text{ almost surely;}$$

in other words, for almost every ω, for every t in $\mathbb{R}_+$ there is n (depending on t and ω) such that $T_n(\omega) \le t < T_{n+1}(\omega)$; it is possible that $T_{n+1}(\omega) = \infty$. Finally, with the conventions in 2.18, we define

3.40 $$Y_n = X_{T_n}, \qquad n \in \mathbb{N}.$$

The strong Markov property at the stopping times T_n implies that $(Y, T) = (Y_n, T_n)_{n \in \mathbb{N}}$ is a Markov chain with state space $\bar{E} \times \bar{\mathbb{R}}_+$. It has a special structure: For A in $\mathcal{E}$ and B in $\mathcal{B}_{\mathbb{R}+}$,

3.41 $$\mathbb{P}\{Y_{n+1} \in A,\ T_{n+1} - T_n \in B\,|\,\mathcal{F}_{T_n}\} = Q\,(Y_n,\ A \times B)\,,$$

that is, given (Y_n), the conditional law of (T_n) is that of an increasing process with independent increments. The process (Y, T) is called a *Markov renewal chain*, and the times T_n are said to be regeneration times for X.

The kernel Q is specified by Theorem 3.35: take $n = 0$ in 3.41 and recall that $Y_0 = X_0$ and $T_1 - T_0 = R$. So,

3.42 $$Q(x, A \times B) = \int_B ds \int_E \hat{P}_s(x,\, dy)\, L\,(y,\, A) = \int_B\, ds\, \hat{P}_s L\,(x, A)\,.$$

It specifies the finite dimensional distributions of (Y, T) via 3.41. In particular, we have the iterative formula

3.43 $$\begin{aligned} Q^n(x, A \times B) &= \mathbb{P}^x \{Y_n \in A, T_n \in B\} \\ &= \int_{E\times\mathbb{R}_+} Q\,(x, dy, ds)\, Q^{n-1}\,(y, A \times (B - s)) \end{aligned}$$

for $n \geq 1$, and obviously, $Q^1 = Q$ and $Q^0(x, A \times B) = I(x, A)\delta_0(B)$. Computationally, in terms of the Laplace transforms

3.44 $$Q_p^n\,(x, A) = \int_{\mathbb{R}_+} Q^n\,(x, A \times ds)\, e^{-ps} = \mathbb{E}^x\; e^{-pT_n} 1_A \circ Y_n,$$

we see from 3.43 and 3.42 that $Q_p^1 = Q_p$, and $Q_p^\circ = I$, and

3.45 $$Q_p = \hat{U}_p L, \qquad Q_p^n = (Q_p)^n, \qquad n \in \mathbb{N},\ p \in \mathbb{R}_+,$$

where $(\hat{U}_p)$ is the resolvent of the semigroup $(\hat{P}_t)$.

Transition function

This is to give an explicit formula for the transition function (P_t) in terms of the known objects $\hat{P}_t$, L, and Q^n. The result is made possible by 3.39, that is, by the fact (guaranteed by 3.2b) that there can be at most finitely many jumps during $[0,\, t]$.

3.46 Theorem. *Let $x \in E,\ f \in \mathcal{E}_+,\ t \in \mathbb{R}_+$. Then,*

$$P_t f(x) = \sum_{n=0}^{\infty} \int_{E\times[0,t]} Q^n(x, dy, ds)\ \hat{P}_{t-s} f(y).$$

Proof. By the strong Markov property at T_n,

$$\begin{aligned} &\mathbb{E}^x\ f \circ X_t\ 1_{\{T_n \leq t < T_{n+1}\}} \\ &= \int_{E\times[0,t]} \mathbb{P}^x \{Y_n \in dy, T_n \in ds\}\, \mathbb{E}^y f \circ X_{t-s}\ 1_{\{R > t-s\}} \\ &= \int_{E\times[0,t]} Q^n(x, dy, ds)\ \hat{P}_{t-s} f(y), \end{aligned}$$

where we used the definition of Q^n in 3.43, and Proposition 3.33 for the meaning of $\hat{P}_t$. Summing both sides over n in $\mathbb{N}$ yields the claimed formula since $\lim T_n = \infty$; see 3.39 *et seq.* □

The formula in the preceding theorem gives the unique solution of the integro-differential equation 3.28 for (P_t). The uniqueness follows from 3.39. By avoiding generators, we have also avoided the continuity condition on j used in Theorem 3.27.

Nevertheless, it may be useful to characterize (P_t) as the unique solution to something resembling the backward equations.

3.47 THEOREM. *Let $x \in E,\ f \in \mathcal{E}_+,\ t \in \mathbb{R}_+$. Then,*

$$P_t f(x) = \hat{P}_t f(x) + \int_0^t ds\ \hat{P}_s L P_{t-s} f(x).$$

Proof. We use the so-called renewal argument at the time R of first jump:

$$\begin{aligned} P_t f(x) &= \mathbb{E}^x f \circ X_t \cdot 1_{\{R>t\}} + \mathbb{E}^x f \circ X_t\ 1_{\{R\le t\}} \\ &= \hat{P}_t f(x) + \int_{[0,t]\times E} \mathbb{P}^x\{R \in ds, X_R \in dz\}\, \mathbb{E}^z\ f \circ X_{t-s} \\ &= \hat{P}_t f(x) + \int_0^t ds \int_E \hat{P}_s(x,dy) \int_E L(y,dz) P_{t-s} f(z) \end{aligned}$$

where we used the strong Markov property at R followed by the distribution provided by Theorem 3.35. □

Resolvent

Let (U_p) be the resolvent of (P_t), defined by 2.34, but for the present X and (P_t). Taking Laplace transforms on both sides of 3.26, assuming the same conditions hold, we get

3.48 $$p\, U_p f = f + U_p\, G f, \qquad t \in \mathcal{C}_K^2.$$

with the generator as defined by 3.19. It is usual to write this in the form $U_p(p-G)f = f$, thus emphasizing that U_p is the inverse of $p - G$.

We can avoid differentials by using the probabilistic derivations for (P_t). With $Q_p = \hat{U}_p L$ as in 3.45, we see from Theorems 3.46 and 3.47 that

3.49 $$U_p f\ = \sum_{n=0}^{\infty} (Q_p)^n\, \hat{U}_p f, \qquad f \in \mathcal{E}_+,$$

3.50 $$U_p f\ = \hat{U}_p\ f + Q_p\ U_p f, \qquad f \in \mathcal{E}_+.$$

Indeed, 3.49 is the unique solution to the integral equation 3.50.

3.51 PROPOSITION. *Let $p > 0$ and $f \in \mathcal{E}_{b+}$. Then, $g = U_p f$ given by 3.49 is the unique bounded solution to*

$$g = \hat{U}_p f + Q_p g.$$

Proof. Replace g on the right side with $\hat{U}_p f + Q_p g$ repeatedly. With the notation of 3.45, we get

$$\begin{aligned} g &= \hat{U}_p f + Q_p\, g \\ &= \hat{U}_p f + Q_p \hat{U}_p f + Q_p^2\, g = \cdots = \left(I + Q_p + \cdots + Q_p^n\right) \hat{U}_p f + Q_p^{n+1}\, g. \end{aligned}$$

In the last member, the first term is increasing to $U_p f$ given by 3.49. Thus, there remains to show that

$$\lim_{n\to\infty} Q_p^n\, g = 0.$$

for every g bounded positive, say, bounded by b. But, then,

$$Q_p^n\, g(x) \le b\; Q_p^n(x, E) = b\; \mathbb{E}^x\; e^{-pT_n} \to 0$$

as $n \to \infty$, because $p > 0$ and $T_n \to \infty$ almost surely. □

Simple step processes

These are pure-jump processes obtained by setting the vector fields $u_0, \ldots, u_m$ equal to zero, and keeping the condition 3.2b on the jump function j. Then, the Wiener processes W^n have no rôle to play, and the diffusion $\bar{X}$ satisfying 3.3 becomes motionless: $\bar{X}_t = X_0$ for all t. Thus, the process X is a right-continuous step process with a bounded Lévy kernel. In the next subsection, we shall discuss dropping the boundedness condition on the Lévy kernel and thus weakening the condition 3.2b.

The Markov and strong Markov properties remain unchanged. Theorem 3.13 on the Lévy kernel L stays the same, as is Corollary 3.18. The generator G is simpler: since $\bar{X}$ is motionless, $\bar{G}$ disappears;

3.52 $$Gf(x) = \int_E L\,(x, dy)\,[f(y) - f(x)]\,, \qquad f \in \mathcal{E}_b.$$

Theorem 3.20 becomes stronger; the claim holds for every f in $\mathcal{E}_b$. Similarly, 3.26 holds for every f in $\mathcal{E}_b$. Theorem 3.27 is stronger:

3.53 $$\frac{d}{dt} P_t f = P_t G f, \qquad f \in \mathcal{E}_b,$$

without the continuity assumption on j. That assumption was used in the proof to show that $s \mapsto Gf \circ X_s$ is bounded and right-continuous; we have boundedness via 3.52 and the boundedness of f and L; and $s \mapsto Gf \circ X_s$

is right-continuous because $s \mapsto X_s$ is a right-continuous step function (and thus $s \mapsto g \circ X_s$ is a right-continuous step function for arbitrary g).

Since $\bar{X}$ is motionless, killing it becomes simpler. For the time R of first jump, since $X_t = \bar{X}_t = X_0$ on $\{R > t\}$, Lemma 3.31 and 3.33 and 3.34 become

$$\mathbb{P}^x \{R > t\} = e^{-k(x)t}, \qquad \hat{P}_t f(x) = \mathbb{E}^x f \circ X_t \ 1_{\{R>t\}} = e^{-k(x)t} \ f(x),$$

and Theorem 3.35 becomes elementary (we drop $X_{R-} = X_0$ and use the notation of Remark 3.37)

$$\mathbb{P}^x \{R \in ds, X_R \in dy\} = (ds \ k(x) \ e^{-k(x)s})K(x, dy).$$

3.54 *Heuristics.* Suppose that the initial state is x. The particle stays there an exponential amount of time with parameter $k(x)$, and, independent of that amount, jumps to a new point y with probability $K(x, dy)$; then, it has a sojourn at y of exponential duration with parameter $k(y)$, followed by a jump to a new point chosen in accord with the probability law $K(y, \cdot)$; and so on. It is possible that, somewhere along its path, the particle lands at a point z with $k(z) = 0$; that z is a trap, and the particle stays there forever after.

For the transition function (P_t) and the resolvent (U_p), it is possible to give explicit and easy to interpret formulas. Heuristically, instead of the jumps of X, the idea is to concentrate on the kicks by the atoms of the Poisson M. Some kicks cause jumps, some not. If the particle is at x when it is kicked by an atom (s, v), it jumps to $x + j(x, v)$ if $j(x, v) \neq 0$, and it stays put if $j(x, v) = 0$. Following this reasoning as in the proof of Theorem 3.4, we obtain (with c as a bound for k)

3.55 $$P_t(x, A) = \sum_{n=0}^{\infty} \frac{e^{-ct}(ct^n)}{n!} Q^n(x, A), \qquad t \in \mathbb{R}_+, \ x \in E, \ A \in \mathcal{E},$$

where Q^n is the n^{th} power of the Markov kernel Q on $(E, \mathcal{E})$ given by

3.56 $$Q(x, A) = \left(1 - \frac{k(x)}{c}\right) I(x, A) + \frac{k(x)}{c} \ K(x, A).$$

Thus, X has the form of a Markov chain subordinated to a Poisson process; see 1.26. The corresponding resolvent is

3.57 $$U_p = \frac{1}{c+p} \sum_{n=0}^{\infty} \left(\frac{c}{c+p}\right)^n \ Q^n, \qquad p > 0.$$

Step processes and extensions

We continue with the vector fields u_n set to zero, and we weaken condition 3.2b: instead of assuming that the Lévy kernel is bounded, we shall assume only that it is finite. The classical examples are the processes with discrete state spaces.

The process X of interest has state space $E = \mathbb{R}^d$ as before. It is adapted to the filtration $\mathcal{F}$, it is right-continuous and has left-limits in E, and it satisfies

3.58 $$X_t = X_0 + \int_{[0,t]\times\mathbb{R}_+} M\,(ds, dv)\;\; j(X_{s-}, v),$$

with the same Poisson random measure M as before. Theorem 3.13 remains true (as remarked in 3.15a) with the Lévy kernel L defined by 3.11. We assume throughout that the following holds.

3.59 Condition. The Lévy kernel L is finite.

In other words, $k(x) = L(x, E) < \infty$ for every x in E. Further, we may assume that $j(x, v) \neq 0$ for $0 \leq v \leq k(x)$ only. Then, 3.58 is easier to visualize; see Exercise 3.84 also.

The process X has well-defined times $T_1, T_2, \ldots$ of the first jump, the second jump, ... It is a step process if and only if

3.60 $$T_\alpha = \lim_n \; T_n$$

is almost surely equal to $+\infty$. Otherwise, $t \mapsto X_t$ is a step function only over the interval $[0, T_\alpha)$. In either case, the evolution of X over $[0, T_\alpha)$ is as described in 3.54. The following two examples are instructive; see Exercises 3.90 and 3.91 as well.

3.61 Example. *Upward staircase.* Take $E = \mathbb{R}$. Let $D = \{x_0, x_1, \ldots\}$ where $0 = x_0 < x_1 < \cdots$ and $\lim x_n = 1$. Let $k(x) \in (0, \infty)$ for each x in D, and put $k(1) = 0$. If $x = x_n$ for some n and $v \leq k(x_n)$, then put $j(x, v) = x_{n+1} - x_n$; put $j(x, v) = 0$ for all other x and v.

If $X_0 = x_0 = 0$, then X stays at x_0 until T_1 and jumps to x_1, stay at x_1 until T_2 and jumps to x_2, and so on. The sojourn lengths $T_1,\ T_2 - T_1, \ldots$ are independent exponential variables with respective parameters $k(x_0),\ k(x_1), \ldots$. Their sum is the variable T_α defined by 3.60. So,

3.62 $$\mathbb{E}^0\; T_\alpha = \sum_{x\in D} \frac{1}{k(x)}.$$

If $\mathbb{E}^0 T_\alpha < \infty$, then $T_\alpha < \infty$ almost surely, and we let $X_t = 1$ for $t \geq T_\alpha$. We show next that, if $\mathbb{E}^0 T_\alpha = +\infty$, then $T_\alpha = +\infty$ almost surely and $X_t \in D$ for all t. In either case, we call X a staircase over $[0, T_\alpha)$ with steps at $x_0, x_1, \ldots$.

For the main computation, we use 3.16. For fixed n, let A consist of x_n, and B of x_{n+1}; then $N_t(A \times B)$ becomes the indicator of $\{X_t > x_n\}$, and 3.16 yields

3.63 $$\mathbb{P}^0\,\{X_t > x_n\} = \int_0^t \; ds\; \mathbb{P}^0\,\{X_s = x_n\}\;\; k(x_n).$$

Pass the factor $k(x_n)$ to the left side and note that $\{T_\alpha \le t\} \subset \{X_t > x_n\}$. Thus,

$$\frac{1}{k(x_n)} \mathbb{P}^0 \{T_\alpha \le t\} \le \int_0^t ds \; \mathbb{P}^0 \{X_s = x_n\}$$

Sum both sides over all n, note 3.62, and note that the sum of the right side is at most t. We get

$$\mathbb{E}^0 T_\alpha \; \mathbb{P}^0 \{T_\alpha \le t\} \le t.$$

We conclude that, if $\mathbb{E}^0 T_\alpha = +\infty$ then $\mathbb{P}^0\{T_\alpha \le t\} = 0$ for all t, which means that $T_\alpha = +\infty$ almost surely.

We re-state the essential content of the preceding example:

3.64 Lemma. *Let S be the sum of a countable independency of exponentially distributed random variables. If $\mathbb{E}S < \infty$ then $S < \infty$ almost surely; if $\mathbb{E}S = +\infty$ then $S = +\infty$ almost surely.*

3.65 Example. Let E, D, k, j be as in the last example, but with

3.66
$$\sum_{x \in D} \frac{1}{k(x)} = 1.$$

Let μ be a probability measure on D. We now describe a process that is a concatenation of staircases.

Started at $x = x_i$ for some i, the process is a staircase over $[0, T_\alpha)$ with steps at $x_i,\ x_{i+1}, \ldots$; in view of 3.66, we have $\mathbb{E}^x T_\alpha \le 1$, and thus $T_\alpha < \infty$ almost surely. At T_α, we deviate from Example 3.61: we choose the random variable X_{T_α} independent of $\mathcal{F}_{T_\alpha}$ according to the distribution μ on D. If X_{T_α} turns out to be x_j, we let X form a staircase over $[T_\alpha, T_{2\alpha})$ with steps at $x_j, x_{j+1}, \ldots$; note that $\mathbb{E}^x T_{2\alpha} \le 2$ and thus $T_{2\alpha} < \infty$ almost surely. We choose $X_{T_{2\alpha}}$ independent of $\mathcal{F}_{T_{2\alpha}}$ and with distribution μ again, and proceed to form another staircase over $[T_{2\alpha}, T_{3\alpha})$. And we repeat this over and over. The result is a process whose jump times can be ordered as

3.67 $\quad T_1, T_2, \ldots, T_\alpha; T_{\alpha+1}, T_{\alpha+2}, \ldots, T_{2\alpha}; T_{2\alpha+1}, T_{2\alpha+2}, \ldots, T_{3\alpha}; \ldots.$

Each $T_{n\alpha}$ is the limit of a strictly increasing sequence of jump times; at each $T_{n\alpha}$ the process jumps from its left-limit 1 to its right-hand value $X_{T_{n\alpha}}$, the latter being independent of $\mathcal{F}_{T_{n\alpha}}$ and having the distribution μ. It follows that $T_{2\alpha} - T_\alpha, T_{3\alpha} - T_{2\alpha}, \ldots$ are independent and identically distributed, and, hence, $\lim_n T_{n\alpha} = +\infty$ almost surely. So, X_t is well-defined for every t in $\mathbb{R}_+$; the process X is right-continuous, is left-limited (as a process with state space E), and satisfies 3.58. Incidentally, this example shows that 3.58 can have many solutions.

We resume the treatment of the process X of 3.58–3.59, concentrating on the transition semigroups (P_t) and (P_t^*), where

3.68
$$P_t f(x) = \mathbb{E}^x \; f \circ X_t, \qquad P_t^* f(x) = \mathbb{E}^x \; f \circ X_t \; 1_{\{T_\alpha > t\}}.$$

Recall that for the time $R = T_1$ of first jump, we have

3.69 $Q(x, dy, ds) = \mathbb{P}^x \{X_R \in dy, R \in ds\} = ds\ k(x)e^{-k(x)s}\ K(x, dy),$

as with simple step processes; and $k(x) < \infty$ by assumption. Also as before, T_n is the time of n^{th} jump and $Y_n = X_{T_n}$ for $n \in \mathbb{N}$, and

3.70 $Q^n(x, dy, ds) = \mathbb{P}^x \{Y_n \in dy, T_n \in ds\}, \qquad x, y \in E, s \in \mathbb{R}_+.$

Obviously, $Q^\circ(x, dy, ds) = I(x, dy)\,\delta_0(ds)$ and $Q^1 = Q$, and Q^n can be computed recursively via 3.43.

3.71 Proposition. *Let $x \in E,\ f \in \mathcal{E}_+,\ t \in \mathbb{R}_+$. Then,*

$$P_t^* f(x) = \sum_{n=0}^{\infty} \int_{E \times [0,t]} Q^n(x, dy, ds)\ e^{-k(y)(t-s)} f(y).$$

Proof. This is essentially as in the proof of Theorem 3.46: On the set $\{T_\alpha > t\}$ we have $T_n \le t < T_{n+1}$ for some n in $\mathbb{N}$. Hence,

$$\begin{aligned} P_t^* f(x) &= \sum_{n=0}^{\infty} \mathbb{E}^x\ f \circ X_t\ 1_{\{T_n \le t < T_{n+1}\}} \\ &= \sum_{n=0}^{\infty} \int_{E \times [0,t]} \mathbb{P}^x \{Y_n \in dy,\ T_n \in ds\}\ f(y)\ \mathbb{P}^y \{R > t - s\}, \end{aligned}$$

which is the claim. □

If $T_\alpha = +\infty$ almost surely, then $P_t^* = P_t$ for all t; see 3.68. A simple criterion for ensuring this condition is obtained from Lemma 3.64: Since the sojourn lengths $T_1, T_2 - T_1, \ldots$ are conditionally independent given (Y_n) and are conditionally exponential with parameters $k \circ Y_0,\ k \circ Y_1, \ldots$, Lemma 3.64 applies to the conditional law of (T_n) given (Y_n). The result is put next.

3.72 Proposition. *If $\sum_n 1/k \circ Y_n = +\infty$ almost surely, then $T_\alpha = +\infty$ almost surely and $P_t = P_t^*$ for all t.*

The preceding proposition is effective in a number of special situations: If k is bounded, then $T_\alpha = +\infty$ almost surely. If there is a recurrent point x for the chain Y, that is, if

$$\mathbb{P}^x \{Y_n = x \text{ for infinitely many } n\} = 1,$$

then $k \circ Y_n = k(x)$ for infinitely many n, and hence $T_\alpha = +\infty$ almost surely under $\mathbb{P}^x$. Similarly, if there is a recurrent set A (which Y visits infinitely often) and if k is bounded on A, then $T_\alpha = +\infty$ almost surely and $P_t^* = P_t$.

The next proposition is a summary of easy observations and a criterion for deciding whether $T_\alpha = +\infty$ almost surely.

3.73 PROPOSITION. *The following are equivalent:*

a) X *is a step process, that is,* $T_\alpha = +\infty$ *almost surely.*

b) $P_t^* = P_t$ *for all* t.

c) There exists $\tau > 0$ *such that* $P_\tau^*(x, E) = 1$ *for all* x *in* E.

d) For some (and therefore all) $p > 0$, *the only solution to*

3.74 $$h = Q_p h, \qquad 0 \le h \le 1, \qquad h \in \mathcal{E},$$

is $h = 0$; *here* $Q_p(x, A) = \frac{k(x)}{k(x)+p} K(x, A)$; *see 3.69.*

Proof. Obviously, $(a) \Leftrightarrow (b) \Rightarrow (c)$. To see that $(c) \Rightarrow (a)$, fix $\tau > 0$ such, that is, $P_\tau^* 1 = 1$. Then, $P_{s+\tau}^* 1 = P_s^* P_\tau^* 1 = P_s^* 1$ for all s. Replacing s with $\tau, 2\tau, \ldots$ we see that $P_{n\tau}^* 1 = 1$ for every n, which means that $\mathbb{P}^x\{T_\alpha > n\tau\} = 1$ for all x and n. Hence, $T_\alpha = +\infty$ almost surely.

Finally, we show that $(d) \Leftrightarrow a$. It follows from 3.69, 3.70, 3.43 that, for fixed $p > 0$,

$$h^*(x) = \mathbb{E}^x \; e^{-pT_\alpha} = \lim_n \mathbb{E}^x e^{-pT_n} = \lim_n \; Q_p^n 1(x),$$

where $Q_p^n = (Q_p)^n$. Thus, $Q_p \; h^* = \lim \; Q_p^{n+1} 1 = h^*$, that is, h^* is a solution to 3.74. Moreover, it is the maximal solution to it: if h is a solution, then $h \le h^*$, since

$$h = Q_p^n h \le Q_p^n 1 \to h^*.$$

Hence, $h^* = 0$ if and only if $h = 0$ is the only solution to 3.74. This shows that $(d) \Leftrightarrow (a)$, since $h^* = 0$ if and only if $T_\alpha = +\infty$ almost surely. □

The next theorem lists the backward equations for the derivatives of (P_t) and (P_t^*). We re-introduce the generator G:

3.75 $$Gf(x) = \int_E L\,(x, dy)\,[f(y) - f(x)], \qquad x \in E, f \in \mathcal{E}_b.$$

3.76 THEOREM. *Let* $f \in \mathcal{E}_b$. *We have the backward equations*

$$\frac{d}{dt} \; P_t f = G P_t f, \qquad \frac{d}{dt} P_t^* f = G P_t^* f.$$

These equations have a unique solution (and $P_t f = P_t^* f$ *for all* t*) if and only if* X *is a step process.*

The proof will be given together with the proof of the following, more comprehensive, result on the backward equations in integral form.

3.77 Theorem. *Let $f \in \mathcal{E}_{b+}$. For bounded Borel $g : E \times \mathbb{R}_+ \mapsto \mathbb{R}_+$, consider the equation*

3.78 $$g(x,t) = e^{-k(x)t} f(x) + \int_0^t ds\, k(x)\, e^{-k(x)s} \int_E K(x,dy)\, g(y,t-s)$$

for x in E and t in $\mathbb{R}_+$. This equation holds for both g^o and g^*, where

$$g^o(x,t) = P_t f(x), \qquad g^*(x,t) = P_t^* f(x).$$

If X is a step process, then 3.78 has exactly one solution: $g = g^o = g^*$. Otherwise, the uniqueness fails, and g^* is the minimal solution.

3.79 Remark. If X is not a step process, the backward equation characterizes $P_t^* f$ as the minimal solution, but does not specify $P_t f$. For instance, for Example 3.65, there are as many solutions as there are choices of μ. See Exercise 3.91 for the computation of $P_t f$.

Proof of 3.76, assuming 3.77. We re-write 3.78:

$$g(x,t) = e^{-k(x)t} \Big[f(x) + \int_0^t ds\ e^{k(x)s} \int_E L(x,dy)\, g(y,s) \Big].$$

On the right side, since g is bounded Borel, the integration over E yields a bounded Borel function, and the integration over $[0, t]$ yields a continuous function in t. Thus, on the left, $t \mapsto g(x,t)$ must be bounded continuous. We put this back into the right side: since $s \mapsto g(y,s)$ is bounded continuous, the integration over E yields a bounded continuous function in s, and the integration over $[0, t]$ yields a differentiable function in t. So, $t \mapsto g(x,t)$ is differentiable. Taking derivatives, we get

$$\frac{\partial}{\partial t} g(x,t) = -k(x) g(x,t) + \int_E L(x,dy)\ g(y,t) = \int_E L(x,dy)\, [g(y,t) - g(x,t)]\,.$$

Assuming Theorem 3.77, then, the functions g^o and g^* must satisfy the preceding; hence the backward equations of Theorem 3.76. The other assertion, on uniqueness, is immediate from Theorem 3.77. □

Proof of Theorem 3.77

We start by showing that g^* satisfies 3.78; showing the same for g^o is similar and simpler. We use the strong Markov property at the time R of first jump. Since $R < T_\alpha$ and $X_t = X_0$ on $\{R > t\}$,

$$\begin{aligned} P_t^* f(x) &= \mathbb{E}^x\ f \circ X_t\ 1_{\{R>t\}} + \mathbb{E}^x\ f \circ X_t\ 1_{\{R \le t\}} 1_{\{T_\alpha > t\}} \\ &= e^{-k(x)t}\ f(x) + \int_{E \times [0,t]} Q(x,dy,ds)\, P_{t-s}^* f(y), \end{aligned}$$

where Q is as in 3.69. This is the same as 3.78 for g^*.

Consider the solutions g to 3.78. We employ Laplace transforms with Q_p as in 3.74 (see also 3.69, 3.70, 3.43) and

$$g_p(x) = \int_0^\infty dt\ e^{-pt}\ g(x,t), \qquad f_p(x) = \frac{1}{k(x)+p} f(x).$$

Then, 3.78 becomes

$$g_p = f_p + Q_p\ g_p.$$

Replace g_p on the right side with $f_p + Q_p g_p$, and repeat n times. We get

$$g_p = \left(I + Q_p + \cdots + Q_p^n\right) f_p + Q_p^{n+1} g_p.$$

Since f is positive, the first term on the right is increasing in n; the limit is, in view of 3.71,

$$\sum_{n=0}^\infty Q_p^n f_p = \int_0^\infty dt\ e^{-pt} P_t^*\ f$$

Hence, g^* is the minimal solution to 3.78. The uniqueness has to do with

$$h_p = \lim_n\ Q_p^n\ g_p.$$

We note that h_p is bounded, positive, and satisfies

$$h_p = Q_p\ h_p.$$

Thus, the remaining assertions of the theorem follow from Proposition 3.73. □

Forward equations are more sensitive to whether X is a steps process. The next theorem shows that (P_t^*) satisfies the forward equation, but (P_t) does not. We list a needed result first; see 3.70 and 3.71.

3.80 Lemma. $\sum_{n=1}^\infty\ Q^n(x,dy,ds) = ds\ P_s^* L(x,dy),\ y \in E,\ s \in \mathbb{R}_+.$

Proof. In view of 3.70, what we need to show can be stated more precisely as

$$\mathbb{E}^x \sum_{n=1}^\infty\ 1_A \circ Y_n\ 1_{\{T_n \le t\}} = \int_0^t ds \int_E P_s^*(x,dy) L(y,A).$$

The left side is the same as the left side of 3.14 with $F_s = 1_{\{s \le t \wedge T_\alpha\}}$ and $f(x,y) = 1_A(y)$. Thus, the left side is equal to

$$\begin{aligned}
&\mathbb{E}^x \int_0^\infty ds\ 1_{\{s \le t\wedge T_\alpha\}} \int_E L(X_s,dy)\ 1_A(y) \\
&= \mathbb{E}^x \int_0^t ds \quad 1_{\{T_\alpha > s\}} L(X_s,A) = \int_0^t ds \int_E P_s^*(x,dy) L(y,A),
\end{aligned}$$

where we used the definition of P_t^* in 3.68 for the last equality. □

3.81 THEOREM. *Let $f \in \mathcal{E}_{b_+}$ and let the generator G be as in 3.75. Then, $P_t^* f$ satisfies the equation*

$$\frac{d}{dt}\, P_t^* f = P_t^* G f$$

and is the minimal solution of it with $P_0^* f = f$. For (P_t) we have

$$\frac{d}{dt}\, P_t f \geq P_t G f;$$

the equality holds if and only if X is a step process, and, then, $P_t f = P_t^* f$.

Proof. Combining Proposition 3.71 and Lemma 3.80, we have

3.82 $$P_t^* f(x) = e^{-k(x)t} f(x) + \int_0^t ds \int_E P_s^* L(x, dy) e^{-k(y)(t-s)} f(y).$$

By Theorem 3.76, this is differentiable. Taking derivatives on both sides we obtain

$$\begin{aligned}\tfrac{d}{dt}\, P_t^* f(x) &= -k(x) e^{-k(x)t} f(x) + P_t^* L f(x) \\ &\quad - \textstyle\int_0^t ds\, P_s^* L(x, dy) e^{-k(y)(t-s)} k(y) f(y) \\ &= -k(x) e^{-k(x)t} f(x) + P_t^* L f(x) - \left[P_t^*(kf)(x) - e^{-k(x)t} k(x) f(x)\right]\end{aligned}$$

where we used 3.82 at the last step. Hence, we have

3.83 $$\frac{d}{dt} P_t^* f = P_t^* L f - P_t^*(kf) = P_t^* G f.$$

For $P_t f$, we have differentiability by Theorem 3.76. And,

$$P_{t+s}\, f - P_t f = P_t (P_s f - f) \geq P_t \left(P_s^* f - f\right)$$

by 3.68 and positivity of f. Thus, using the boundedness of f,

$$\frac{d}{dt} P_t f = \lim_{s \to 0} \frac{P_{t+s} f - P_t f}{s} \geq \lim_{s \to 0} P_t \frac{P_s^* f - f}{s} = P_t\, G f,$$

where we used 3.83 at the last step. The other assertions are repetitions of some claims in Theorem 3.77. □

Much of the foregoing are classical results for Markov processes with discrete state spaces. We have chosen the state space to be $E = \mathbb{R}^d$. Obviously, every countable set D with the discrete topology can be embedded in E, but our requirement of right-continuity for X leaves out an interesting class of processes which have discrete state spaces but permit discontinuities of the second kind; see the notes for this chapter.

Exercises and complements

3.84 *Lévy kernel and the jump function.* Let L be a finite kernel, that is, $k(x) = L(x, E) < \infty$ for every x in E, and take $E = \mathbb{R}$. Suppose that $L(x, \{x\}) = 0$ for each x and define $j(x, v)$ by

$$x + j(x, v) = \inf \{y \in \mathbb{R} \ : \ L(x, (-\infty, y]) > v\}$$

for v in $(0,k(x))$, and set $j(x, v) = 0$ for other v in $\mathbb{R}_+$.

a) Show that $y \mapsto L(x, (-\infty, y])$ and $v \mapsto x + j(x, v)$ are functional inverses of each other.

b) Show that L is the Lévy kernel defined from j by 3.11.

3.85 *Exponential decay with jumps.* Let $E = \mathbb{R}$, let $X_0 \in (0, \infty)$, and let X satisfy 3.1 with

$$a(x) = -cx, \qquad b(x) = 0, \qquad j(x, v) = xv\ 1_{(0,1)}(xv),$$

for $x > 0$. Note that X remains in $(0, \infty)$ forever. Plot the atoms of $M(\omega, \cdot)$ for a typical ω. Draw the path $t \mapsto X_t(\omega)$ corresponding to $M(\omega,\ \cdot)$ and with $X_0(\omega) = 1$.

3.86 *Continuation.* In the preceding exercise, replace j with

$$j(x, v) = j_0(v) 1_{(0,3)}(v)$$

for some increasing right-continuous function j_0 on (0,3). Describe the evolution of X with special attention to jump times, jump amounts, dependence and independence. Show that X satisfies the integral equation

$$X_t(\omega) = X_0(\omega) - c \int_0^t ds\ X_s(\omega) + Z_t(\omega), \qquad \omega \in \Omega, t \in \mathbb{R}_+,$$

where Z is a compound Poisson process. Solve this equation for X.

3.87 *Piecewise deterministic processes.* In 3.1, let $b = 0$, and let $a = u_0$ satisfy the Lipschitz condition 2.4, and j satisfy 3.2b. Then, between two consecutive jumps, the path $t \mapsto X_t(\omega)$ satisfies the ordinary differential equation

$$\frac{d}{dt} x_t = a(x_t),$$

whose solution is unique and deterministic given its initial condition. Show that Theorem 3.27 holds with the generator

$$Gf(x) = \sum_{i=1}^{d} a^i(x)\ \partial_i f(x) + \int_0^\infty\ dv \left[f(x + j(x, v)) - f(x) \right], \qquad f \in \mathcal{C}_K^1.$$

3.88 *Probably no jumps.* This is to give an example of X that has at most finitely many jumps. Suppose that $E = \mathbb{R}$,

$$a(x) = -1, \qquad b(x) = 1, \qquad j(x,v) = -1_{\mathbb{R}_+}(x)1_{(0,c)}(v).$$

Note that $\bar{X}$ of 3.3 is a Brownian motion with downward drift. Show (see Chapter V for this) that

$$A = \text{Leb}\left\{t \geq 0 : \bar{X}_t \in \mathbb{R}_+\right\}$$

is almost surely finite. Show that, for the time R of first jump,

$$\mathbb{P}^x\left\{R = +\infty | \bar{X}\right\} = e^{-cA}.$$

Conclude that $\mathbb{P}^x\{R = +\infty\} = \mathbb{E}^x e^{-cA}$ is strictly positive.

3.89 *Brownian motion plus jumps.* Let X be as in 3.1–3.2. Suppose that $E = \mathbb{R}$ and $\bar{X} = X_0 + W$, a standard Brownian motion. Define $j(x,v) = -ax$ for $0 \leq v \leq 1 - e^{-|x|}$, and $j(x,v) = 0$ otherwise, where a is a constant in (0,1).

a) Describe the motion X during $[0,R)$ and at R, where R is the time of first jump.
b) Specify the Lévy kernel L
c) Specify the generator G given by 3.19.

3.90 *Downward staircase.* Let X satisfy 3.58 with $X_0 = x_0 \in (0,1]$ and

$$j(x,v) = x^3 v \; 1_{[0,1]}(x^2 v), \qquad x \in [0,1], \; v \in \mathbb{R}_+.$$

Let $T_1, T_2, \ldots$ be the successive jump times and define $T_\alpha = \lim\; T_n$.

a) Describe the Markov chain (Y_n), where $Y_n = X_{T_n}$.
b) Show that $Y_1, Y_2, \ldots$ are the atoms of a Poisson random measure on the interval $(0, x_0)$. What is the mean measure?
c) Compute $\mathbb{E}^{x_0}\; T_\alpha$. Note that $X_t = 0$ on $\{t \geq T_\alpha\}$.
d) Compute the transition function (P_t) for X.

3.91 *Transition function for 3.65.* Let X be the process described in Example 3.65. Let P_t and P_t^* be as defined by 3.68. In view of Proposition 3.71, we assume that (P_t^*) is known. This is to compute (P_t). We use $\mathbb{P}^\mu = \int_D \mu(dx)\mathbb{P}^x$.

a) Show that $\mathbb{P}^x\{T_\alpha \leq t\} = 1 - P_t^*(x,\; D)$. Thus,

$$\nu(B) = \mathbb{P}^\mu\left\{T_\alpha \in B\right\} = \int_D \mu(dx)\; \mathbb{P}^x\left\{T_\alpha \in B\right\}, \qquad B \in \mathcal{B}_{\mathbb{R}_+},$$

is well-specified. Let ν^n be the n-fold convolution of ν with itself, with $\nu^0 = \delta_0$ obviously.

b) Define $\rho = \sum_{n=0}^{\infty} \nu^n$. Obviously,

$$\rho(B) = \mathbb{E}^\mu \sum_{n=0}^{\infty} I(T_{n\alpha}, B).$$

Show that $\rho(B) < \infty$ for B compact. Hint:

$$\mathbb{P}^\mu \{T_{n\alpha} \le t\} = \mathbb{P}^\mu \left\{e^{-Tn\alpha} \ge e^{-t}\right\} \le e^t \ \mathbb{E}^\mu e^{-T_{n\alpha}} = e^t \ (\mathbb{E}^\mu \ e^{-T_\alpha})^n.$$

c) Show that $\mathbb{E}^\mu f \circ X_t = \int_{[0,t]} \rho(ds) \int_D \mu(dx) P^*_{t-s} f(x)$.

d) Show that $P_t f(x) = P_t{}^* f(x) + \int_{[0,t]} \nu(ds) \mathbb{E}^\mu f \circ X_{t-s}$.

3.92 *Step processes with discrete state spaces.* Let D be a countable set; we identify it with $\mathbb{N}$ or a subset of $\mathbb{N}$, and regard D as a subset of $E = \mathbb{R}$. We use the notational principles mentioned in Exercise 1.35.

Let X be a step process (right-continuous) satisfying 3.58 and whose values are in D. Then, its Lévy kernel satisfies

$$L(x, A) = \sum_{y \in A} \ell(x, y), \qquad x \in D, A \subset D$$

for some positive numbers $\ell(x, y)$ with

$$\ell(x, x) = 0, \qquad k(x) = \sum_{y \in D} \ell(x, y) < \infty.$$

We may assume that the jump function j has the following form: For each x, let $\{A_{xy} : y \in D\}$ be a partition of $[0, k(x)]$ such that, for each y, the set A_{xy} is an interval of length $\ell(x, y)$. Then, put

$$j(x, v) = \sum_{y \in D} (y - x) 1_{A_{xy}}(v), \qquad x \in D, v \in \mathbb{R}_+.$$

Show that the generator G of X has the form

$$Gf(x) = \sum_{y \in D} g(x, y) f(y), \qquad x \in D,$$

and identify the entries $g(x, y)$ of the matrix G. Let $p_t(x, y) = \mathbb{P}^x\{X_t = y\}$ as before in 1.35. show that

$$\frac{d}{dt}\, p_t(x, y) = \sum_{z \in D} p_t(x, z)\ g(z, y), \qquad x, y \in D,$$

and also

$$\frac{d}{dt}\, p_t(x, y) = \sum_{z \in D} g(x, z)\ p_t(z, y), \qquad x, y \in D.$$

If the kernel L is bounded, that is, if the function k is bounded, then we have

$$P_t = e^{tG} = \sum_{n=0}^{\infty} \frac{t^n}{n!} G^n,$$

where P_t is the matrix whose (x,y)-entry is $p_t(x, y)$, and G is the matrix with entries $g(x, y)$ similarly.

3.93 *Semigroups on discrete spaces.* Let D be a countable set. Let the matrices $P_t = [p_t(x, y)]$ satisfy $P_t P_u = P_{t+u}$. Without reference to a Markov process, suppose that $\lim_{t\to 0} p_t(x, x) = 1$ for every x in D. Then, it can be shown that

$$g(x, y) = \lim_{t\to 0} \frac{d}{dt}\, p_t(x, y), \qquad x, y \in D,$$

exist and satisfy $g(x, x) = -k(x)$, with $k(x) \in [0, +\infty]$, and $g(x, y) \in [0, \infty)$ for $x \neq y$, and $\sum\limits_{y\neq x} g(x, y) \leq k(x)$. The state x is a trap if $k(x) = 0$, is holding (stable is another term for the same) if $0 < k(x) < \infty$, and is instantaneous if $k(x) = +\infty$. (P_t) is said to be conservative $\sum_{y\neq x} g(x, y) = k(x)$ for every x.

3.94 *Continuation.* Let $D = \mathbb{N}$. For x and y in D, let

$$p_t(x, y) = \mathbb{P}\left\{Y_{s+t} = y | Y_s = x\right\},$$

where Y is the process defined by 1.33. Show that $p_t(x, x) \to 1$ as $t \to 0$. Show that $k(x) \in (0, \infty)$ for each x; identify $k(x)$ in terms of the data $m(q)$, rational q. Show that

$$g(x, y) = 0, \qquad x \neq y.$$

3.95 *Itô processes.* These are Markov processes X that satisfy a stochastic integral equation of the form

$$\begin{aligned} X_t = X_0 &+ \int_0^t a \circ X_s\, ds + \int_0^t b \circ X_s\, dW_s \\ &+ \int_{[0,t]\times\mathbb{R}_+} (M(ds, dv) - ds\, dv) j(X_{s-}, v) 1_{\{j(X_{s-},v)\leq 1\}} \\ &\qquad + \int_{[0,t]\times\mathbb{R}_+} M(ds, dv) j(X_{s-}, v) 1_{\{j(X_{s-},v)>1\}}. \end{aligned}$$

Here, a, b, M, W, j are as in 3.1, but without the condition 3.2, and j must satisfy

$$\int_{\mathbb{R}_+} du \left[\left(j(x,\, v)^2 \wedge 1\right)\right] < \infty,$$

and the third integral is a stochastic integral, defined as a limit in probability. This class of processes includes all Lévy processes (see Itô-Lévy decomposition), all Itô diffusions, all jump-diffusions, and more. See the complement 5.51 for more.

4 Markov Systems

This section is to introduce Markov processes in the modern setting. We shall introduce a probability measure $\mathbb{P}^x$ for each state x; it will serve as the conditional probability law given that the process X is at x initially. We shall introduce a shift operator θ_t for each time t; it will indicate that t is the present time. And, we shall think of X as the motion of a particle that lives in E, but might die or be killed at some random time; this will require an extra point to serve as the cemetery.

The space E will be kept fairly general. Although the Markov property has nothing to do with the topology of E, the analytical machinery requires that E be topological and X right-continuous. The reader is invited to take $E = \mathbb{R}$ on a first reading. This section is independent of Sections 2 and 3; but some familiarity with at least Section 2 would be helpful as motivation. Also helpful is the formalism of Lévy processes; the connections are spelled out in Exercises 4.31 and 4.32.

The system

This is to describe the setting for Markov processes. The time-set is $\mathbb{R}_+$; it will be extended to $\bar{\mathbb{R}}_+$.

4.1 *State space.* Let E be a locally compact separable metrizable space, and $\mathcal{E}$ the Borel σ-algebra on it. If E is compact, we let ∂ be an isolated point outside E. If E is not compact, ∂ will be the "point at infinity" in the one point compactification of E. We put

$$\bar{E} = E \cap \{\partial\}, \qquad \bar{\mathcal{E}} = \sigma(\mathcal{E} \cup \{\bar{E}\}).$$

4.2 *Convention.* Every function $f:\ E \mapsto \bar{\mathbb{R}}$ is extended onto $\bar{E}$ automatically by setting $f(\partial) = 0$. Thus, writing $f \in \mathcal{E}$ indicates also a function in $\bar{\mathcal{E}}$ with $f(\partial) = 0$; otherwise, we write $\bar{f} \in \bar{\mathcal{E}}$ to mean that $\bar{f}$ is defined on $\bar{E}$ and is $\bar{\mathcal{E}}$-measurable without an assumption on $\bar{f}(\partial)$.

4.3 *Transition semigroups.* Let (P_t) be a family of sub-Markov kernels on $(E, \mathcal{E})$ such that $P_t P_u = P_{t+u}$. Each P_t is extended to become a Markov kernel $\bar{P}_t$ on $(\bar{E}, \bar{\mathcal{E}})$ by putting

$$\bar{P}_t(x, B) = P_t(x, B \cap E) + (1 - P_t(x, E))I(\partial, B), \qquad x \in \bar{E}, B \in \bar{\mathcal{E}}.$$

Note that $\bar{P}_t(\partial, B) = I(\partial, B) = 1_B(\partial)$ by the preceding convention applied to the function $x \mapsto P_t(x, E)$ on E, namely, the convention that puts $P_t(\partial, E) = 0$. It is easy to check that $\bar{P}_t \bar{P}_u = \bar{P}_{t+u}$.

4.4 *Stochastic base.* Let $(\Omega, \mathcal{H})$ be a measurable space, $\mathcal{F} = (\mathcal{F}_t)$ a filtration over it, and $\theta = (\theta_t)$ a family of "shift" operators $\theta_t : \Omega \mapsto \Omega$ such that $\theta_0 \omega = \omega$ and

$$\theta_u(\theta_t \omega) = \theta_{t+u}\, \omega$$

for every ω in Ω. We assume that there is a special point ω_∂ in Ω, and that $\theta_t\omega_\partial = \omega_\partial$ for all t, and $\theta_\infty\omega = \omega_\partial$ for all ω. Finally, let $\mathbb{P}^\bullet = (\mathbb{P}^x)$ be a family of probability measures $\mathbb{P}^x$ on $(\Omega, \mathcal{H})$ such that $(x, H) \mapsto \mathbb{P}^x(H)$ is a transition kernel from $(\bar{E}, \bar{\mathcal{E}})$ into $(\Omega, \mathcal{H})$.

4.5 *Stochastic process.* Let $X = (X_t)$ be a process with state space $(\bar{E}, \bar{\mathcal{E}})$, adapted to the filtration $\mathcal{F}$, and with the point ∂ as a trap; the last phrase means that if $X_t(\omega) = \partial$ then $X_{t+u}(\omega) = \partial$ for all $u \geq 0$. We assume that $X_0(\omega_\partial) = \partial$, and that $X_\infty(\omega) = \partial$ for all ω, and that

$$X_u(\theta_t\omega) = X_{t+u}(\omega), \qquad \omega \in \Omega, \quad t, u \in \bar{\mathbb{R}}_+.$$

We let $\mathcal{G}^o = (\mathcal{G}^o_t)$ be the filtration generated by X, and put $\mathcal{G}^o_\infty = \vee_t\ \mathcal{G}^o_t$ as usual.

Markov system

Throughout this section and further we are working with the system described in 4.1–4.5 above. In conditional expectations and probabilities, we use the old conventions (see V.2.21 *et seq.*) and put

4.6 $$\mathbb{P}^x_T = \mathbb{P}^x(\cdot|\mathcal{F}_T), \qquad \mathbb{E}^x_T = \mathbb{E}^x(\cdot|\mathcal{F}_T).$$

The following is the enhanced version of Markovness.

4.7 DEFINITION. *The system* $\mathfrak{X} = (\Omega, \mathcal{H}, \mathcal{F}, \theta, X, \mathbb{P}^\bullet)$ *is said to be* Markov *with* living space E *and transition semigroup* (P_t) *if the following hold:*

Normality. $\mathbb{P}^x\{X_0 = x\} = 1$ *for every* x *in* $\bar{E}$*.*

Right-continuity for $\mathcal{F}$. *The filtration* $(\mathcal{F}_t)$ *is right-continuous.*

Regularity of paths. *For every* ω*, the path* $t \mapsto X_t(\omega)$ *is right-continuous and has left-limits as a function from* $\mathbb{R}_+$ *into* $\bar{E}$*.*

Markov property. *For every* x *in* E *and every* t *and* u *in* $\mathbb{R}_+$*,*

4.8 $$\mathbb{E}^x_t\ f \circ X_{t+u} = P_u f \circ X_t, \qquad f \in \mathcal{E}_+.$$ □

The normality condition makes $\mathbb{P}^x$ the probability measure on $(\Omega, \mathcal{H})$ under which X is started at x. The right-continuity of $\mathcal{F}$ enriches the pool of stopping times and will be of further use with the strong Markov property; note that $\mathcal{G}^o_{t+} \subset \mathcal{F}_{t+} = \mathcal{F}_t$.

4.9 REMARK. In terms of the definitions of Section 1, the Markov property of the preceding definition implies the following for each x in $\bar{E}$: Over the probability space $(\Omega, \mathcal{H}, \mathbb{P}^x)$, the process X is a (time-homogeneous) Markov process with state space $(\bar{E}, \bar{\mathcal{E}})$ and transition function $(\bar{P}_t)$ given in 4.3. This can be seen by noting that, in view of the conventions, 4.8 implies that

$$\mathbb{E}^x\ \bar{f} \circ X_{t+u} = \bar{P}_u \bar{f} \circ X_t, \qquad \bar{f} \in \bar{\mathcal{E}}_+.$$

Thus, Definition 4.7 introduces a family of Markov processes, one for each x in $\bar{E}$, but all these processes have the same transition function and are intertwined in a systematic manner.

4.10 REMARK. The meaning of P_t is implicit in 4.8. There, putting $t = 0$, applying the expectation operator $\mathbb{E}^x$ to both sides, and using the normality,

4.11 $$\mathbb{E}^x\ f \circ X_u = P_u f(x), \qquad x \in E, u \in \mathbb{R}_+,\ f \in \mathcal{E}_+.$$

This remains true for $x = \partial$ as well, because $X_u = \partial$ almost surely under $\mathbb{P}^\partial$, and the conventions yield $f(\partial) = 0$ and $P_u f(\partial) = 0$. Hence, we may re-write the condition 4.8, using $X_{t+u} = X_u \circ \theta_t$ as in 4.5, in the form

4.12 $$\mathbb{E}_t^x\ f \circ X_u \circ \theta_t = \mathbb{E}^{X_t}\ f \circ X_u, \qquad f \in \mathcal{E}_+,$$

but with a cautionary remark: the right side stands for $g \circ X_t$ where

4.13 $$g(y) = \mathbb{E}^y f \circ X_u = P_u f(y).$$

Almost surely

Since there is a multitude of probability measures $\mathbb{P}^x$, it is convenient to say "almost surely" to mean "almost surely under $\mathbb{P}^x$ for every x in $\bar{E}$". Similarly, a proposition $\pi(\omega)$ for ω is said to hold for almost every ω in H, and then we write

4.14 $$\pi \textit{ a.e. on } H, \text{ or, } \pi(\omega) \quad \text{for } \textit{a.e. } \omega \text{ in } H,$$

if $H \in \mathcal{H}$ and for every x in $\bar{E}$ there is a subset H_x of H in $\mathcal{H}$ such that $\mathbb{P}^x(H \setminus H_x) = 0$ and $\pi(\omega)$ holds for every ω in H_x.

Lifetime of X

According to 4.5, the "boundary" point ∂ is a trap; it is the final resting place for X. Thus

4.15 $$\zeta = \inf \{t \in \mathbb{R}_+ :\ X_t = \partial\}$$

is called the *lifetime* of X. Note that $\zeta(\omega) > u$ if and only if $X_u(\omega) \in E$; hence the term "living space" for E. When $\mathfrak{X}$ is Markov, it follows from 4.11 with $f = 1_E$ that

$$\mathbb{P}^x \{\zeta > u\} = P_u (x, E), \qquad x \in E, \qquad u \in \mathbb{R}_+.$$

This gives meaning to the defect $1 - P_u(x,\ E)$ when P_u is sub-Markov. The process X is said to be *conservative* if $P_t(x,\ E) = 1$ for all x in E, that is, if every P_t is Markov.

Markov property

Here we explore the essential condition in Definition 4.7, the Markov property.

4.16 REMARK. The collection of functions f in $\mathcal{E}$ for which 4.8 holds is a monotone class. Thus, in order for 4.8 to hold for all f in $\mathcal{E}_+$, it is sufficient that it hold for the indicators of Borel subsets of E, or for the indicators of open subsets of E, or for bounded continuous functions on E, or for continuous functions on E with compact support.

The next theorem captures the essence of Markov property by replacing $f \circ X_u$ in 4.12 with an arbitrary functional V of X. Note, for example, that

$$V = f(X_{u_1}, \ldots, X_{u_n}) \Rightarrow V \circ \theta_t = f(X_{t+u_1}, \ldots, X_{t+u_n}).$$

Thus, the proof of the next theorem is immediate from Remark 4.10, Theorem 1.2, and 4.12 above.

4.17 THEOREM. (Markov property). *Suppose that $\mathfrak{X}$ is a Markov system. Then, for every x in $\bar{E}$ and t in $\mathbb{R}_+$ and positive V in $\mathcal{G}^o_\infty$,*

$$\mathbb{E}^x_t \ V \circ \theta_t = \mathbb{E}^{X_t} V.$$

4.18 EXAMPLE. This is to illustrate the preceding theorem with a specific V. The aim is to clarify some technical matters which are implicit, and also to re-iterate the heuristics.

a) Let $f \in \mathcal{E}_b$, that is, let $f : E \mapsto \mathbb{R}$ be a bounded Borel function. Since X is right-continuous and each X_t is measurable with respect to $\mathcal{G}^o_\infty$ and $\bar{\mathcal{E}}$, the mapping $(t, \omega) \mapsto X_t(\omega)$ is measurable with respect to $\mathcal{B}_{\mathbb{R}_+} \otimes \mathcal{G}^o_\infty$ and $\bar{\mathcal{E}}$. Thus, $(t, \omega) \mapsto f \circ X_t(\omega)$ is in $\mathcal{B}_{\mathbb{R}_+} \otimes \mathcal{G}^o_\infty$.

b) Hence, Fubini's theorem shows that, for fixed $p > 0$,

$$V = \int_{\mathbb{R}_+} du \ e^{-pu} \ f \circ X_u$$

defines a bounded variable V in $\mathcal{G}^o_\infty$; and

$$g(y) = \mathbb{E}^y \ V, \qquad y \in E,$$

defines a function g in $\mathcal{E}_b$; this is because $\mathbb{P}^\bullet$ is a transition kernel, and $g(y) = \mathbb{P}^y V$ in the kernel-function notation. Similarly, and since $X_u \circ \theta_t = X_{t+u}$ by 4.5,

$$V \circ \theta_t = \int_{\mathbb{R}_+} du \ e^{-pu} \ f \circ X_{t+u}$$

is a well-defined bounded random variable in $\mathcal{G}^o_\infty$.

c) Now, the heuristic part: $g(y)$ is our estimate of V made at time 0 if the initial state is y. The initial state of the process $X \circ \theta_t$ is X_t. The variable $V \circ \theta_t$ is the same functional of $X \circ \theta_t$ as V is of X. Thus, our estimate of $V \circ \theta_t$ made at time t should be $g(X_t)$ if we think that t is the origin of time and all the past is past. □

Blumenthal's zero-one law

This useful result is a consequence of the normality, the Markov property, and the right-continuity of the filtration $\mathcal{F}$.

4.19 THEOREM. *Let $\mathfrak{X}$ be Markov. Let H be an event in $\mathcal{G}^o_{0+}$. For each x in $\bar{E}$, then, $\mathbb{P}^x(H)$ is either 0 or 1.*

Proof. Put $V = 1_H$. Clearly, $V = V \cdot V$, and $V = V \circ \theta_0$ since $\theta_0 \omega = \omega$ for all ω; hence, $V = V \cdot (V \circ \theta_0)$. On the other hand, the filtration $(\mathcal{G}^o_t)$ generated by X is coarser than $(\mathcal{F}_t)$ by the adaptedness of X to $\mathcal{F}$; and, thus, $\mathcal{G}^o_{t+} \subset \mathcal{F}_{t+} = \mathcal{F}_t$, the last equality being the definition of right-continuity for $\mathcal{F}$. This implies, since $V \in \mathcal{G}^o_{0+}$ by assumption, that $V \in \mathcal{F}_0$. It now follows from the Markov property 4.17 at $t = 0$ that, since $\mathbb{E}^x = \mathbb{E}^x \mathbb{E}^x_0$,

$$\mathbb{E}^x\ V = \mathbb{E}^x V \cdot V \circ \theta_0 = \mathbb{E}^x V \mathbb{E}^{X_0} V.$$

But, by normality, $X_0 = x$ with $\mathbb{P}^x$-probability one. Hence, $\mathbb{E}^x V = \mathbb{E}^x V\ \mathbb{E}^x V$, which implies that $\mathbb{E}^x V = \mathbb{P}^x(H)$ is either 0 or 1. □

Holding points, instantaneous points

Started at a point x, the process either exits x instantaneously or stays at x some strictly positive amount of time. This dichotomy is a consequence of the preceding zero-one law.

Suppose that the system $\mathfrak{X}$ is Markov. Define

4.20 $$R = \inf\{t > 0 : X_t \neq X_0\}.$$

Then, R is a stopping time of $\left(\mathcal{G}^o_{t+}\right)$, and thus, the event $\{R = 0\}$ belongs to $\mathcal{G}^o_{0+}$. Hence, for fixed x in $\bar{E}$, the zero-one law applies to show that

4.21 $$\mathbb{P}^x\{R = 0\}$$

is either 0 or 1. It this probability is 0, then x is said to be a *holding point*; if it is 1, then x is said to be *instantaneous.* A holding point x is called a *trap*, or an *absorbing* point, if $\mathbb{P}^x\{R = \infty\} = 1$.

The point ∂ is a trap; there may be other traps. For step processes of Section 3, and for Poisson and compound Poisson processes, every point of E is a holding point. For Brownian motions in $E = \mathbb{R}^d$, every point of E is instantaneous; similarly for Itô diffusions. For Lévy processes other than compound Poisson, every point of $E = \mathbb{R}^d$ is instantaneous.

Let x be a holding point. Started at x, the process stays there for a strictly positive amount R of time. The next theorem shows that the distribution of R is exponential, and the state X_R is independent of R. We shall show later (see 5.23) that when X is strong Markov, it must exist x by a jump.

4.22 THEOREM. *Let $\mathfrak{X}$ be Markov, and let x be a holding point. Then,*

$$\mathbb{P}^x\{R > t, X_R \in B\} = e^{-k(x)t} K(x, B),\ t \in \mathbb{R}_+,\ B \in \bar{\mathcal{E}},$$

for some number $k(x) < \infty$ and some measure $B \mapsto K(x, B)$ on $(\bar{E}, \bar{\mathcal{E}})$.

Proof. For every ω, the definition of R implies that

$$R(\omega) > t + u \Leftrightarrow R(\omega) > t \quad \text{and} \quad R(\theta_t \omega) > u,$$

and, then,

$$X_R(\omega) = X_{t+R(\theta_t\omega)}(\omega) = X_{R(\theta_t\omega)}(\theta_t\omega) = X_R(\theta_t\omega).$$

Thus, for B in $\bar{\mathcal{E}}$,

4.23 $$\begin{aligned}\mathbb{P}^x\{R > t + u,\ X_R \in B\} &= \mathbb{P}^x\{R > t,\ R \circ \theta_t > u,\ X_R \circ \theta_t \in B\} \\ &= \mathbb{E}^x 1_{\{R>t\}} \mathbb{P}^{X_t}\{R > u, X_R \in B\} \\ &= \mathbb{P}^x\{R > t\}\, \mathbb{P}^x\{R > u, X_R \in B\};\end{aligned}$$

here, the second equality is justified by the Markov property of 4.17 at time t, and the third equality by the observation that $X_t = X_0$ on $\{R > t\}$ followed by the normality condition.

In 4.23, take $B = \bar{E}$. The result is

$$\mathbb{P}^x\{R > t + u\} = \mathbb{P}^x\{R > t\}\, \mathbb{P}^x\{R > u\}$$

for all t and u in $\mathbb{R}_+$; and $t \mapsto \mathbb{P}^x\{R > t\}$ is obviously right-continuous and is equal to 1 at $t = 0$ (since x is holding). Thus, there exists $k(x)$ in $\mathbb{R}_+$ such that

4.24 $$\mathbb{P}^x\{R > t\} = e^{-k(x)t}, \qquad t \in \mathbb{R}_+.$$

Next, put this into 4.23 and set $u = 0$. Since x is holding, $\mathbb{P}^x\{R > 0\} = 1$, and 4.23 becomes

$$\mathbb{P}^x\{R > t, X_R \in B\} = e^{-k(x)t}\ \mathbb{P}^x\{X_R \in B\},$$

which has the form claimed. □

4.25 REMARK. The point x is a trap if and only if $k(x) = 0$. In fact, 4.24 holds for instantaneous x as well; then, $k(x) = +\infty$.

Measures P^μ

For each x in $\bar{E}$, the distribution of X_0 under $\mathbb{P}^x$ is the Dirac measure δ_x; this is by the normality of $\mathfrak{X}$. Thus, for an arbitrary probability measure on $(\bar{E}, \bar{\mathcal{E}})$,

4.26 $$\mathbb{P}^\mu(H) = \int_E \mu(dx)\, \mathbb{P}^x(H), \qquad H \in \mathcal{H},$$

defines a probability measure on $(\Omega, \mathcal{H})$, under which X_0 has the distribution μ. This follows from Theorem I.6.3 via the hypothesis in 4.4 that $(x,\ H) \mapsto \mathbb{P}^x(H)$ is a transition probability kernel.

Exercises

4.27 *Compound Poissons.* Let $X = X_0 + Y$, where Y is a compound Poisson process (with $Y_0 = 0$) whose jump times form a Poisson process with rate c, and whose jump sizes have the distribution μ. Classify the states as holding or instantaneous. What are $k(x)$ and $K(x, B)$ of Theorem 4.22 in this case?

4.28 *Step processes.* Let X be a step process as in Section 3. Show that every point x in $E = \mathbb{R}^d$ is a holding point. Compute $k(x)$ and $K(x, B)$ of Theorem 4.22 in terms of the Lévy kernel L of X.

4.29 *Brownian motion with holding boundary points.* Let $a < 0 < b$ be fixed. Started in the interval $(a,\ b)$, the motion X is standard Brownian until it exits the interval; if the exit is at a, then X stays at a an exponential time with parameter $k(a)$ and then jumps to the point 0; if the exit is at b, then X stays at b an exponential time with parameter $k(b)$ and then jumps to 0. Once at 0, the motion resumes its Brownian character, and so on. Classify the points of $E = [a, b]$; identify the distributions $K(x, \cdot)$ for the holding points x.

4.30 *Achilles' run.* The living space E is (0,1]; then $\partial = 0$ necessarily. Started at x in E, the particle stays at x an exponential amount of time with mean x and, then, jumps to $y = x/2$; at y, it stays an exponential time with mean y and, then, jumps to $z = y/2$; and so on. Let $T_1,\ T_2, \ldots$ be the successive jump times, put $\zeta = \lim\ T_n$, and define $X_t = \partial = 0$ for t in $[\zeta, +\infty]$. Show that all points are holding points. Identify the parameters $k(x)$ and $K(x, B)$. Compute $\mathbb{E}^x\zeta$.

4.31 *Lévy processes.* Let $\mathfrak{X} = (\Omega, \mathcal{H}, \mathcal{F}, \theta, X, \mathbb{P}^\bullet)$ be a Markov system with living space $E = \mathbb{R}^d$. Suppose that its transition semigroup (P_t) is such that, for each t,

$$P_t f(x) = \int_E \pi_t(dy)\ f(x+y), \qquad x \in E,\ f \in \mathcal{E}_+.$$

for some probability measure π_t on E. Show that, then, X has stationary and independent increments under each $\mathbb{P}^x$. More precisely, for each x in E, the process $Y = (X_t - X_0)_{t \in \mathbb{R}_+}$ is a Lévy process over the stochastic base $(\Omega, \mathcal{H}, \mathcal{F}^x, \theta, \mathbb{P}^x)$ in the sense of Definition VII.3.3; here $\mathcal{F}^x$ is the augmentation of $\mathcal{F}$ with respect to the probability measure $\mathbb{P}^x$. Show this.

4.32 *Continuation.* This is a converse to the preceding. Let X and $\mathcal{B} = (\Omega, \mathcal{H}, \mathcal{F}, \theta, \mathbb{P})$ be as in Definitions VII.3.1 and VII.3.3. Put $E = \mathbb{R}^d,\ \mathcal{E} = \mathcal{B}(\mathbb{R}^d)$, and set ∂ to be the "point at infinity." Define

$$\hat\Omega = \bar{E} \times \Omega,\ \hat{\mathcal{H}} = \bar{\mathcal{E}} \otimes \mathcal{H},\ \hat{\mathcal{F}}_t = \bar{\mathcal{E}} \otimes \mathcal{F}_t, \hat{\mathbb{P}}^x = \delta_x \times \mathbb{P}$$

for x in $\bar{E}$; and, for $\hat\omega = (x, \omega)$ in $\hat\Omega$, put

$$\hat{X}_t(\hat\omega) = x + X_t(\omega), \qquad \hat\theta_t\hat\omega = (\hat{X}_t(\hat\omega), \theta_t\omega).$$

Show that $\hat{\mathfrak{X}} = (\hat{\Omega}, \hat{\mathcal{H}}, \hat{\mathcal{F}}, \hat{\theta}, \hat{X}, \hat{\mathbb{P}}^\bullet)$ is a Markov system, in the sense of Definition 4.7, with living space E and transition function (P_t) given by

$$P_t(x, B) = \mathbb{P}\{x + X_t \in B\}.$$

5 Hunt Processes

These are Markov processes which have almost all the properties desired of a Markov process. Itô diffusions, jump-diffusions, simple step processes, and all Lévy processes (including, of course, Poisson and Brownian motions) are Hunt processes. We choose them to form the central reference system for the theory; even when a Markov process is not Hunt, it is best to describe it by telling how it differs from a Hunt process.

Throughout this section, $\mathfrak{X} = (\Omega, \mathcal{H}, \mathcal{F}, \theta, X, \mathbb{P}^\bullet)$ is a Markov system with living space E and transition semigroup (P_t); see Definition 4.7 and the setup 4.1–4.6. Recall, in particular, that the filtration $\mathcal{F}$ is right-continuous and that the path $t \mapsto X_t$ is right-continuous and has left-limits in $\bar{E}$. Recall also that $(\mathcal{G}_t^o)$ is the filtration generated by X.

In preparation for the definition of strong Markov property next, we note that $\bar{f} \circ X_T$ is $\mathcal{F}_T$-measurable for every $\mathcal{F}$-stopping time T and every $\bar{\mathcal{E}}$-measurable $\bar{f} : \bar{E} \mapsto \bar{\mathbb{R}}_+$. For continuous $\bar{f}$, this follows from Theorem V.1.14 via the right-continuity of $\bar{f} \circ X$. Then, a monotone class argument extends it to all $\bar{\mathcal{E}}$-measurable $\bar{f}$. For positive V in $\mathcal{G}_\infty^o$, putting $\bar{f}(y) = \mathbb{E}^y V$ yields a function $\bar{f}$ that is $\bar{\mathcal{E}}$-measurable; and, then, $\bar{f} \circ X_T$ belongs to $\mathcal{F}_T$ as required for it to be a conditional expectation given $\mathcal{F}_T$.

5.1 Definition. *The Markov system $\mathfrak{X}$ is said to be* strong Markov *if, for every $\mathcal{F}$-stopping time T and every positive random variable V in $\mathcal{G}_\infty^o$,*

5.2 $$\mathbb{E}_T^x\, V \circ \theta_T = \mathbb{E}^{X_T}\, V, \qquad x \in E.$$

It is said to be quasi-left-continuous if, for every increasing sequence (T_n) of $\mathcal{F}$-stopping times with limit T,

5.3 $$\lim_n X_{T_n} = X_T \text{ almost surely on } \{T < \infty\}.$$

It is said to be a Hunt system *if it is strong Markov and quasi-left-continuous.*

We shall explore the contents of these definitions and their ramifications. We start with the less familiar concept.

Quasi-left-continuity

If X is continuous, then $\mathfrak{X}$ is quasi-left-continuous automatically. The continuity is not necessary. For instance, if $X - X_0$ is a Poisson process, then $\mathfrak{X}$ is quasi-left-continuous even though X has infinitely many jumps.

Similarly, if $X - X_0$ is a Lévy process, or if X is a jump-diffusion or a step process, then $\mathfrak{X}$ is quasi-left-continuous. These comments will become clear shortly.

Recall the notation X_{t-} for the left-limit of X at t; we put $X_{0-} = X_0$ for convenience. For a random time T, then, X_{T-} is the random variable $\omega \mapsto X_{T-}(\omega) = X_{T(\omega)-}(\omega)$. Suppose now that (T_n) is an increasing sequence of $\mathcal{F}$-stopping times with limit T, and pick ω such that $T(\omega) < \infty$. If $T_n(\omega) = T(\omega)$ for all n large enough, then $\lim X_{T_n}(\omega) = X_T(\omega)$ trivially. Otherwise, if

5.4 $$T_n(\omega) < T(\omega) \quad \text{for all } n,$$

then $\lim X_{T_n}(\omega) = X_{T-}(\omega)$, and quasi-left-continuity would require that $X_{T-}(\omega) = X_T(\omega)$, unless ω happens to be in the negligible exceptional set of 5.3. In other words, if $\mathfrak{X}$ is quasi-left-continuous, then 5.4 is incompatible with

5.5 $$T(\omega) < \infty, \; X_{T-}(\omega) \neq X_T(\omega),$$

except for a negligible set of ω.

We may interpret 5.4 as "predictability" for $T(\omega)$, because the sequence of times $T_n(\omega)$ enables the observer to foresee $T(\omega)$. So, heuristically, quasi-left-continuity is about the continuity of paths at predictable times and, equivalently, about the unpredictability of jump times. We make these remarks precise next.

Predictable times, total unpredictability

We recall some definitions introduced in passing in Chapter V, adapted to the newer meaning of "almost everywhere" given around 4.14. We shall use the notation (read T on H)

5.6 $$T_H(\omega) = \begin{cases} T(\omega) & \text{if } \omega \in H, \\ +\infty & \text{otherwise,} \end{cases}$$

for $\mathcal{F}$-stopping times T and events H in $\mathcal{F}_T$; and, then, T_H is also an $\mathcal{F}$-stopping time.

5.7 Definition. *Let T be an $\mathcal{F}$-stopping time. It is said to be* predictable *if there exists an increasing sequence (T_n) of $\mathcal{F}$-stopping times with limit T such that*

5.8 $$T_n < T \text{ for all } n \qquad \text{a.e.} \quad \text{on } \{0 < T < \infty\}.$$

It is said to be totally unpredictable *if, for every predictable $\mathcal{F}$-stopping time S,*

5.9 $$T = S \quad \textit{almost nowhere on } \{T < \infty\}.$$

For Brownian motion, every hitting time is predictable, and more. For a Poisson process, Proposition VI.5.20 implies that the first jump time is totally unpredictable; see also V.7.31. In Example 3.65, the time T_α is predictable, so are $T_{2\alpha}$, $T_{3\alpha}$, etc. The other times in 3.67 are totally unpredictable. The following enhances the definition.

5.10 Lemma. *Let T be a totally unpredictable $\mathcal{F}$-stopping time. Suppose that (T_n) is an increasing sequence of $\mathcal{F}$-stopping times with limit T on the event $\{T < \infty\}$. Then, 5.4 fails for almost every ω in $\{T < \infty\}$.*

Proof. Let $H_n = \{T_n < T\}$ and $H = \cap_n H_n$. We need to show that

5.11 $$\mathbb{P}^x\left(H \cap \{T < \infty\}\right) = 0$$

for every x. Define $S = T_H$, see 5.6, and define S_n similarly from T_n and H_n; these are all $\mathcal{F}$-stopping times. Since (T_n) is increasing, the sequence (H_n) is shrinking to H, and (S_n) is increasing to S. Moreover, if $\omega \in H$, then $S_n(\omega) = T_n(\omega) < T(\omega) = S(\omega)$. Since $\{0 < S < \infty\} \subset H$, we conclude that S is predictable. It follows from the total unpredictability of T that 5.9 holds; and 5.9 is the same as 5.11 for all x. □

Total unpredictability of jumps

Let T be an $\mathcal{F}$-stopping time. We call T a *time* of *continuity* for X if (recall that $X_{0-} = X_0$)

5.12 $$X_{T-} = X_T \qquad \text{a.e. on } \{T < \infty\},$$

and *a jump time* for X if

5.13 $$X_{T-} \neq X_T \qquad \text{a.e. on } \{T < \infty\}.$$

The following clarifies the true meaning of quasi-left-continuity.

5.14 Theorem. *The following are equivalent:*

a) The Markov system $\mathfrak{X}$ is quasi-left-continuous.

b) Every predictable $\mathcal{F}$-stopping time is a time of continuity.

c) Every jump time is totally unpredictable.

Proof. Suppose (a). Let T be predictable. Then, there is (T_n) increasing to T such that 5.8 holds. Therefore, $\lim X_{T_n} = X_{T-}$ a.e. on $\{T < \infty\}$. But the limit is X_T a.e. on $\{T < \infty\}$ by the assumed quasi-left-continuity. Thus, 5.12 holds. Hence $(a) \Rightarrow (b)$.

Suppose (b). Let T be a jump time, that is, let 5.13 hold, and let S be predictable. Then, $X_{T-} = X_{S-} = X_S = X_T$ almost everywhere on $\{T = S,\ T < \infty\}$ in view of (b) for S. This means, in view of 5.13 for T, that 5.9 holds. Hence, T is totally unpredictable. So, $(b) \Rightarrow (c)$.

Suppose (c). Let (T_n) be an increasing sequence of $\mathcal{F}$-stopping times with limit T. On $\{X_{T-} = X_T,\ T < \infty\}$, we have $\lim X_{T_n} = X_T$ obviously. To show quasi-left-continuity at T, we show next that

5.15 $$\lim\ X_{T_n} = X_T \qquad \text{a.e. on } H = \{X_{T-} \neq X_T,\ T < \infty\}.$$

Define $S = T_H$ as in 5.6; it is obviously a jump time. By the assumed (c), then, S is totally unpredictable. Moreover, $\lim T_n = T = S$ on $\{S < \infty\}$, because $\{S < \infty\} = H \subset \{T < \infty\}$. It follows from Lemma 5.10 that, almost surely on $\{S < \infty\} = H$, we have $T_n = S$ for all n large enough. Since $S = T$ on H, we have 5.15. So, $(c) \Rightarrow (a)$. □

Classification of stopping times

A hasty reading of the last theorem might suggest that a stopping time is predictable if and only if it is a continuity time. This is false in general; see Example 5.22. However, it is true provided that we limit ourselves to the stopping times of $\left(\mathcal{G}^o_{t+}\right)$. We offer this without proof. Note that, when $\mathfrak{X} = (\dots, \mathcal{F}, \dots)$ is quasi-left-continuous, then so is the Markov system $\mathfrak{X}^0 = \left(\dots, \left(\mathcal{G}^o_{t+}\right), \dots\right)$; so, half of the statements next follow from the last theorem.

5.16 THEOREM. *Suppose that $\mathfrak{X}$ is quasi-left-continuous. Consider a stopping time of $\left(\mathcal{G}^o_{t+}\right)$. It is predictable if and only if it is a continuity time for X; it is totally unpredictable if and only if it is a jump time for X.* □

5.17 REMARK. Let S be an $\mathcal{F}$-stopping time. We call it *σ-predictable* if there is a sequence (S_n) of predictable $\mathcal{F}$-stopping times such that, for almost every ω with $S(\omega) < \infty$, we have $S(\omega) = S_n(\omega)$ for some n. An arbitrary $\mathcal{F}$-stopping time R can be written as

5.18 $$R = S \wedge T,$$

where S is σ-predictable and T totally unpredictable. The preceding theorem implies, in particular, that for a quasi-left-continuous system, every σ-predictable stopping time of $\left(\mathcal{G}^o_{t+}\right)$ is necessarily predictable. Hence, every $\left(\mathcal{G}^o_{t+}\right)$-stopping time R has the form 5.18 with S predictable and T totally unpredictable; indeed, in the notation 5.6,

$$S = R_{\{X_{R-} = X_R,\, R<\infty\}}, \qquad T = R_{\{X_{R-} \neq X_R,\, R<\infty\}}.$$

Examples

All continuous Markov processes are obviously quasi-left-continuous. So, we concentrate on processes with jumps. The reader will see that quasi-left-continuity at a jump time depends on whether that jump is endogeneous (as in the first example below) or exogeneous (and is caused by kicks from a Poisson).

5.19 *Brownian motion with jump boundaries.* This is a variation on Exercise 4.29. The motion X is Brownian inside the interval (a,b) until the time T of exit; if X_{T-} is a, then X_T has some distribution μ_a on (a,b); if X_{T-} is b, then X_T has some distribution μ_b on (a,b). This $\mathfrak{X}$ is strong Markov; it is not quasi-left-continuous. To see the latter point, let T_n be the time of exit from $(a + 1/n,\ b - 1/n)$; then, (T_n) increases to T, but $X_{T-} \neq X_T$. Note that, in this example, the jumps are triggered by the particle itself. The process of Example 4.29, by contrast, is quasi-left-continuous; its jumps are exogeneous; they are caused by kicks from a Poisson.

5.20 *Step processes.* Suppose that X is a step process. Then $\mathfrak{X}$ is strong Markov, we show now that it is quasi-left-continuous. So, $\mathfrak{X}$ is Hunt.

We start by showing that quasi-left-continuity hold at the time R of first jump. Let (R_n) be a sequence of stopping times increasing to R. Fix x in E and recall that, under $\mathbb{P}^x$, the time R has the exponential distribution with some parameter $k(x) < \infty$.

If x is a trap, then $k(x) = 0$ and $\mathbb{P}^x\{R = +\infty\} = 1$, and thus the condition 5.3 holds at R by default. Suppose that x is not a trap. Observe that, on $\{R_n < R\}$, we have $R = R_n + R \circ \theta_{R_n}$. Thus,

$$\begin{aligned}\mathbb{E}^x R &= \mathbb{E}^x\ R1_{\{R_n = R\}} + \mathbb{E}^x\left(R_n + R \circ \theta_{R_n}\right)\ 1_{\{R_n < R\}}\\ &= \mathbb{E}^x\ R_n 1_{\{R_n = R\}} + \mathbb{E}^x\ R_n 1_{\{R_n < R\}} + \mathbb{E}^x 1_{\{R_n < R\}} \mathbb{E}^{X(R_n)} R\\ &= \mathbb{E}^x R_n + \mathbb{E}^x R\ \mathbb{P}^x\left\{R_n < R\right\};\end{aligned}$$

here, we used the strong Markov properly at R_n and noted that $X_{R_n} = x$ on $\{R_n < R,\ X_0 = x\}$. Since (R_n) is increasing to R and $\mathbb{E}^x R < \infty$ (since x is not a trap), we conclude that

$$\lim_n \mathbb{P}^x\left\{R_n < R\right\} = 0.$$

But, the events $\{R_n < R\}$ are shrinking to $\{R_n < R \text{ for all } n\}$. So,

$$\mathbb{P}^x\left\{R_n < R \text{ for all n}\right\} = 0,$$

that is, for $\mathbb{P}^x$-almost every ω, we have $R_n(\omega) = R(\omega)$ for all n large enough, and hence, $\lim X_{R_n}(\omega) = X_R(\omega)$. Thus, quasi-left-continuity holds at R.

Fix m in $\mathbb{N}$, let T be the $(m+1)^{th}$ jump time and let (T_n) an increasing sequence of stopping times with limit T. Let S denote the m^{th} jump time and put $R_n = S \vee T_n$. Then, (R_n) is increasing to T, and $S \leq R_n \leq T$, and $T = S + R \circ \theta_S$ with R as before (the time of first jump). The arguments of the last paragraph apply with the conditional law $\mathbb{P}^x_S$ replacing $\mathbb{P}^x$; this is by the strong Markov property at S. Thus, almost surely on $\{T < \infty\}$, $R_n = T$ for all n large enough and

$$\lim_n X_{T_n} = \lim X_{R_n} = X_T.$$

So, quasi-left-continuity holds at T; and since m is arbitrary and the process X is a step process, this implies quasi-left-continuity for $\mathfrak{X}$.

5.21 *Lévy processes.* Suppose that $X - X_0$ is a Lévy process; see Exercise 4.31. Then, $\mathfrak{X}$ is strong Markov. We now show that it is quasi-left-continuous and, hence, a Hunt process.

If X is continuous, there is nothing to prove. Suppose that it has jumps. Fix an integer $m \geq 1$, and consider the successive jump times at which X jumps by an amount whose magnitude is in the interval $[1/m, 1/m-1)$. Those jump times form a Poisson process N; in fact,

$$\mathfrak{X}_m = (\Omega, \mathcal{H}, \mathcal{F}, \theta, X_0 + N, \mathbb{P}^\bullet)$$

is a Markov system and $X_0 + N$ is a simple step process. It follows from the preceding example that $\mathfrak{X}_m$ is quasi-left-continuous. So, every one of its jump times is totally unpredictable.

Since this is true for every $m \geq 1$, and since all those jump times put together exhaust all the jump times of X, we conclude that $\mathfrak{X}$ is quasi-left-continuous.

5.22 *Brown and Poisson.* This is to show the necessity, in Theorem 5.16, of restriction to $(\mathcal{G}^o_{t+})$-stopping times. Let $\mathfrak{X}^* = (\Omega, \mathcal{H}, \mathcal{F}, \theta, X^*, \mathbb{P}^\bullet)$ be a Lévy, where $X^* = X_0 + W + N$, with W Wiener, N Poisson, and W and N independent. So, X^* is Hunt. Consider $\mathfrak{X} = (\Omega, \mathcal{H}, \mathcal{F}, \theta, X, \mathbb{P}^\bullet)$ where $X = X_0 + W$, which is also a Hunt process. Let T be the time of first jump for the Poisson process N; it is an $\mathcal{F}$-stopping time and it is totally unpredictable (since $\mathfrak{X}^*$ is Hunt). But, for the Brownian motion X, we have $X_{T-} = X_T$. This is possible because T is not a stopping time of $(\mathcal{G}^o_{t+})$, the filtration of X itself.

Exiting a holding point

This is to supplement Theorem 4.22 by showing that a strong Markov process exits a holding point only by a jump.

5.23 Proposition. *Suppose that the system $\mathfrak{X}$ has the strong Markov property. Let R be the time of exit from X_0 as in 4.20. Then, for every holding point x in E,*

$$\mathbb{P}^x \{X_{R-} \neq X_R\} = 1.$$

Proof. If x is a trap in E, then $R = \infty$ and $X_{R-} = X_0 = x$ and $X_R = \partial$ almost surely; thus the claim holds trivially. Suppose that x is a holding point but not a trap. Observe that, for every ω in $\{X_0 = x\}$,

$$R(\omega) = r, \quad X_r(\omega) = x \;\Rightarrow\; R(\theta_r \omega) = 0$$

by the definition of R. Thus,

5.24 $$\mathbb{P}^x \{X_R = x,\ R \circ \theta_R = 0\} = \mathbb{P}^x \{X_R = x\}.$$

On the other hand, R is a stopping time of $(\mathcal{G}^o_{t+})$ and, therefore, of $(\mathcal{F}_t)$. By the strong Markov property applied at R,

5.25 $$\begin{aligned}\mathbb{P}^x \{X_R = x,\ R \circ \theta_R = 0\} &= \mathbb{E}^x 1_{\{X_R = x\}} \mathbb{P}^{X_R} \{R = 0\} \\ &= \mathbb{E}^x 1_{\{X_R = x\}} \mathbb{P}^x \{R = 0\} = 0\end{aligned}$$

since $\mathbb{P}^x\{R = 0\} = 0$ by the assumption that x is a holding point. It follows from 5.24 and 5.25 that

$$\mathbb{P}^x\{X_R = x\} = 0,$$

which proves the claim since $\mathbb{P}^x\{X_{R-} = x\} = 1$. □

The preceding propositions supplies the rigorous reason for the failure of strong Markov property for Examples 1.28 and 1.29, the delayed uniform motion and Lévy's continuous increasing process. Of course, the proposition has further implications: for instance, if T is a stopping time, on the event that X_T is a holding point, we have $X_R \circ \theta_T \neq X_{R_-} \circ \theta_T = X_T$ almost surely. A somewhat stronger result is next.

No rest for a continuous strong Markov

5.26 PROPOSITION. *Suppose that $\mathfrak{X}$ is strong Markov, and X continuous. Then, almost surely, $t \mapsto X_t$ has no flat segments of finite duration.*

Proof. We are to show that, for almost every ω, there exists no interval $[r,t]$ with $0 \le r < t < \infty$ such that $X_s(\omega) = X_t(\omega)$ for all s in $[r,t]$ and $X_u(\omega) \neq X_t(\omega)$ for some u in (t,∞). Define

$$Q_t(\omega) = t - inf\,\{r \ge 0:\ X_s(\omega) = X_t(\omega) \text{ for all } s \text{ in } [r,t]\}\,.$$

If there were such an interval, then there would exist t such that $Q_t(\omega) > \varepsilon$ for some rational number $\varepsilon > 0$ and that $R(\theta_t\omega) < \infty$; note that $R(\theta_t\omega)$ is the length of the interval from t until the exit from $X_t(\omega)$. Hence, with

$$T_\varepsilon = \inf\{t : Q_t > \varepsilon\}\,, \qquad U_\varepsilon = T_\varepsilon + R \circ \theta_{T_\varepsilon},$$

it is enough to show that, for every x in E and every $\varepsilon > 0$,

5.27 $$\mathbb{P}^x\{U_\varepsilon < \infty\} = 0.$$

Fix x and ε such, and drop ε from the notations T_ε and U_ε. The process (Q_t) is adapted to $(\mathcal{G}_t^\circ)$; thus, T is a stopping time of $\left(\mathcal{G}_{t+}^\circ\right)$ and so is U consequently. Observe that, on the event $\{U < \infty\}$ we have, by the definitions of R, T, U,

$$T < \infty,\ R \circ \theta_T > 0,\ X_T = X_U,\ R \circ \theta_U = 0,$$

the last being due to the continuity of X. So,

$$\mathbb{P}^x\{U < \infty\} = \mathbb{P}^x\{U < \infty, R \circ \theta_U = 0\} = \mathbb{E}^x 1_{\{U<\infty\}} \mathbb{P}^{X_U}\{R = 0\}$$

by the strong Markov property at U; and on the event $\{U < \infty\}$,

$$\mathbb{P}^{X_U}\{R = 0\} = \mathbb{P}^{X_T}\{R = 0\} = \mathbb{P}_T^x\{R \circ \theta_T = 0\} = 0$$

since $X_U = X_T$ and $R \circ \theta_T > 0$ on $\{U < \infty\}$. Hence, 5.27 holds. □

The paths are locally bounded

5.28 PROPOSITION. *Suppose that $\mathfrak{X}$ is quasi-left-continuous. Then, for almost every ω and every $t < \zeta(\omega)$, the set $\{X_s(\omega) : 0 \le s \le t\}$ is contained in some compact subset K^ω of E.*

Proof. Since E is locally compact, there is a sequence (K_n) of compact subsets increasing to E and such that K_n is contained in the interior of K_{n+1} for each n. Let T_n be the time of exit from K_n for each n. Then, (T_n) is an increasing sequence of $(\mathcal{G}^\circ_{t+})$-stopping times, and its limit T is again a stopping time. Thus, by the assumed quasi-left-continuity, $\lim X_{T_n} = X_T$ almost surely on $\{T < \infty\}$. By the right-continuity of X and the way the K_n are picked, $X_{T_{n+1}}$ is outside K_n for every n. Hence, the limit X_T is outside E on $\{T < \infty\}$; in other words, $T \ge \zeta$ on $\{T < \infty\}$ and, therefore, on Ω almost surely. Consequently, for almost every ω, if $t < \zeta(\omega)$, then $t < T_n(\omega)$ for some n, in which case $X_s(\omega) \in K_n$ for all $s \le t$.

Strong Markov property

In Definition 5.1, the strong Markov property is stated in its most useful, intuitive form. Several uses of it appeared in the development above. But, how does one tell whether the given system $\mathfrak{X}$ is strong Markov?

For primary processes such as Poisson, Brownian, and Lévy, the strong Markov property was proved directly. For Itô diffusions and jump-diffusions, its proof exploited the dynamics of the motion and the same property for Poisson and Wiener. Next we aim at processes $\mathfrak{X}$ introduced axiomatically; after some preliminaries, we state a condition on (P_t) that ensures both the strong Markov property and the quasi-left-continuity, see Definition 5.36.

5.29 PROPOSITION. *The Markov system $\mathfrak{X}$ is strong Markov if and only if*

5.30 $$\mathbb{E}^x_T\ f \circ X_{T+u} = P_u f \circ X_T, \qquad x \in E,\ u \in \mathbb{R}_+,$$

for every f in $\mathcal{E}_+$ and every stopping time T of $\mathcal{F}$.

Proof. Necessity is obvious. Sufficiency is essentially as in the proof of Theorem 1.2: It is enough to show that 5.30 implies 5.2 for V having the form

$$V_n = f_1 \circ X_{t_1} \cdots f_n \circ X_{t_n}$$

for some $n \ge 1$, times $0 \le t_1 < \ldots < t_n$, and functions $f_1, \ldots, f_n$ in $\mathcal{E}_+$. This is done by induction on n, whose steps are the same as those of the proof of 1.2; basically, replace u_i there with $T + t_i$. We leave out the details. □

5.31 LEMMA. *The system $\mathfrak{X}$ is strong Markov if and only if*

5.32 $$\mathbb{E}^x\ f \circ X_{T+u} = \mathbb{E}^x\ P_u\ f \circ X_T, \qquad x \in E,\ u \in \mathbb{R}_+,$$

for every f in $\mathcal{E}_+$ and every $\mathcal{F}$-stopping time T.

Proof. Applying $\mathbb{E}^x$ to both sides of 5.30 yields 5.32. For the converse, fix f and T, let H be an event in $\mathcal{F}_T$, and consider the stopping time T_H (T on H defined in 5.6). Assuming 5.32, we get

$$\begin{aligned}\mathbb{E}^x 1_H\ f \circ X_{T+u} &= \mathbb{E}^x\ f \circ X_{T_H+u} \\ &= \mathbb{E}^x\ P_u f \circ X_{T_H} = \mathbb{E}^x 1_H\ P_u f \circ X_T .\end{aligned}$$

Since H in $\mathcal{F}_T$ is arbitrary, this is equivalent to 5.30. □

5.33 LEMMA. *Let T be an $\mathcal{F}$-stopping time that takes values in a countable subset of $\bar{\mathbb{R}}_+$. Then, the strong Markov property holds at T.*

Proof. For fixed t in $\mathbb{R}_+$,

5.34 $$\mathbb{E}^x 1_{\{T=t\}}\ f \circ X_{t+u} = \mathbb{E}^x 1_{\{T=t\}} P_u f \circ X_t$$

by the Markov property, since $\{T = t\} \in \mathcal{F}_t$. The same holds (and both sides vanish) for $t = +\infty$ as well, via the conventions on X_∞ and $f(\partial)$ and $P_u f(\partial)$. Now, summing both sides of 5.34 over the countably many possible values t for T, we obtain 5.32 via the monotone convergence theorem. □

5.35 REMARK. Consider the strong Markov property in the form 5.32. For an arbitrary $\mathcal{F}$-stopping time T, Proposition V.1.20 provides a sequence (T_n) of countably-valued stopping times decreasing to T. By the preceding lemma, 5.32 holds for each T_n. By the right-continuity of X, we have $X_{T_n+u} \to X_{T+u}$ and $X_{T_n} \to X_T$ as $n \to \infty$. Thus, if f and $P_u f$ are continuous and bounded, then 5.32 will hold for T. And, if 5.32 holds for f continuous, then it will hold for all f in $\mathcal{E}_+$ by a monotone class argument, and hence the strong Markov property. For this program to work, we need an assumption that the function $P_u f$ be continuous for f continuous, both regarded as functions on the compact space $\bar{E}$. We take this up next.

Feller processes

Let $\mathcal{C}_0 = C_0(E \mapsto \mathbb{R})$, the set of all continuous functions $f: E \mapsto \mathbb{R}$ with $\lim_{x\to\partial} f(x) = 0$. Elements of $\mathcal{C}_0$ are called continuous functions *vanishing at infinity*. These are functions on E whose automatic extensions (with $f(\partial) = 0$) onto $\bar{E}$ yield continuous functions on $\bar{E}$. Since $\bar{E}$ is compact, every such function is bounded. Every continuous function $\bar{f}: \bar{E} \mapsto \mathbb{R}$ has the form $\bar{f}(x) = f(x) + c$ for some f in $\mathcal{C}_0$ and some constant c, namely, $c = \bar{f}(\partial)$.

5.36 DEFINITION. *The Markov system $\mathfrak{X}$ is called a* Feller system *if*

5.37 $$f \in \mathcal{C}_0 \quad \Rightarrow \quad P_t f \in \mathcal{C}_0 \text{ for every } t \text{ in } \mathbb{R}_+.$$

5.38 REMARK. a) Since X is right-continuous, the condition 5.37 implies that

5.39 $$f \in \mathcal{C}_0 \Rightarrow \lim_{t\to 0} P_t f(x) = f(x), \qquad x \in E.$$

For, as $t \to 0$, we have $X_t \to X_0$, and $f \circ X_t \to f \circ X_0$ by the continuity of f; thus

$$P_t f(x) = \mathbb{E}^x\ f \circ X_t \to \mathbb{E}^x\ f \circ X_0 = f(x)$$

by the bounded convergence theorem and the normality of $\mathfrak{X}$.

b) In the absence of $\mathfrak{X}$, a sub-Markov semigroup (P_t) is said to satisfy the *Feller condition* if 5.37 and 5.39 hold. Given such a semigroup, and given E as in 4.1, it is possible to construct a system $\mathfrak{X}$ that satisfies 4.1–4.6 and is a Markov system in the sense of Definition 4.7. The construction is long and tedious. Following the modern sensibilities, we have defined the Markov system $\mathfrak{X}$ axiomatically, rather than treating the semigroup (P_t) (which is rarely explicit except for Wiener and Poisson) as the primary object. The next theorem shows that every Feller process is a Hunt process, that is, it is strong Markov and quasi-left-continuous.

5.40 Theorem. *If $\mathfrak{X}$ is a Feller system, then it is a Hunt system.*

Proof. Suppose that $\mathfrak{X}$ has the Feller property 5.37, we need to show that, then, $\mathfrak{X}$ is strong Markov and quasi-left-continuous.

a) For the first, we follow the program outlined in Remark 5.35. Let T be an $\mathcal{F}$-stopping time, choose stopping times T_n decreasing to T such that each T_n is countably-valued. By Lemma 5.33,

$$\mathbb{E}^x\ f \circ X_{T_n+u} = \mathbb{E}^x\ P_u\ f \circ X_{T_n}$$

for f in $\mathcal{E}_b$. Now let $f \in \mathcal{C}_0$ and let $n \to \infty$. We obtain 5.32 through the right-continuity of X, the continuity of f and $P_u f$ in $\mathcal{C}_0$ when extended onto $\bar{E}$, and the bounded convergence theorem. Finally, 5.32 extends to f in $\mathcal{E}_+$ by a monotone class argument. Thus, $\mathfrak{X}$ is strong Markov.

b) To show quasi-left-continuity, let T be a stopping time of $\mathcal{F}$, and (T_n) an increasing sequence of such times with limit T; we need to show that

5.41 $$\lim X_{T_n} = X_T \text{ almost surely on } \{T < \infty\}.$$

It is enough to show that it is so almost surely on $\{T \le b\}$ for every $b < \infty$; then, letting $b \to \infty$ over the integers yields the desired end. But, on $\{T \le b\}$, we have $T = T \wedge b$ and $T_n = T_n \wedge b$, which are all bounded stopping times. Thus, we may and do assume that T is bounded.

Since X is left-limited in $\bar{E}$, the limits

5.42 $$L = \lim_n X_{T_n}, \qquad L_u = \lim_n\ X_{T_n+u}$$

exist, the latter for every $u > 0$. For $u > 0$, we have $T_n(\omega) + u > T(\omega)$ for all n large enough; thus, $L_u \to X_T$ as $u \to 0$, by the right-continuity of X. Hence, for continuous $\bar{f}$ and $\bar{g}$ on $\bar{E}$,

5.43 $$\mathbb{E}^x\ \bar{f} \circ L \quad \bar{g} \circ X_T = \lim_{u\to 0}\ \lim_{n\to\infty} \mathbb{E}^x\ \bar{f} \circ X_{T_n} \quad \bar{g} \circ X_{T_n+u}.$$

We can write $\bar{g} = c + g$, where $g \in \mathcal{C}_0$ and $c = \bar{g}(\partial)$. Using the already proved strong Markov property at T_n, we see that the right side of 5.43 is equal to

$$\begin{aligned}\lim_{u \to 0} \lim_{n \to \infty} \ & \mathbb{E}^x \left(\bar{f} \circ X_{T_n}\right)(c + P_u g \circ X_{T_n}) \\ &= \lim_{u \to 0} \mathbb{E}^x \left(\bar{f} \circ L\right)(c + P_u g \circ L),\end{aligned}$$

where the last equality follows from the bounded continuity of $\bar{f}$ and $P_u g$, the latter being through the assumed Feller property 5.37 applied to g in $\mathcal{C}_0$. Thus, using Remark 5.38a to the effect that $P_u g \to g$ as $u \to 0$, we see that 5.43 becomes,

5.44 $$\mathbb{E}^x \bar{f} \circ L \quad \bar{g} \circ X_T = \mathbb{E}^x \ \bar{f} \circ L \quad \bar{g} \circ L.$$

Since continuous functions of the form $\bar{f} \times \bar{g}$ generate the Borel σ-algebra on $\bar{E} \times \bar{E}$, a monotone class argument applied to 5.44 shows that

$$\mathbb{E}^x \ \bar{h} \circ (L, X_T) = \mathbb{E}^x \ \bar{h} \circ (L, L)$$

for every bounded Borel function $\bar{h}$ on $\bar{E} \times \bar{E}$. Taking $\bar{h}$ to be the indicator of the diagonal of $\bar{E} \times \bar{E}$, and noting the definition of L in 5.42, we obtain the desired result 5.41. □

Markovian bestiary

Poisson processes are the quintessential Markov processes with jumps. Brownian motions are the continuous Markov processes par excellence. They are both Lévy processes.

All Lévy processes are Itô processes; the latter are processes that satisfy stochastic integral equations like 3.1, but with a further term that define a compensated sum of jumps; see 3.95. Itô diffusions, jump-diffusions, and simple processes are special cases of Itô processes.

All Itô processes are Feller processes. The latter are introduced through their transition functions, with conditions on how the transition kernels P_t treat continuous functions. From those conditions follow the real objectives: regularity properties of the sample paths, strong Markov property, quasi-left-continuity, etc.

All Feller processes are Hunt processes. The latter are introduced axiomatically by saying that we have a process and it has the following desirable properties. This is the straightforward approach; it puts the process as the central object, the axioms can be checked directly in practical situations or in the case of Itô processes.

All Hunt processes are "standard;" the latter allow quasi-left-continuity to fail at ζ, at the end of life. Finally, all standard processes are "right processes," the latter form a class of Markov processes that is invariant under certain useful transformations such as killing, time changes, spatial transformations.

These are the objects of the general theory of Markov processes. (See the notes for this chapter for references.)

There is a class of processes that is totally outside of all the previous classes: It consists of Markov processes (in continuous-time) with discrete state spaces, but *without* the sample path regularities such as right-continuity. When the state space is discrete (with the discrete topology), every right-continuous left-limited path is necessarily a step process; too simple, theoretically. On the other hand, on a general state space, it is impossible to build a theory without right-continuity etc. for the paths. But, with a discrete state space, it is possible to create a rich theory that allows sample paths to have discontinuities of the second type. Such processes should be called *Chung processes*.

Exercises and Complements

5.45 *Additive functionals.* Let $\mathfrak{X}$ be a Markov system with living space E. Let $f \in \mathcal{E}_{b+}$ and put

$$A_t = \int_0^t ds\ f \circ X_s, \qquad t \in \mathbb{R}_+.$$

Show that, for every ω,

$$A_{t+u}(\omega) = A_t(\omega) + A_u(\theta_t \omega), \qquad t, u \in \mathbb{R}_+;$$

then, A is said to be *additive.*

5.46 *Continuation.* Let $\mathfrak{X}$ be a Markov system. Let $A = (A_t)$ be an increasing right-continuous process with $A_o = 0$. It is said to be an *additive functional* of X if it is additive and is adapted to $\left(\mathcal{G}^\circ_{t+}\right)$. The preceding exercise gave an example of a continuous additive functional. If X is a Brownian motion, the local time at 0 is another example of a continuous additive functional. If X is a jump-diffusion as in Section 3, then

$$A_t = \sum_{s \le t} f \circ (X_{s-}, X_s)\ 1_{\{X_{s-} \neq X_s\}}, \quad t \in \mathbb{R}_+,$$

is an additive functional of the pure-jump type.

5.47 *Time changes.* Let $\mathfrak{X}$ be a Hunt system with living space E. Let $f : E \mapsto (0,1)$ be Borel and define

$$C_t = \int_0^t ds\ f \circ X_s, \qquad t \in \mathbb{R}_+.$$

Then, C is a strictly increasing continuous additive functional. Using C as a random clock, let S be its functional inverse (that is, $S_u = \inf\{t \ge 0 : C_t > u\},\ u \in \mathbb{R}_+$). Each S_t is a stopping time of $\left(\mathcal{G}^\circ_{t+}\right)$ and of $(\mathcal{F}_t)$. Define

$$\hat{X}_t = X_{S_t},\ \hat{\theta}_t = \theta_{S_t},\ \hat{\mathcal{F}}_t = \mathcal{F}_{S_t}, \quad t \in \mathbb{R}_+.$$

Show that $\hat{\mathfrak{X}} = (\Omega,\ \mathcal{H},\ \hat{\mathcal{F}},\ \hat{\theta},\ \hat{X},\ \mathbb{P}^\bullet)$ is again a Hunt system with living space E.

5.48 *Increasing continuous processes.* Let $\mathfrak{X}$ be a Hunt system with living space $E = \mathbb{R}_+$. Suppose that $t \mapsto X_t$ is increasing and continuous, and that $\zeta = +\infty$.

a) Show that $t \mapsto X_t$ is strictly increasing.

b) Put $C_t = X_t - X_0$. Show that C is an strictly increasing continuous additive functional of X.

c) Let $\hat{X}$ be the time-changed process. Note that $\hat{X}_t = X_0 + t$, deterministic, except for the initial state. Show that (S_t) is a continuous additive functional of $\hat{X}$. In particular, this means that S_t is determined by $\{\hat{X}_s : s \le t\}$.

d) Conclude that (S_t) and, therefore, (C_t) and (X_t) are deterministic except for the dependence on X_0. Here is the form of X: Let f be a continuous strictly increasing function on $\mathbb{R}_+$ with $f(0) = 0$. If $X_0(\omega) = x$, choose the unique time t_0 such that $f(t_0) = x$; Then $X_t(\omega) = f(t_0 + t)$ for all $t \ge 0$.

5.49 *Step processes.* Let $\mathfrak{X}$ be such that X is a step process; let $(Y_n), (T_n)$, and $k(x)$ be as in 3.69 *et seq.*, with no traps. Let

$$C_t = \int_0^t ds\ k \circ X_s, \qquad t \in \mathbb{R}_+,$$

and consider the process $\hat{X}$ obtained as in 5.47, but from this C.

a) Show that $\hat{X}$ has the form

$$\hat{X}_t = Y_n \qquad on \quad \left\{\hat{T}_n \le t < \hat{T}_{n+1}\right\},$$

where $\hat{T}_0 = 0$, and $\{\hat{T}_{n+1} - \hat{T}_n : n \in \mathbb{N}\}$ is an independency of standard exponential variables that is independent of (Y_n). Thus, $\hat{X}_t = Y_{N_t}$, where N is a standard Poisson process independent of Y.

5.50 *Continuation.* Let $\mathfrak{X}$ be as in Example 5.19 above, where X is a Brownian motion inside $(a,\ b)$ and has sojourns at a and b before jumping into $(a,\ b)$. Define

$$C_t = \int_0^t ds\ 1_{(a,b)} \circ X_s.$$

Note that C remains flat during sojourns of X. Now, C is still a continuous additive functional, but not strictly increasing. Define (S_t) and $\hat{\mathfrak{X}}$ as in the preceding exercise.

a) Show that $\hat{\mathfrak{X}}$ is a Markov system with living space $[a,\ b]$ except that the normality fails (for $\mathbb{P}^a$ and $\mathbb{P}^b$). Of course, the actual state space for $\hat{X}$ is the interval $(a,\ b)$; and $\hat{\mathfrak{X}}$ is a Markov system with living space $(a,\ b)$, since normality does hold for x in $(a,\ b)$ and $x = \partial$.

b) Describe the process $\hat{X}$.

c) Is $\hat{\mathfrak{X}}$ with living space (a,b) a Hunt process?

5.51 *Semimartingale Hunt processes.* Let X be an Itô process; see 3.95. Let $f, C, S, \hat{\mathfrak{X}}$ be as in Exercise 5.47 above. Then, $\hat{\mathfrak{X}}$ is a Hunt system as mentioned in 5.47. Moreover, every component $\hat{X}^i$ of $\hat{X}$ is a semimartingale.

This has a converse. Every Hunt process $\hat{X}$ whose components are semimartingales has the structure described. Somewhat more explicitly, let $\hat{X}$ be a Hunt process with state space $E = \mathbb{R}^d$. Then, there are deterministic measurable functions a, b, j, f and (on an enlargement of the original probability space) a Wiener process W and a Poisson random measure M such that $\hat{X}_t = X_{S_t}$ where X is an Itô process as described in 3.95 and S is defined from X and f as in 5.47. See the notes for this chapter.

6 Potentials and Excessive Functions

This section is independent of Section 5, and its dependence on the earlier sections is slight. Throughout, $\mathfrak{X} = (\Omega, \mathcal{H}, \mathcal{F}, \theta, X, \mathbb{P}^\bullet)$ is a Markov system with living space E and transition semigroup (P_t); see 4.1–4.6 and Definition 4.7 for these and the attendant conventions.

For f in $\mathcal{E}_b$ and $p > 0$, by the arguments of Example 4.18,

6.1 $$U_p f(x) = \mathbb{E}^x \int_0^\infty dt\; e^{-pt} \quad f \circ X_t, \qquad x \in \mathbb{E},$$

defines a function $U_p f$ in $\mathcal{E}_b$. The same makes sense for f in $\mathcal{E}_+$ and $p \geq 0$, and the result is a function $U_p f$ in $\mathcal{E}_+$. In both cases,

6.2 $$U_p f(x) = \int_E \quad U_p(x, dy)\; f(y), \qquad x \in E,$$

where

6.3 $$U_p(x, A) = \mathbb{E}^x \int_0^\infty dt\; e^{-pt}\; 1_A \circ X_t = \int_0^\infty dt\; e^{-pt}\; P_t(x, A).$$

If $p > 0$, then U_p is a bounded kernel: $U_p(x, E) \leq {}^1/_p$. When $p = 0$, writing the integral over t as a sum of integrals over $[n,\; n+1)$ shows that U_0 is Σ-bounded, but generally not σ-finite.

The function $U_p f$ is called the *p-potential* of *f*, and U_p the *p-potential kernel* or *p-potential operator* depending on the role it plays. The family $(U_p)_{p>0}$ of operators $U_p : \mathcal{E}_b \mapsto \mathcal{E}_b$ is called the *resolvent* of the semigroup (P_t) or of the Markov process X.

6.4 Theorem. *a) The resolvent equation*

6.5 $$U_p - U_q + (p - q)\, U_p\; U_q = 0$$

holds for $p, q > 0$; in particular, $U_p\; U_q = U_q\; U_p$.

b) For each $p > 0$, the kernel pU_p is sub-Markov; and

$$\lim_{p\to\infty} pU_p\; f(x) = f(x)$$

for every f in $\mathcal{E}_b$ that is continuous at the point x of E.

Proof. Follows from the same arguments as in 2.36–2.39 and Theorem 2.40. □

Potentials and supermartingales

6.6 THEOREM. *Let $p > 0$ and $f \in \mathcal{E}_{b+}$. Then, for each x in E,*

$$M_t = \int_0^t ds\; e^{-ps}\; f \circ X_s + e^{-pt}\; U_p\; f \circ X_t, \qquad t \in \mathbb{R}_+,$$

is a uniformly integrable $\mathcal{F}$-martingale over $(\Omega, \mathcal{H}, \mathbb{P}^x)$.

Proof. Define

6.7
$$A_t = \int_0^t ds \quad e^{-ps}\; f \circ X_s, \qquad t \in \mathbb{R}_+.$$

The process (A_t) is increasing, and the limit A_∞ is a bounded positive variable in $\mathcal{G}_\infty$ since $p > 0$ and f is bounded. Note that

$$A_\infty = A_t + \int_t^\infty ds\; e^{-ps}\; f \circ X_s = A_t + e^{-pt}\; A_\infty \circ \theta_t.$$

Thus, by the Markov property at t,

$$\mathbb{E}_t^x\; A_\infty = A_t + e^{-pt}\; \mathbb{E}^{X_t}\; A_\infty = A_t + e^{-pt}\; U_p f \circ X_t = M_t,$$

since $\mathbb{E}^y A_\infty = U_p f(y)$ by definition. Via Theorem V.5.13, this shows that M is a uniformly integrable martingale, with respect to $\mathcal{F}$, under every $\mathbb{P}^x$. □

6.8 COROLLARY. *Under each $\mathbb{P}^x$,*

$$V_t = e^{-pt}\; U_p f \circ X_t, \qquad t \in \mathbb{R}_+,$$

is a positive supermartingale with $\lim_{t\to\infty} V_t = 0$.

Proof. We have $V = M - A$ with the definitions in 6.6 and 6.7, and the process A is increasing. Thus, V is a supermartingale. And, $\lim_{t\to\infty} V_t = 0$, because, by the martingale convergence theorem,

$$\lim_{t\to\infty} M_t = \lim_t\; \mathbb{E}_t^x\; A_\infty = A_\infty = \lim_{t\to\infty} A_t.$$ □

In martingale terminology, the process (V_t) is a potential; see V.4.53. The decomposition

$$V = M - A$$

is an instance of *Doob-Meyer decomposition* for supermartingales, which is the continuous-time version of Doob's decomposition given in Theorem V.3.2.

Going back to Theorem 6.6, the uniform integrability of M implies that M is a Doob martingale on $[0, \infty]$, and thus

$$\mathbb{E}^x\, M_T = \mathbb{E}^x M_0 = U_p f(x)$$

for every stopping time T of $\mathcal{F}$. This proves the following corollary to 6.6.

6.9 THEOREM. *Let $p > 0$ and $f \in \mathcal{E}_{b+}$. Then, for every x in E,*

$$U_p f(x) = \mathbb{E}^x \int_0^T ds\; e^{-ps}\; f \circ X_s \;+\; \mathbb{E}^x\; e^{-pT}\; U_p f \circ X_T$$

for every $\mathcal{F}$-stopping time T. In particular, if $f \circ X_s = 0$ on $\{s < T\}$, then

$$\mathbb{E}^x\; e^{-pT}\; U_p\; f \circ X_T = U_p\; f(x), \quad x \in E.$$

The particular case is useful in computing the distributions of T and X_T by choosing f appropriately. The theorem is the potential counterpart of Dynkin's formula using generators; see Theorem 2.27 for Dynkin's formula for Itô diffusions.

Excessive functions

6.10 DEFINITION. *Let $p \in \mathbb{R}_+$. A function f in $\mathcal{E}_+$ is said to be p-excessive if*

a) $f \geq e^{-pt}\; P_t f$ for every t in $\mathbb{R}_+$, and

b) $\lim_{t \to 0} e^{-pt}\; P_t f(x) = f(x)$ for every x in E.

The condition (a) is called the *p-supermedian property* for f; other terms in use are *p-super-mean-valued*, *p-superaveraging*. It implies, via the semigroup property $P_t P_u = P_{t+u}$, that the mapping $t \mapsto e^{-pt} P_t f(x)$ is decreasing. Hence, the limit in (b) is an increasing limit, and the condition (b) can be written as

6.11
$$\sup_t\; e^{-pt}\; P_t\; f(x) = f(x).$$

If the conditions (a) and (b) hold for Borel f, without requiring that f be positive, then f is said to be *p-superharmonic*. In all this, when $p = 0$, it is dropped both from notation and terminology. The following is the connection to supermartingales.

6.12 PROPOSITION. *Let $p \geq 0$. Let f be p-supermedian. Then, for each x in E with $f(x) < \infty$, the process*

$$Y_t = e^{-pt}\; f \circ X_t, \qquad t \in \mathbb{R}_+,$$

is an $\mathcal{F}$-supermartingale over $(\Omega, \mathcal{H}, \mathbb{P}^x)$.

Proof. Obviously, Y is adapted to $\mathcal{F}$. Fix x such that $f(x) < \infty$. Then, since f is p-supermedian,

$$f(x) \geq e^{-pt}\ P_t f(x) = \mathbb{E}^x\ e^{-pt}\ f \circ X_t = \mathbb{E}^x\ Y_t,$$

showing that Y_t is integrable under $\mathbb{P}^x$. And, by the Markov property of X,

$$\begin{aligned}\mathbb{E}_t^x\ Y_{t+u} &= e^{-p(t+u)}\ \mathbb{E}_t^x\ f \circ X_{t+u} \\ &= e^{-pt}\ e^{-pu}\ P_u\ f \circ X_t \leq e^{-pt}\ f \circ X_t = Y_t,\end{aligned}$$

where the inequality is via the p-supermedian property of f. □

In the preceding proposition, if f is p-excessive, it can be shown that Y is right-continuous. Thus, p-excessive functions are, roughly speaking, continuous over the paths of X.

Potentials are excessive

6.13 Proposition. *Let $p \geq 0$ and $f \in \mathcal{E}_+$. Then, $U_p f$ is p-excessive.*

Proof. Clearly, $U_p f \in \mathcal{E}_+$. Also, by Fubini's theorem,

$$e^{-pt}\ P_t\ U_p\ f = e^{-pt} \int_0^\infty du\ e^{-pu}\ P_t\ P_u\ f = \int_t^\infty ds\ e^{-ps}\ P_s f.$$

The last integral is dominated by $U_p f$ and increases to $U_p f$ as t decreases to 0. Hence, $U_p f$ is p-excessive. □

The following is one-half of a theorem characterizing excessive functions in terms of the resolvent. But it is sufficient for our purposes.

6.14 Proposition. *Let $p \geq 0$ and let f be p-supermedian. Then $q \mapsto qU_{p+q} f$ is increasing and dominated by f. Its limit is f as $q \to \infty$ if f is p-excessive.*

Proof. Fix $p \geq 0$. For $q > 0$,

$$q\ U_{p+q}\ f = \int_0^\infty dt\ q\ e^{-qt}\ e^{-pt}\ P_t f = \int_0^\infty du\ e^{-u} e^{-pu/q}\ P_{u/q}\ f.$$

As q increases, u/q decreases and the integrand increases by the p-supermedian property of f. By the same property, the last integrand is dominated by $e^{-u} f$, and hence, the integral is dominated by f. Finally, if f is p-excessive, the integrand increases to $e^{-u} f$ as $q \to \infty$, and the monotone convergence theorem implies that the integral becomes f in the limit. □

Approximation by bounded potentials

Let f be p-excessive. Then, for each n, the function $nU_{p+n}f = U_{p+n}(nf)$ is a potential, and as $n \to \infty$ the limit is f. So, every p-excessive function is the limit of an increasing sequence of potentials. The following sharpens the result when $p > 0$.

6.15 THEOREM. *Let $p > 0$. Let f be p-excessive. Then, there exists a sequence (g_n) in $\mathcal{E}_{b+}$ such that the sequence $(U_p g_n)$ increases to f.*

Proof. For each integer $n \geq 1$, put $f_n = f \wedge n$; each f_n is bounded and p-supermedian (since f is such and the constant n is such a function). By the resolvent equation 6.5,

$$U_{p+q}\ f_n = U_p\ f_n - q\ U_p\ U_{p+q}\ f_n;$$

the right side is well-defined as the difference of two bounded functions since f_n is bounded and $p > 0$. Thus, with $g_n = n(f_n - nU_{p+n}f_n)$, we have

6.16 $$nU_{p+n}\ f_n = U_p\ g_n.$$

Since f_n is p-supermedian, Proposition 6.14 yields $f_n \geq nU_{p+n}f_n$, and hence $g_n \in \mathcal{E}_{b+}$ for every n. There remains to show that the left side of 6.16 increases to f as $n \to \infty$. To that end, we note that (f_n) is increasing to f, and that $qU_{p+q}f_n$ is increasing in q, since f_n is p-supermedian (see Proposition 6.14). Thus, the left side of 6.16 is increasing in n, and

$$\lim_n\ nU_{p+n}\ f_n = \lim_q\ \lim_n\ qU_{p+q}\ f_n = \lim_q\ q\ U_{p+q}\ f = f,$$

the last equality being via Proposition 6.14 applied to the p-excessive function f. □

Supermedian property at stopping times

This is essentially Doob's stopping theorem for the supermartingale Y of Proposition 6.12. See Exercises 6.23–6.26 for its interpretation in optimal stopping games.

6.17 THEOREM. *Let $p \geq 0$. Let f be p-excessive. Then, for every $\mathcal{F}$-stopping time T,*

6.18 $$f(x)\ \geq\ \mathbb{E}^x\ e^{-pT}\ f \circ X_T, \qquad x \in E.$$

Proof. Suppose that $p > 0$. Let (g_n) be as in Theorem 6.15, so that $U_p g_n \nearrow f$. By Theorem 6.9,

$$U_p\ g_n(x) \geq \mathbb{E}^x\ e^{-pT}\ U_p g_n \circ X_T.$$

Letting $n \to \infty$ we obtain 6.18 when $p > 0$.

If $p = 0$ and f is excessive, then f is p-excessive for every $p > 0$, and thus 6.18 holds for every $p > 0$. Let p decrease to 0 strictly. By the monotone convergence theorem,

$$f(x) \geq \lim_{p\to 0} \mathbb{E}^x \, e^{-pT} \, f \circ X_T = \mathbb{E}^x \, 1_{\{T<\infty\}} f \circ X_T = \mathbb{E}^x \, f \circ X_T,$$

since $X_T = \partial$ on $\{T = \infty\}$ and $f(\partial) = 0$.

Exercises

6.19 *Poisson process.* Let $X = X_0 + N$ where N is a Poisson process with rate c; we take $E = \mathbb{R}_+$. Show that

$$Uf(x) = \mathbb{E}^x \int_0^\infty dt \, f \circ X_t = \frac{1}{c} \sum_{j=0}^\infty f(x+j), \qquad x \in E,$$

for every f positive Borel, more generally, for $p \geq 0$, show that

$$U_p f(x) = \frac{1}{c+p} \sum_{j=0}^\infty \left(\frac{c}{c+p}\right)^j f\,(x+j)\,.$$

6.20 *Stable processes.* Suppose that $X = X_0 + S$, where S is an increasing stable process with index a in (0,1). Suppose that its Lévy measure is given as $\lambda(dx) = (c/x^{a+1})dx$, with $c = a/\Gamma(1-a)$; see Example VII.7.6b. Show that, with $E = \mathbb{R}_+$,

$$Uf(x) = \frac{1}{\Gamma(a)} \int_x^\infty dy \, (y-x)^{a-1} \, f(y), \qquad x \in E.$$

6.21 *Brownian motion.* Suppose that $X = X_0 + W$, the standard Brownian motion in $\mathbb{R}^d$.

a) For $d = 1$ or $d = 2$, show that

$$Uf\;(x) = +\infty, \qquad x \in \mathbb{R}^d,$$

for every f positive Borel on $\mathbb{R}^d$, except for $f = 0$ in which case $Uf = 0$.

b) Show, if $d \geq 3$, that

$$Uf(x) = \frac{\Gamma\left(\frac{d}{2}-1\right)}{2\;\pi^{d/2}} \int_{\mathbb{R}^d} dy \, |y-x|^{2-d} \; f(y)$$

for all x in $\mathbb{R}^d$ and positive Borel f on R^d. Thus, except for multiplication by a constant, Uf is the Newtonian potential of f in classical potential theory.

6.22 *Excessive functions.* For the Markov system $\mathfrak{X}$, prove the following.

a) If $f = c$ for some constant $c \geq 0$, then f is p-excessive for every $p \geq 0$.

b) If f is p-excessive, then it is q-excessive for every $q \geq p$.

c) If f is p-excessive and $c \geq 0$ is a constant, then cf is p-excessive. If f and g are p-excessive, then so is $f + g$.

d) If (f_n) is an increasing sequence of p-excessive functions with limit f, then f is p-excessive. Hint: $e^{-pt}\ P_t\ f_n$ is increasing in n, and increasing with decreasing t.

e) If f and g are p-supermedian, then so is $f \wedge g$.

6.23 *Brownian motion on* $\mathbb{R}$. Let X be a standard Brownian motion on $\mathbb{R}$. Let f be excessive ($p = 0$). Show that $f = c$ for some constant $c \geq 0$. Hint: Use 6.18 with $T = T_y$, the hitting time of the point y.

6.24 *Continuation.* Let $\hat{X}$ be a standard Brownian motion on $\mathbb{R}$, and let $X_t = \hat{X}_{\tau \wedge t}$, where τ is the time of exit for $\hat{X}$ from the fixed interval (a,b). Thus, X lives in $E = [a, b]$, and the boundary points a and b are traps. Show that every excessive function for X is a concave function on $[a,b]$. Hint: Recall the formula for $\mathbb{E}^x\ f \circ X_T$ for $y < x < z$ and T the time of exit from the interval $(y, z) \subset [a, b]$. Fix y and z, take $x = \alpha y + (1 - \alpha)z$ for $0 \leq \alpha \leq 1$.

6.25 *Optimal stopping.* We are to receive a one-time reward of $f \circ X_T$ dollars if we choose time T to ask for the reward. We want to choose a stopping time T_o that maximizes

$$\mathbb{E}^x\ e^{-rT}\ f \circ X_T, \qquad T \ \text{ is an } \mathcal{F}\text{-stopping time},$$

if possible, or, if this proves impossible, come close to the value

$$v(x) = \sup_T \mathbb{E}^x\ e^{-rT}\ f \circ X_T,$$

where the supremum is over all $\mathcal{F}$-stopping times. Here, f is a positive Borel function on E, called the payoff function. We interpret r as the interest rate, and v is called the value of the game.

a) If f is r-excessive, then $v = f$ and $T_0 = 0$ is an optimal stopping time.

b) In general, v is the minimal r-excessive function that dominates f.

6.26 *Continuation.* Suppose that X is the standard Brownian motion on $\mathbb{R}$. Let f be a bounded positive function on $\mathbb{R}$, and take $r = 0$. Show that $v = c$, no computations needed, where

$$c = \sup_{y \in \mathbb{R}}\ f(y).$$

If the supremum is attained, that is, there exists x^* in $\mathbb{R}$ such that $f(x^*) = c$, then the time T_0 of hitting x^* is an optimal stopping time. If $f(x) = 1 - e^{-x}$ for $x > 0$ and is 0 otherwise, then $v(x) = c = 1$ for all x in $\mathbb{R}$; but there is no optimal stopping time; recall that $X_\infty = \partial$ and $f(\partial) = 0$ – the dead pay nothing. As this example indicates, there might be no optimal stopping time. But, for every $\varepsilon > 0$ there is a stopping time T_ε such that

$$\mathbb{E}^x\ f \circ X_{T_\varepsilon} \geq v(x) - \varepsilon.$$

7 Appendix: Stochastic Integration

This is a quick introduction to stochastic calculus. It is driven by the needs of Section 2 on Itô diffusions. We limit ourselves mostly to continuous processes and omit almost all proofs.

Throughout, $(\Omega, \mathcal{H}, \mathbb{P})$ is a complete probability space, and $\mathcal{F} = (\mathcal{F}_t)$ is an augmented right-continuous filtration. All processes are adapted to this $\mathcal{F}$, without further mention. Also, all processes have the state space $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$. We introduce the term *Stieltjes process* as a short substitute for a process whose almost every path is right-continuous, is left-limited, and has bounded total variation over every bounded interval. Then, to repeat Definition V.5.18, a process X is a semimartingale if it can be written as the sum of a local martingale and a Stieltjes process.

Stochastic integrals

For our current purposes, we call $\sigma = (t_i)$ a *subdivision* of $\mathbb{R}_+$ if $0 = t_0 < t_1 < \cdots$ and $\lim t_n = +\infty$. A subdivision of $[0,t]$ is a finite sequence $\sigma = (t_0,\ t_1, \ldots, t_n)$ with $0 = t_0 < t_1 < \cdots < t_n = t$. In both cases, $\|\sigma\| = \sup_i\, (t_{i+1} - t_i)$ is called the mesh of σ.

Let F be a left-continuous process, and X continuous. For every subdivision $\sigma = (t_i)$ of $\mathbb{R}_+$, we define a new process Y^σ by putting $Y_0^\sigma = 0$ and

7.1 $$Y_t^\sigma = \sum_{i=1}^{j} F_{t_{i-1}} \cdot \left(X_{t_i} - X_{t_{i-1}}\right) + F_{t_j} \cdot \left(X_t - X_{t_j}\right) \qquad \text{if } t_j < t \le t_{j+1}.$$

Then, Y^σ is a continuous process. Note the resemblance to V.3.4, the integral in discrete time. We omit the proof of the following fundamental result.

7.2 Theorem. *Let F be a left-continuous process, and X a continuous semimartingale. Then, there exists a unique process Y such that*

$$\lim_{\|\sigma\|\to 0} \mathbb{P}\left\{ \sup_{0\le t\le u} |Y_t^\sigma - Y_t| > \varepsilon \right\} = 0$$

for every $\varepsilon > 0$ and $u < \infty$. The process Y is a continuous semimartingale. □

7.3 Definition. *The process Y of the preceding theorem is called the* stochastic integral *of F with respect to X, and the notations*

$$\int F\, dX \quad \text{and} \quad \int_0^t F_s\, dX_s$$

are used, respectively, for the process Y and the random variable Y_t.

7.4 REMARK. a) The theorem can be re-phrased: as $\|\sigma\| \to 0$, $Y_t^\sigma \to Y_t$ in probability, uniformly in t over compacts. The uniqueness of Y is up to indistinguishability.

b) If X is a Stieltjes process, then the stochastic integral coincides with the path-by-path ordinary integral, that is, for almost every ω, the number $Y_t(\omega)$ is the Riemann-Stieltjes integral of the function $s \mapsto F_s(\omega)$ with respect to the bounded variation function $s \mapsto X_s(\omega)$ over the interval $[0,t]$. Of course, then Y is a continuous Stieltjes process.

c) If X is not Stieltjes, if $X = W$ Wiener for instance, then $Y_t(\omega)$ is not the limit of $Y_t^\sigma(\omega)$ with ω held fixed. In fact, in most cases, $\lim_{\|\sigma\|\to 0} Y_t^\sigma(\omega)$ will not exist.

7.5 EXAMPLE. *Wiener driven integrals.* Suppose that F is left-continuous and bounded, and $X = W$, a Wiener process. Then, Y is an L^2-martingale and

7.6
$$\mathbb{E}\left|\int_0^t F_s \, dW_s\right|^2 = \mathbb{E}\int_0^t |F_s|^2 \, ds.$$

Here is the explanation. Given a subdivision $\sigma = (t_i)$, define the left-continuous step process F^σ by letting $F_t^\sigma = F_{t_{i-1}}$ for $t_{i-1} < t \le t_i$, and $F_0^\sigma = F_0$. Note that, in fact, Y^σ is the integral of F^σ with respect to X by every reasonable definition of integration. It is evident from 7.1 that Y^σ is now a martingale with

7.7
$$E\,|Y_t^\sigma|^2 = \mathbb{E}\int_0^t |F_s^\sigma|^2 \, ds;$$

this is an easy computation recalling that the increments $W_{t_i} - W_{t_{i-1}}$ are independent with mean 0 and variances $t_i - t_{i-1}$. In fact, since F is bounded, $Y_t^\sigma \to Y_t$ in the sense of L^2-convergence as $\|\sigma\| \to 0$. And, $F^\sigma \to F$ by the left-continuity of F. Thus, letting $\|\sigma\| \to 0$ on both sides of 7.7 we obtain 7.6.

Arithmetic of integration

Stochastic integrals are the same as ordinary integrals in linearity etc. The next proposition shows them; proofs are immediate from 7.1–7.3.

7.8 THEOREM. *Let F and G be left-continuous processes, X and Y continuous semimartingales, and a and b constants. Then,*

$$\int (aF + b\,G)\, dX = a\int F\, dX + b\int G\, dX,$$
$$\int F\, d(aX + bY) = a\int F\, dX + b\int F\, dY,$$
$$Y = \int F\, dX \Rightarrow \int G\, dY = \int (F\cdot G)\, dX.$$

7.9 REMARK. Let X be a continuous semimartingale. Then,

7.10 $$X = L + V,$$

where L is a continuous local martingale and V is a continuous Stieltjes process. For F left-continuous, then

7.11 $$\int F\, dX = \int F\, dL + \int F\, dV,$$

and, on the right side, the first term is a continuous local martingale, and the second is a continuous Stieltjes. In particular, if L is a martingale and F is bounded, then the first term is a martingale; Example 7.5 is a special case.

Cross variation, quadratic variation

Given processes X and Y, and a subdivision $\sigma = (t_i)$ of $\mathbb{R}_+$, let

7.12 $$C_t^\sigma = \sum_{t_i<t} \left(X_{t_{i+1}} - X_{t_i}\right)\left(Y_{t_{i+1}} - Y_{t_i}\right), \qquad t \in \mathbb{R}_+.$$

When $X = Y = W$, Wiener, we have seen in Theorem VIII.7.2 that $C_t^\sigma \to t$ in probability as $\|\sigma\| \to 0$. The following is the general case.

7.13 THEOREM. *Let X and Y be continuous semimartingales. Then, there is a continuous Stieltjes process C such that $C_t^\sigma \to C_t$ in probability for every t in the limit as $\|\sigma\| \to 0$.* □

The process C of the preceding theorem is called the *cross variation* of X and Y, and the notation $\langle X, Y\rangle$ is employed for it, that is,

7.14 $$\langle X, Y\rangle_t = C_t, \qquad t \in \mathbb{R}_+.$$

In particular, $\langle X, X\rangle$ is called the *quadratic variation* for the continuous semimartingale X; it is an increasing process in view of 7.12. The approximation 7.12 shows as well that

7.15 $$\langle X + Y,\ X + Y\rangle = \langle X, X\rangle + 2\langle X, Y\rangle + \langle Y, Y\rangle.$$

Solving this for $\langle X, Y\rangle$, since all other terms are increasing processes, we see that $\langle X, Y\rangle$ is indeed the difference of two increasing processes (as a Stieltjes process must be).

7.16 PROPOSITION. *Let X and Y be continuous semimartingales; if X or Y is Stieltjes, then $\langle X, Y\rangle = 0$. In particular, if $X = L + V$ as in 7.10, then $\langle X, X\rangle = \langle L, L\rangle$.*

Proof. Suppose Y is Stieltjes. From 7.12, we have

$$|C_t^\sigma| \le \sup_{t_j<t} \left|\ X_{t_{j+1}} - X_{t_j}\ \right| \sum_{t_i<t} \left|Y_{t_{i+1}} - Y_{t_i}\right|.$$

As $\|\sigma\| \to 0$, the supremum goes to 0 by the continuity of X, and the sum goes to the total variation of Y over $[0,t]$. The latter is finite since Y is assumed to be Stieltjes. Thus, $C_t^\sigma \to 0$, that is, $\langle X, Y\rangle = 0$. The particular statement follows from 7.15 for $\langle L+V,\ L+V\rangle$, because $\langle L, V\rangle = \langle V, V\rangle = 0$ since V is continuous Stieltjes.

Stochastic differentials

In analogy with ordinary calculus, we write

7.17 $$dY = F\ dX \ \Leftrightarrow\ Y_t = Y_0 + \int_0^t F_s\ dX_s, \quad t \in \mathbb{R}_+.$$

Similarly, in view of 7.12–7.14, we introduce the notation

7.18 $$dX\ dY = d\langle X, Y\rangle.$$

In particular, $dX\ dX = (dX)^2$ becomes the notation for the differential of the increasing continuous process $\langle X, X\rangle$. Next are the rules of stochastic differential calculus.

7.19 THEOREM. *Let F and G be left-continuous processes, X and Y continuous semimartingales, and a and b constants. Then,*

a) $d(aX + bY) = a\ dX + b\ dY$,

b) $F \cdot (dX + dY) = F\ dX + b\ dY$,

c) $(aF + b\ G)\ dX = a\ F\ dX + b\ G\ dX$,

d) $F \cdot (G\ dX) = (F \cdot G)\ dX$,

e) $(F\ dX)(G\ dY) = (F \cdot G)\ dX\ dY$, *and*

f) if X or Y is Stieltjes, then $dX\ dY = 0$.

Proof. (a) is direct from the definitions; (b), (c), (d) are the differential versions of the properties listed in Theorem 7.8; (e) follows by a simple computation from 7.12 upon replacing X there with $\int F\ dX$, and Y with $\int G\ dY$; finally, (f) is a re-statement of Proposition 7.16.

Itô's formula

This is the chain rule of differentiation for stochastic calculus. Recall that $C^2(\mathbb{R}^d \mapsto \mathbb{R})$ is the class of function $f : \mathbb{R}^d \mapsto \mathbb{R}$ that are twice continuously differentiable, and that we write $\partial_i f(x)$ for $\frac{\partial}{\partial x_i} f(x)$, and $\partial_{ij} f(x)$ for $\frac{\partial^2}{\partial x_i \partial x_j} f(x)$. When $d = 1$, we write f' for the first derivative, and f'' for the second. The next is Itô's formula for continuous semimartingales.

7.20 THEOREM. *Let $X^1, \ldots, X^n$ be continuous semimartingales, and put $X = (X^1, \ldots, X^n)$. For f in $C^2(\mathbb{R}^n \mapsto \mathbb{R})$, then, $f \circ X$ is a semimartingale, and*

$$d(f \circ X) = \sum_{i=1}^{n} (\partial_i f \circ X)\ dX^i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} (\partial_{ij} f \circ X)\ dX^i\, dX^j.$$

7.21 REMARK. a) For $n = 1$, Itô's formula becomes

$$d(f \circ X) = (f' \circ X)\ dX + \frac{1}{2} (f'' \circ X)\ (dX)^2.$$

b) If $X^1, \ldots, X^n$ are continuous Stieltjes processes, then $dX^i\ dX^j = 0$ for all i and j, and Itô's formula becomes the chain rule of differential calculus:

$$d\,(f \circ X) = \sum_{i=1}^{n} (\partial_i f \circ X)\ dX^i.$$

c) We re-state the conclusion of the theorem above in the formal notation of stochastic integrals:

$$f \circ X_t = f \circ X_0 + \sum_{i=1}^{n} \int_0^t (\partial_i f \circ X_s)\ dX_s^i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \int_0^t (\partial_{ij} f \circ X_s)\, d\langle X^i, X^j \rangle_s.$$

d) Taking $n = 2$ and $f(x, y) = xy$ in the preceding formula, we obtain the following formula for integration by parts for continuous semimartingales X and Y:

$$X_t Y_t = X_0 Y_0 + \int_0^t X_s\ dY_s + \int_0^t Y_s\ dX_s + \langle X, Y \rangle_t.$$

Wiener driven integrals

7.22 LEMMA. *Let X and Y be independent Wiener processes. Then, $dX\ dY = 0$.*

Proof. We have $\langle X, X \rangle_t = \langle Y, Y \rangle_t = t$ by Theorem VIII.7.2. By the same theorem, since $X + Y = \sqrt{2}\, W$ for some Wiener W, we have $\langle X+Y, X+Y \rangle_t = 2t$. The claim now follows from 7.15. □

Adding the preceding lemma to Itô's formula proves the following.

7.23 THEOREM. *Let $W^1, \ldots, W^n$ be independent Wiener processes and put $W = (W^1, \ldots, W^n)$. For f in $C^2(\mathbb{R}^n \mapsto \mathbb{R})$, then,*

$$d\,(f \circ W_t) = \sum_{i=1}^{n} (\partial_i f \circ W_t)\, dW_t^i + \frac{1}{2} \sum_{i=1}^{n} (\partial_{ii} f \circ W_t)\ dt.$$

Characterizations for Wiener processes

Recall the martingale characterization for Wiener processes; see Theorem V.2.19 and Proposition V.6.21: The processes W and $Y = \left(W_t^2 - t\right)$ are continuous martingales with $W_0 = Y_0 = 0$ if and only if W is a Wiener process. Itô's formula in 7.20 with $n = 1$ and $f(x) = x^2$ identifies the process Y,

$$W_t^2 - t = 2 \int_0^t W_s \, dW_s,$$

and shows, furthermore, that $\left(W_t^2 - t\right)$ is a martingale if and only if $\langle W, W \rangle_t = t$. This last property characterizes Wiener processes among all continuous local martingales:

7.24 THEOREM. *Let X be a continuous local martingale with $X_0 = 0$. Then, X is a Wiener process if and only if $\langle X, X \rangle_t = t$ for all $t \geq 0$.*

Proof. Necessity is by Theorem VIII.7.2. We show the sufficiency next. Suppose that $\langle X, X \rangle_t = t$ for all t. Then, by Itô's formula with $n = 1$ and f in $C^2(\mathbb{R} \longmapsto \mathbb{R})$,

$$f \circ X_t = f(0) + \int_0^t (f' \circ X_s) \, dX_s + \frac{1}{2} \int_0^t (f'' \circ X_s) \, ds.$$

Assuming further that f, f', f'' are bounded, the stochastic integral term on the right side defines a martingale. This is the content of Lemma V.6.22, and the proof of Proposition V.6.21 applies to show that X is a Wiener process. □

The next theorem is the n-dimensional version of the preceding. The necessity part of its proof is by Lemma 7.22; we omit the proof of sufficiency (it is similar to that of V.6.21).

7.25 THEOREM. *Let $X^1, \ldots, X^n$ be continuous local martingales with $X_0^i = 0$ for every i. Then, $X^1, \ldots, X^n$ are independent Wiener processes if and only if*

$$\langle X^i, X^j \rangle_t = I\,(i,j)\;\; t, \qquad t \in \mathbb{R}_+,$$

where I is the identity matrix in n-dimensions.

Itô's formula and the characterization above in terms of cross variations form a summary of stochastic integrals driven by Wiener processes.

Local martingales as stochastic integrals

If $dX = F \, dW$, then $(dX)^2 = F^2 \, (dW)^2 = F^2 \, dt$. The following theorem provides a converse as well.

7.26 THEOREM. *Let X be a continuous local martingale. Suppose that $(dX_t)^2 = (F_t)^2 \, dt$ for some left-continuous process F. Then, there is a*

Wiener process W (possibly on an enlargement of the original probability space) such that

$$X_t = X_0 + \int_0^t F_s \, dW_s$$

Proof. For the case $F > 0$. Then, since $1/F$ is left-continuous,

$$dW = \frac{1}{F} \, dX, \qquad W_0 = 0,$$

defines a continuous local martingale W. Since $(dW)^2 = (1/F)^2 \, (dX)^2 = dt$ by the assumption on $(dX)^2$, we see from the characterization theorem 7.24 that W is Wiener. Obviously, $dX = F \, dW$ as claimed. □

Local martingales are time changed Wieners

7.27 THEOREM. *Let X be a continuous local martingale. Let $C = \langle X, X \rangle$. Then, there is a Wiener process W such that*

$$X_t = X_0 + W_{C_t}, \qquad t \in \mathbb{R}_+.$$

Proof is omitted, but its essentials can be seen in Figure 17 below. If time is reckoned with the random clock $C = \langle X, X \rangle$, then $X - X_0$ appears as a Wiener process.

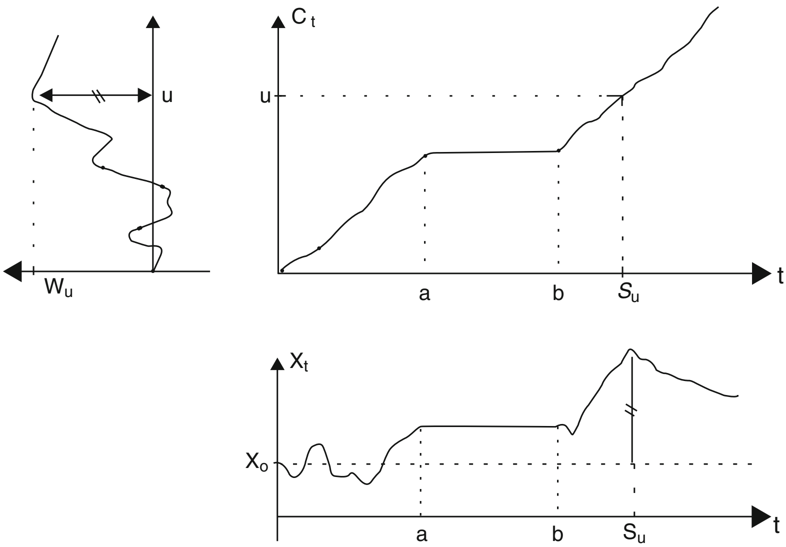

Figure 17: Using $C = \langle X, X \rangle$ as a random clock converts the local martingale X into a Wiener process. Conversely, X is obtained from the Wiener process by reversing the procedure.

For the outcome ω pictured in Figure 17, the path $t \mapsto C_t(\omega)$ remains flat over the interval $[a(\omega),\ b(\omega)]$, and then, $t \mapsto X_t(\omega) = W_{C_t(\omega)}(\omega)$ must remain constant over the same interval. This observation proves the following.

7.28 Proposition. *Let X be a continuous local martingale. Suppose that it is also a Stieltjes process. Then, for almost every ω, we have $X_t(\omega) = X_0(\omega)$ for all $t \geq 0$.* □

To put it another way, if X is a continuous local martingale and shows some signs of life, then its total variation must be infinite over some intervals.