2

Stochastic Processes

2.1 Definition

We commence along the lines of the founding work of Kolmogorov by regarding stochastic processes as a family of random variables defined on a probability space and thereby define a probability law on the set of trajectories of the process. More specifically, stochastic processes generalize the notion of (finite-dimensional) vectors of random variables to the case of any family of random variables indexed in a general set T. Typically, the latter represents "time" and is an interval of $\mathbb{R}$ (in the continuous case) or $\mathbb{N}$ (in the discrete case). For a nice and elementary introduction to this topic the reader may refer to Parzen (1962).

Definition 2.1. Let $(\Omega, \mathcal{F}, P)$ be a probability space, T an index set, and $(E, \mathcal{B})$ a measurable space. An $(E, \mathcal{B})$-valued *stochastic process* on $(\Omega, \mathcal{F}, P)$ is a family $(X_t)_{t\in T}$ of random variables $X_t : (\Omega, \mathcal{F}) \to (E, \mathcal{B})$ for $t \in T$.

$(\Omega, \mathcal{F}, P)$ is called the underlying *probability space* of the process $(X_t)_{t\in T}$, while $(E, \mathcal{B})$ is the *state space* or *phase space*. Fixing $t \in T$, the random variable X_t is the *state of the process at "time" t*. Moreover, for all $\omega \in \Omega$, the mapping $X(\cdot, \omega) : t \in T \to X_t(\omega) \in E$ is called the *trajectory* or *path of the process* corresponding to ω. Any trajectory $X(\cdot, \omega)$ of the process belongs to the space E^T of functions defined in T and valued in E. Our aim is to introduce a suitable σ-algebra $\mathcal{B}^T$ on E^T that makes the family of trajectories of our stochastic process a random function $X : (\Omega, \mathcal{F}) \to (E^T, \mathcal{B}^T)$.

More generally, let us consider the family of measurable spaces $(E_t, \mathcal{B}_t)_{t\in T}$ (as a special case, all E_t may coincide with a unique E) and define $W^T = \prod_{t\in T} E_t$. If $S \in \mathcal{S}$, where $\mathcal{S} = \{S \subset T |\ \text{S is finite}\}$, then the product σ-algebra $\mathcal{B}^S = \bigotimes_{t\in S} \mathcal{B}_t$ is well defined as the σ-algebra generated by the family of rectangles with sides in $\mathcal{B}_t$, $t \in S$.

Definition 2.2. If $A \in \mathcal{B}^S$, $S \in \mathcal{S}$, then the subset $\pi_{ST}^{-1}(A)$ is a *cylinder* in W^T with base A, where π_{ST} is the canonical projection of W^T on W^S.

V. Capasso and D. Bakstein, *An Introduction to Continuous-Time Stochastic Processes*, MSSET, DOI 10.1007/978-0-8176-8346-7_2,

It is easy to show that if C_A and $C_{A'}$ are cylinders with bases $A \in \mathcal{B}^S$ and $A' \in \mathcal{B}^{S'}$, $S, S' \in \mathcal{S}$, respectively, then $C_A \cap C_{A'}$, $C_A \cup C_{A'}$, and $C_A \setminus C_{A'}$ are cylinders with base in $W^{S \cup S'}$. From this it follows that the set of cylinders with a finite-dimensional base is a *ring* of subsets of W^T (or, better, an *algebra*). We denote by $\mathcal{B}^T$ the σ-algebra generated by it (See, e.g., Métivier 1968).

Definition 2.3. The measurable space $(W^T, \mathcal{B}^T)$ is called the *product space of the measurable spaces* $(E_t, \mathcal{B}_t)_{t \in T}$.

From the definition of $\mathcal{B}^T$ we have the following result.

Theorem 2.4. *$\mathcal{B}^T$ is the smallest σ-algebra of the subsets of W^T that makes all canonical projections π_{ST} measurable.*

Furthermore, the following lemma is true.

Lemma 2.5. *The canonical projections π_{ST} are measurable if and only if $\pi_{\{t\}T}$ for all $t \in T$ are measurable as well.*

Moreover, from a well-known result of measure theory, we have the following proposition.

Proposition 2.6. *A function $f : (\Omega, \mathcal{F}) \to (W^T, \mathcal{B}^T)$ is measurable if and only if for all $t \in T$ the composite mapping $\pi_{\{t\}T} \circ f : (\Omega, \mathcal{F}) \to (E_t, \mathcal{B}_t)$ is measurable.*

For proofs of Theorem 2.4, Lemma 2.5, and Proposition 2.6, see, e.g., Métivier (1968).

Remark 2.7. Let $(\Omega, \mathcal{F}, P, (X_t)_{t \in T})$ be a stochastic process with state space $(E, \mathcal{B})$. Since the function space $E^T = \prod_{t \in T} E$, the mapping $f : \Omega \to E^T$, which associates every $\omega \in \Omega$ with its corresponding trajectory of the process, is $(\mathcal{F} - \mathcal{B}^T)$-measurable, and in fact we have that

$$\forall t \in T: \qquad \pi_{\{t\}T} \circ f(\omega) = \pi_{\{t\}T}(X(\cdot, \omega)) = X_t(\omega),$$

where $\pi_{\{t\}T} \circ f = X_t$, which is a random variable, is obviously measurable.

Definition 2.8. A function $f : \Omega \to E^T$ defined on a probability space $(\Omega, \mathcal{F}, P)$ and valued in a measurable space $(E^T, \mathcal{G})$ is called a *random function* if it is $(\mathcal{F}\text{-}\mathcal{G})$-measurable.

How can we define a probability law P^T on $(E^T, \mathcal{B}^T)$ for the stochastic process $(X_t)_{t \in T}$ defined on the probability space $(\Omega, \mathcal{F}, P)$ in a coherent way? We may observe that from a *physical* point of view, it is natural to assume that in principle we are able, from experiments, to define all possible *finite-dimensional* joint probabilities

$$P(X_{t_1} \in B_1, \ldots, X_{t_n} \in B_n)$$

for any $n \in \mathbb{N}$, for any $\{t_1, \ldots, t_n\} \subset T$, and for any $B_1, \ldots, B_n \in \mathcal{B}$, i.e., the joint probability laws P^S of all finite-dimensional random vectors $(X_{t_1}, \ldots, X_{t_n})$, for any choice of $S = \{t_1, \ldots, t_n\} \subset \mathcal{S}$, such that

$$P^S(B_1 \times \cdots \times B_n) = P(X_{t_1} \in B_1, \ldots, X_{t_n} \in B_n).$$

Accordingly, we require that, for any $S \subset \mathcal{S}$,

$$P^T(\pi_{ST}^{-1}(B_1 \times \cdots \times B_n)) = P^S(B_1 \times \cdots \times B_n) = P(X_{t_1} \in B_1, \ldots, X_{t_n} \in B_n).$$

A general answer comes from the following theorem. Having constructed the σ-algebra $\mathcal{B}^T$ on E^T, we now define a measure μ_T on $(W^T, \mathcal{B}^T)$, supposing that, for all $S \in \mathcal{S}$, a measure μ_S is assigned on $(W^S, \mathcal{B}^S)$. If $S \in \mathcal{S}$, $S' \in \mathcal{S}$, and $S \subset S'$, we denote the canonical projection of $W^{S'}$ on W^S by $\pi_{SS'}$, which is certainly $(\mathcal{B}^{S'}$-$\mathcal{B}^S)$-measurable.

Definition 2.9. If, for all $(S, S') \in \mathcal{S} \times \mathcal{S}'$, with $S \subset S'$, we have that $\pi_{SS'}(\mu_{S'}) = \mu_S$, then

$$(W^S, \mathcal{B}^S, \mu_S, \pi_{SS'})_{S,S' \in \mathcal{S}; S \subset S'}$$

is called a *projective system* of measurable spaces and $(\mu_S)_{S \in \mathcal{S}}$ is called a *compatible system* of measures on the finite products $(W^S, \mathcal{B}^S)_{S \in \mathcal{S}}$.

Theorem 2.10 (Kolmogorov–Bochner). *Let $(E_t, \mathcal{B}_t)_{t \in T}$ be a family of Polish spaces (i.e., metric, complete, separable) endowed with their respective Borel σ-algebras, and let $\mathcal{S}$ be the collection of finite subsets of T and, for all $S \in \mathcal{S}$ with $W^S = \prod_{t \in S} E_t$ and $\mathcal{B}^S = \bigotimes_{t \in S} \mathcal{B}_t$, let μ_S be a finite measure on $(W^S, \mathcal{B}^S)$. Under these assumptions the following two statements are equivalent:*

1. *There exists a μ_T measure on $(W^T, \mathcal{B}^T)$ such that for all $S \in \mathcal{S}$: $\mu_S = \pi_{ST}(\mu_T)$.*
2. *The system $(W^S, \mathcal{B}^S, \mu_S, \pi_{SS'})_{S,S' \in \mathcal{S}; S \subset S'}$ is projective.*

Moreover, in both cases, μ_T, as defined in 1, is unique.

Proof. See, e.g., Métivier (1968). □

Definition 2.11. The unique measure μ_T of Theorem 2.10 is called the *projective limit* of the projective system $(W^S, \mathcal{B}^S, \mu_S, \pi_{SS'})_{S,S' \in \mathcal{S}; S \subset S'}$.

As a special case consider a family of probability spaces $(E_t, \mathcal{B}_t, P_t)_{t \in T}$. If, for all $S \in \mathcal{S}$, we define $P_S = \bigotimes_{t \in S} P_t$, then $(W^S, \mathcal{B}^S, P_S, \pi_{SS'})_{S,S' \in \mathcal{S}; S \subset S'}$ is a projective system and the projective probability limit $\bigotimes_{t \in T} P_t$ is called the *probability product* of the family of probabilities $(P_t)_{t \in T}$.

With respect to the projective system of finite-dimensional probability laws $P_S = \bigotimes_{t \in S} P_{X_t}$ of a stochastic process $(X_t)_{t \in \mathbb{R}_+}$, the projective limit will be the required *probability law of the process.*

Theorem 2.12. *Two stochastic processes $(X_t)_{t \in \mathbb{R}_+}$ and $(Y_t)_{t \in \mathbb{R}_+}$ that have the same finite-dimensional probability laws have the same probability law.*

Definition 2.13. Two stochastic processes are *equivalent* if and only if they have the same projective system of finite-dimensional joint distributions.

A more stringent notion follows.

Definition 2.14. Two real-valued stochastic processes $(X_t)_{t\in\mathbb{R}_+}$ and $(Y_t)_{t\in\mathbb{R}_+}$ on the probability space $(\Omega, \mathcal{F}, P)$ are called *modifications* or *versions* of one another if,

$$\text{for any } t \in T,\ P(X_t = Y_t) = 1.$$

Remark 2.15. It is obvious that two processes that are modifications of one another are also equivalent.

An even more stringent requirement comes from the following definition.

Definition 2.16. Two processes are *indistinguishable* if

$$P(X_t = Y_t, \forall t \in \mathbb{R}_+) = 1.$$

Remark 2.17. It is obvious that two indistinguishable processes are modifications of each other.

Example 2.18. Let $(X_t)_{t\in T}$ be a family of independent random variables defined on $(\Omega, \mathcal{F}, P)$ and valued in $(E, \mathcal{B})$. [In fact, in this case, it is sufficient to assume that only finite families of $(X_t)_{t\in T}$ are independent.] We know that for all $t \in T$ the probability $P_t = X_t(P)$ is defined on $(E, \mathcal{B})$. Then

$$\forall S = \{t_1, \ldots, t_r\} \in \mathcal{S}: \qquad P_S = \bigotimes_{k=1}^{r} P_{t_k} \text{for some } r \in \mathbb{N}^*,$$

and the system $(P_S)_{S\in\mathcal{S}}$ is compatible with its finite products $(E^S, \mathcal{B}^S)_{S\in\mathcal{S}}$. In fact, let $S, S' \in \mathcal{S}$, with $S = \{t_1, \ldots, t_r\} \subset S' = \{t_1, \ldots, t_{r'}\}$; if B is a rectangle of $\mathcal{B}^S$, i.e., $B = B_{t_1} \times \cdots \times B_{t_r}$, then

$$\begin{aligned} P_S(B) &= P_S(B_{t_1} \times \cdots \times B_{t_r}) = P_{t_1}(B_{t_1}) \cdot \cdots \cdot P_{t_r}(B_{t_r}) \\ &= P_{t_1}(B_{t_1}) \cdot \cdots \cdot P_{t_r}(B_{t_r}) P_{t_{r+1}}(E) \cdot \cdots \cdot P_{t_{r'}}(E) \\ &= P_{S'}(\pi_{SS'}^{-1}(B)). \end{aligned}$$

By the extension theorem we obtain that $P_S = \pi_{SS'}(P_{S'})$. As anticipated above, in this case we will write $P_T = \bigotimes_{t\in T} P_t$.

Remark 2.19. The compatibility condition $P_S = \pi_{SS'}(P_{S'})$, for all $S, S' \in \mathcal{S}$ and $S \subset S'$, can be expressed in an equivalent way by either the distribution function F_S of the probability P_S or its density f_S. For $E = \mathbb{R}$ we obtain, respectively,

1. For $S, S' \in \mathcal{S}$, with $S = \{t_1, \ldots, t_r\} \subset S' = \{t_1, \ldots, t_{r'}\}$; and for $(x_{t_1}, \ldots, x_{t_r}) \in \mathbb{R}^S$:
   $F_S(x_{t_1}, \ldots, x_{t_r}) = F_{S'}(x_{t_1}, \ldots, x_{t_r}, +\infty, \ldots, +\infty)$.

2. For $S, S' \in \mathcal{S}$, with $S = \{t_1, \ldots, t_r\} \subset S' = \{t_1, \ldots, t_{r'}\}$; and for $(x_{t_1}, \ldots, x_{t_r}) \in \mathbb{R}^S$:
$f_S(x_{t_1}, \ldots, x_{t_r}) = \int \cdots \int dx_{t_{r+1}} \cdots dx_{t_{r'}} f_{S'}(x_{t_1}, \ldots, x_{t_r}, x_{t_{r+1}}, \ldots, x_{t_{r'}})$.

Definition 2.20. A real-valued stochastic process $(X_t)_{t \in \mathbb{R}_+}$ is *continuous in probability* if

$$P - \lim_{s \to t} X_s = X_t, \qquad s, t \in \mathbb{R}_+.$$

Definition 2.21. A function $f : \mathbb{R}_+ \to \mathbb{R}$ is *right-continuous* if for any $t \in \mathbb{R}_+$, with $s > t$,

$$\lim_{s \downarrow t} f(s) = f(t).$$

Instead, the function is *left-continuous* if for any $t \in \mathbb{R}_+$, with $s < t$,

$$\lim_{s \uparrow t} f(s) = f(t).$$

Definition 2.22. A stochastic process $(X_t)_{t \in \mathbb{R}_+}$ is *right-(left-)continuous* if its trajectories are right-(left-)continuous almost surely. A stochastic process is continuous if its trajectories are continuous almost surely

Proposition 2.23. *A stochastic process that is continuous* a.s. *is continuous in probability. A stochastic process that is L^2-continuous is continuous in probability.*

Definition 2.24. A stochastic process $(X_t)_{t \in \mathbb{R}_+}$ is said to be *right-continuous with left limits* (RCLL) or *continu à droite avec limite à gauche* (càdlàg) if, almost surely, it has trajectories that are RCLL. The latter is denoted $X_{t^-} = \lim_{s \uparrow t} X_s$.

Theorem 2.25. *Let $(X_t)_{t \in \mathbb{R}_+}$ and $(Y_t)_{t \in \mathbb{R}_+}$ be two RCLL processes. X_t and Y_t are modifications of each other if and only if they are indistinguishable.*

As discussed in Doob (1953, p. 51) and in Billingsley (1986, p. 551), the finite-dimensional distributions, which determine the existence of the probability law of a stochastic process according to the Kolmogorov–Bochner theorem, are not sufficient to determine the properties of the sample paths of the process. On the other hand, it is possible, under rather general conditions, to ensure the property of separability of a process, and from this property various other desirable properties of the sample paths follow, such as continuity for the Brownian paths.

Definition 2.26. A real-valued stochastic process $(X_t)_{t \in \mathbb{R}_+}$ on the probability space $(\Omega, \mathcal{F}, P)$ is called *separable* if

- There exists a $T_0 \subset \mathbb{R}_+$, countable and dense everywhere in $\mathbb{R}_+$
- There exists an $A \in \mathcal{F}, P(A) = 0$ (negligible)

such that

- For all $t \in \mathbb{R}_+$: there exists $(t_n)_{n\in\mathbb{N}} \in T_0^{\mathbb{N}}$ such that $\lim_{n\to\infty} t_n = t$.
- For all $\omega \in \Omega \setminus A$: $\lim_{n\to\infty} X_{t_n}(\omega) = X_t(\omega)$.

The subset T_0 of $\mathbb{R}_+$, as defined previously, is called the *separating set.*

Theorem 2.27. *Let $(X_t)_{t\in\mathbb{R}_+}$ be a separable process, having T_0 and A as its separating and negligible sets, respectively. If $\omega \notin A$, $t_0 \in \mathbb{R}_+$, and $\lim_{t\to t_0} X_t(\omega)$ for $t \in T_0$ exists, then so does the limit $\lim_{t\to t_0} X_t(\omega)$ for $t \in \mathbb{R}_+$, and they coincide.*

Proof. See, e.g., Ash and Gardner (1975). □

Theorem 2.28. *Every real stochastic process $(X_t)_{t\in\mathbb{R}_+}$ admits a separable modification, almost surely finite, for any $t \in \mathbb{R}_+$.*

Proof. See, e.g., Ash and Gardner (1975). □

Remark 2.29. By virtue of Theorem 2.28, we may henceforth only consider separable processes.

In general, it is not true that a function $f(\omega_1, \omega_2)$ is jointly measurable in both variables, even if it is separately measurable in each of them. It is therefore required to impose conditions that guarantee the joint measurability of f in both variables. Evidently, if $(X_t)_{t\in\mathbb{R}_+}$ is a stochastic process, then for all $t \in \mathbb{R}_+ : X(t, \cdot)$ is measurable.

Definition 2.30. Let $(X_t)_{t\in\mathbb{R}_+}$ be a stochastic process defined on the probability space $(\Omega, \mathcal{F}, P)$ and valued in $(E, \mathcal{B}_E)$. The process $(X_t)_{t\in\mathbb{R}_+}$ is said to be *measurable* if it is measurable as a function defined on $\mathbb{R}_+ \times \Omega$ (with the σ-algebra $\mathcal{B}_{\mathbb{R}_+} \otimes \mathcal{F}$) and valued in E.

Proposition 2.31. *If the process $(X_t)_{t\in\mathbb{R}_+}$ is measurable, then the trajectory $X(\cdot, \omega) : \mathbb{R}_+ \to E$ is measurable for all $\omega \in \Omega$.*

Proof. Let $\omega \in \Omega$ and $B \in \mathcal{B}_E$. We want to show that $(X(\cdot, \omega))^{-1}(B)$ is an element of $\mathcal{B}_{\mathbb{R}_+}$. In fact,

$$(X(\cdot, \omega))^{-1}(B) = \{t \in \mathbb{R}_+ | X(t, \omega) \in B\} = \left\{t \in \mathbb{R}_+ | (t, \omega) \in X^{-1}(B)\right\},$$

meaning that $(X(\cdot, \omega))^{-1}(B)$ is the path ω of X^{-1}, which is certainly measurable, because $X^{-1}(B) \in \mathcal{B}_{\mathbb{R}_+} \otimes \mathcal{F}$ (as follows from the properties of the product σ-algebra). □

If the process is measurable, then it makes sense to consider the integral $\int_a^b X(t,\omega)dt$ along a trajectory. By Fubini's theorem, we have

$$\int_\Omega P(d\omega)\int_a^b dtX(t,\omega) = \int_a^b dt\int_\Omega P(d\omega)X(t,\omega).$$

Definition 2.32. The process $(X_t)_{t\in\mathbb{R}_+}$ is said to be *progressively measurable* with respect to the filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$, which is an increasing family of subalgebras of $\mathcal{F}$, if, for all $t \in \mathbb{R}_+$, the mapping $(s,\omega) \in [0,t]\times\Omega \to X(s,\omega) \in E$ is $(\mathcal{B}_{[0,t]}\otimes\mathcal{F}_t)$-measurable. Furthermore, we henceforth suppose that $\mathcal{F}_t = \sigma(X(s), 0 \leq s \leq t)$, $t \in \mathbb{R}_+$, which is called the *generated* or *natural* filtration of the process X_t.

Proposition 2.33. *If the process $(X_t)_{t\in\mathbb{R}_+}$ is progressively measurable, then it is also measurable.*

Proof. Let $B \in \mathcal{B}_E$. Then

$$\begin{aligned} X^{-1}(B) &= \{(s,\omega)\in\mathbb{R}_+\times\Omega | X(s,\omega)\in B\} \\ &= \bigcup_{n=0}^{\infty}\{(s,\omega)\in[0,n]\times\Omega | X(s,\omega)\in B\}. \end{aligned}$$

Since

$$\forall n: \qquad \{(s,\omega)\in[0,n]\times\Omega | X(s,\omega)\in B\}\in\mathcal{B}_{[0,n]}\otimes\mathcal{F}_n,$$

we obtain that $X^{-1}(B)\in\mathcal{B}_{\mathbb{R}_+}\otimes\mathcal{F}$. □

Theorem 2.34. *Let $(X_t)_{t\in\mathbb{R}_+}$ be a real stochastic process continuous in probability; then it admits a separable and progressively measurable modification.*

Proof. See, e.g., Ash and Gardner (1975). □

Definition 2.35. A filtered complete probability space $(\Omega,\mathcal{F},P,(\mathcal{F}_t)_{t\in\mathbb{R}_+})$ is said to satisfy the *usual hypotheses* if

1. $\mathcal{F}_0$ contains all the P-null sets of $\mathcal{F}$.
2. $\mathcal{F}_t = \bigcap_{s>t}\mathcal{F}_s$, for all $t \in \mathbb{R}_+$, i.e., the filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ is right-continuous.

Henceforth we will always assume that the usual hypotheses hold, unless specified otherwise.

Definition 2.36. Let $(\Omega,\mathcal{F},P,(\mathcal{F}_t)_{t\in\mathbb{R}_+})$ be a filtered probability space. The σ-algebra on $\mathbb{R}_+\times\Omega$ generated by all sets of the form $\{0\}\times A$, $A\in\mathcal{F}_0$, and $]a,b]\times A$, $0\leq a<b<+\infty$, $A\in\mathcal{F}_a$, is said to be the *predictable σ-algebra* for the filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$.

Definition 2.37. A real-valued process $(X_t)_{t\in\mathbb{R}_+}$ is called *predictable* with respect to a filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$, or *$\mathcal{F}_t$-predictable*, if as a mapping from $\mathbb{R}_+\times$

$\Omega \to \mathbb{R}$ it is measurable with respect to the predictable σ-algebra generated by this filtration.

Definition 2.38. A *simple predictable process* is of the form

$$X = k_0 I_{\{0\}\times A} + \sum_{i=1}^{n} k_i I_{]a_i,b_i]\times A_i},$$

where $A_0 \in \mathcal{F}_0$, $A_i \in \mathcal{F}_{a_i}$, $i = 1, \ldots, n$, and $k_0, \ldots, k_n$ are real constants.

Proposition 2.39. *Let $(X_t)_{t\in\mathbb{R}_+}$ be a process that is $\mathcal{F}_t$-predictable. Then, for any $t > 0$, X_t is $\mathcal{F}_{t^-}$-measurable.*

Lemma 2.40. *Let $(X_t)_{t\in\mathbb{R}_+}$ be a left-continuous real-valued process adapted to $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$. Then X_t is predictable.*

Lemma 2.41. *A process is predictable if and only if it is measurable with respect to the smallest σ-algebra on $\mathbb{R}_+ \times \Omega$ generated by the adapted left-continuous processes.*

Proposition 2.42. *Every predictable process is progressively measurable.*

Proposition 2.43. *If the process $(X_t)_{t\in\mathbb{R}_+}$ is right-(left-)continuous, then it is progressively measurable.*

Proof. See, e.g., Métivier (1968). □

Let $(X_t)_{t\in\mathbb{R}_+}$ be an $\mathbb{R}^d$-valued stochastic process defined on the probability space $(\Omega, \mathcal{F}, P)$.

We say it is *continuous* (resp. *right-continuous*, *left-continuous*) if, for almost all $\omega \in \Omega$, the trajectory $(X_t(\omega))_{t\in\mathbb{R}_+}$ is *continuous* (resp. *right-continuous*, *left-continuous*) with respect to t.

We say it is $\mathcal{F}$*-adapted* (or simply *adapted*) if, for every $t \in \mathbb{R}_+$, X_t is $\mathcal{F}$-measurable.

We say it is *measurable* if the function $(t, \omega) \in \mathbb{R}_+ \times \Omega \mapsto X_t(\omega) \in \mathbb{R}^d$ is $\mathcal{B}_{\mathbb{R}_+} \times \mathcal{F}$-measurable.

We say it is *progressively measurable* or *progressive* if, for every $T \in \mathbb{R}_+$, the function $(t, \omega) \in [0, T] \times \Omega \mapsto X_t(\omega) \in \mathbb{R}^d$ is $\mathcal{B}_{[0,T]} \times \mathcal{F}_T$-measurable.

Let $\mathcal{O}$ (resp. $\mathcal{P}$) be the smallest σ-algebra on $\mathbb{R}_+ \times \Omega$ with respect to which every càdlàg-adapted process (resp. left-continuous process) is a measurable function of (t, ω). We say that the stochastic process $(X_t)_{t\in\mathbb{R}_+}$ is *optional* (resp. *predictable*) if the process regarded as a function of (t, ω) is $\mathcal{O}$-measurable (resp. $\mathcal{P}$-measurable).

We say that a real-valued stochastic process $(X_t)_{t\in\mathbb{R}_+}$ is an *increasing* process if, for almost all $\omega \in \Omega$, $X_t(\omega)$ is nonnegative nondecreasing right-continuous with respect to $t \in \mathbb{R}_+$.

We say it is a *process of finite variation* if it can be decomposed as $X_t = \overline{X}_t - \widehat{X}_t$, with both $\overline{X}_t$ and $\widehat{X}_t$ increasing processes.

It is obvious that processes of finite variation are cadlag. Hence adapted processes of finite variation are optional.

The relations among various properties of stochastic properties are summarized below.

continuous adapted		continuous adapted		adapted increasing
$\Downarrow$		$\Downarrow$		$\Downarrow$
left-continuous adapted		cadlag adapted	$\Leftarrow$	adapted finite variation
$\Downarrow$		$\Downarrow$		
predictable	$\Rightarrow$	optional		
		$\Downarrow$		
		progressive	$\Rightarrow$	adapted
		$\Downarrow$		
		measurable		

2.2 Stopping Times

In what follows we are given a probability space $(\Omega, \mathcal{F}, P)$ and a filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ on $\mathcal{F}$.

Definition 2.44. A random variable T defined on Ω (endowed with the σ-algebra $\mathcal{F}$) and valued in $\bar{\mathbb{R}}_+$ is called a *stopping time* (or *Markov time*) with respect to the filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$, or simply an $\mathcal{F}_t$-*stopping time*, if

$$\forall t \in \mathbb{R}_+ : \qquad \{\omega | T(\omega) \leq t\} \in \mathcal{F}_t.$$

The stopping time is said to be finite if $P(T = \infty) = 0$.

Remark 2.45. If $T(\omega) \equiv k$ (constant), then T is always a stopping time. If T is a stopping time with respect to the filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ generated by the stochastic process $(X_t)_{t\in\mathbb{R}_+}$, $t \in \mathbb{R}_+$, then T is called the *stopping time of the process.*

Definition 2.46. Let T be an $\mathcal{F}_t$-stopping time. $A \in \mathcal{F}$ is said to *precede* T if, for all $t \in \mathbb{R}_+$: $A \cap \{T \leq t\} \in \mathcal{F}_t$.

Proposition 2.47. *Let T be an $\mathcal{F}_t$-stopping time, and let*

$$\mathcal{F}_T = \{A \in \mathcal{F} | A \text{ precedes } T\} ;$$

then $\mathcal{F}_T$ is a σ-algebra of the subsets of Ω. It is called the σ-algebra of T-preceding events.

Proof. See, e.g., Métivier (1968). □

Theorem 2.48. *The following relationships hold:*

1. *If both S and T are stopping times, then so are $S \wedge T = \inf\{S, T\}$ and $S \vee T = \sup\{S, T\}$.*
2. *If T is a stopping time and $a \in [0, +\infty[$, then $T \wedge a$ is a stopping time.*
3. *If T is a finite stopping time, then it is $\mathcal{F}_T$-measurable.*
4. *If both S and T are stopping times and $A \in \mathcal{F}_S$, then $A \cap \{S \leq T\} \in \mathcal{F}_T$.*
5. *If both S and T are stopping times and $S \leq T$, then $\mathcal{F}_S \subset \mathcal{F}_T$.*

Proof. See, e.g., Métivier (1968). □

Theorem 2.49. *Let $(X_t)_{t \in \mathbb{R}_+}$ be a progressively measurable stochastic process valued in $(S, \mathcal{B}_S)$. If T is a finite stopping time, then the function*

$$X(T) : \omega \in \Omega \to X(T(\omega), \omega) \in E$$

is $\mathcal{F}_T$-measurable (and hence a random variable).

Proof. We need to show that

$$\forall B \in \mathcal{B}_E : \qquad \{\omega | X(T(\omega)) \in B\} \in \mathcal{F}_T,$$

hence

$$\forall B \in \mathcal{B}_E, \forall t \in \mathbb{R}_+ : \qquad \{\omega | X(T(\omega)) \in B\} \cap \{T \leq t\} \in \mathcal{F}_t.$$

Fixing $B \in \mathcal{B}_E$ we have

$$\forall t \in \mathbb{R}_+ : \qquad \{\omega | X(T(\omega)) \in B\} \cap \{T \leq t\} = \{X(T \wedge t) \in B\} \cap \{T \leq t\},$$

where $\{T \leq t\} \in \mathcal{F}_t$, since T is a stopping time. We now show that

$$\{X(T \wedge t) \in B\} \in \mathcal{F}_t.$$

In fact, $T \wedge t$ is a stopping time (by point 2 of Theorem 2.48) and is $\mathcal{F}_{T \wedge t}$- measurable (by point 3 of Theorem 2.48). But $\mathcal{F}_{T \wedge t} \subset \mathcal{F}_t$ and thus $T \wedge t$ is $\mathcal{F}_t$-measurable. Now $X(T \wedge t)$ is obtained as a composite of the mapping

$$\omega \in \Omega \to (T \wedge t(\omega), \omega) \in [0, t] \times \Omega, \tag{2.1}$$

with

$$(s, \omega) \in [0, t] \times \Omega \to X(s, \omega) \in E. \tag{2.2}$$

The mapping (2.1) is $(\mathcal{F}_t - \mathcal{B}_{[0,t]} \otimes \mathcal{F}_t)$-measurable (because $T \wedge t$ is $\mathcal{F}_t$-measurable) and the mapping (2.2) is $(\mathcal{B}_{[0,t]} \otimes \mathcal{F}_t - \mathcal{B}_E)$-measurable since X is progressively measurable. Therefore, $X(T \wedge t)$ is $\mathcal{F}_t$-measurable, completing the proof. □

2.3 Canonical Form of a Process

Let $(\Omega, \mathcal{F}, P, (X_t)_{t\in T})$ be a stochastic process valued in $(E, \mathcal{B})$ and, for every $S \in \mathcal{S}$, let P_S be the joint probability law for the random variables $(X_t)_{t\in S}$ that is the probability on $(E^S, \mathcal{B}^S)$ induced by P through the function

$$X^S : \omega \in \Omega \to (X_t(\omega))_{t\in S} \in E^S = \prod_{t\in S} E.$$

Evidently, if

$$S \subset S'(S, S' \in \mathcal{S}), \qquad X^S = \pi_{SS'} \circ X^{S'},$$

then it follows that

$$P_S = X^S(P) = (\pi_{SS'} \circ X^{S'})(P) = \pi_{SS'}(P_{S'}),$$

and therefore $(E^S, \mathcal{B}^S, P_S, \pi_{SS'})_{S,S'\in\mathcal{S}, S\subset S'}$ is a projective system of probabilities.

On the other hand, the random function $f : \Omega \to E^T$ that associates every $\omega \in \Omega$ with a trajectory of the process in ω is measurable (following Proposition 2.6). Hence we can consider the induced probability P_T on $\mathcal{B}^T$, $P_T = f(P)$; P_T is the projective limit of $(P_S)_{S\in\mathcal{S}}$. From this it follows that $(E^T, \mathcal{B}^T, P_T, (\pi_t)_{t\in T})$ is a stochastic process with the property that, for all $S \in \mathcal{S}$, the random vectors $(\pi_t)_{t\in S}$ and $(X_t)_{t\in S}$ have the same joint distribution.

Definition 2.50. The stochastic process $(E^T, \mathcal{B}^T, P_T, (\pi_t)_{t\in T})$ is called the *canonical form* of the process $(\Omega, \mathcal{F}, P, (X_t)_{t\in T})$.

Remark 2.51. From this it follows that two stochastic processes are equivalent if they admit the same canonical process.

2.4 Gaussian Processes

Definition 2.52. A real-valued stochastic process $(\Omega, \mathcal{F}, P, (X_t)_{t\in\mathbb{R}_+})$ is called a *Gaussian process* if, for all $n \in \mathbb{N}^*$ and for all $(t_1, \ldots, t_n) \in \mathbb{R}_+^n$, the n-dimensional random vector $\mathbf{X} = (X_{t_1}, \ldots, X_{t_n})'$ has a multivariate Gaussian distribution, with probability density

$$f_{t_1,\ldots,t_n}(\mathbf{x}) = \frac{1}{(2\pi)^{n/2}\sqrt{\det K}} \exp\left\{-\frac{1}{2}(\mathbf{x}-\mathbf{m})'K^{-1}(\mathbf{x}-\mathbf{m})\right\}, \tag{2.3}$$

where

$$\begin{cases} m_i = E[X_{t_i}] \in \mathbb{R}, & i = 1, \ldots, n, \\ K_{ij} = Cov[X_{t_i}, X_{t_j}] \in \mathbb{R}, & i, j = 1, \ldots, n. \end{cases}$$

The covariance matrix $K = (\sigma_{ij})$ is taken as positive-definite, i.e., for all $\mathbf{a} \in \mathbb{R}^n\colon \sum_{i,j=1}^n a_i K_{ij} a_j > 0$).

The existence of Gaussian processes is guaranteed by the following remarks. By assigning a real-valued function

$$m : \mathbb{R}_+ \to \mathbb{R},$$

and a positive-definite function

$$K : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R},$$

thanks to well-known properties of multivariate Guassian distributions, we may introduce a projective system of Gaussian laws $(P_S)_{S\in\mathcal{S}}$ (where $\mathcal{S}$ is the set of all finite subsets of $\mathbb{R}_+$) of the form (2.3) such that, for $S = \{t_1, \ldots, t_n\}$,

$$m_i = m(t_i), \quad i = 1, \ldots, n,$$

$$K_{ij} = K(t_i, t_j), \quad i, j = 1, \ldots, n.$$

Since $\mathbb{R}$ is a Polish space, by the Kolmogorov–Bochner Theorem 2.10, we can now assert that there exists a Gaussian process $(X_t)_{t\in\mathbb{R}_+}$ having the preceding $(P_S)_{S\in\mathcal{S}}$ as its projective system of finite-dimensional distributions.

Example 2.53. The *standard Brownian Bridge* is a centered Gaussian process $(X_t)_{t\in[0,1]}$ on $\mathbb{R}$ such that

$$\begin{cases} \forall t \in [0,1]\colon\ E[X_t] = 0; \\ \forall (s,t) \in [0,1] \times [0,1], s \leq t\colon\ Cov[X_s, X_t] = s(1-t). \end{cases}$$

2.5 Processes with Independent Increments

Definition 2.54. The stochastic process $(\Omega, \mathcal{F}, P, (X_t)_{t\in\mathbb{R}_+})$, with state space $(E, \mathcal{B})$, is called a *process with independent increments* if, for all $n \in \mathbb{N}$ and for all $(t_1, \ldots, t_n) \in \mathbb{R}_+^n$, where $t_1 < \cdots < t_n$, the random variables $X_{t_1}, X_{t_2} - X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}}$ are independent.

Theorem 2.55. *If $(\Omega, \mathcal{F}, P, (X_t)_{t\in\mathbb{R}_t})$ is a process with independent increments, then it is possible to construct a compatible system of probability laws $(P_S)_{S\in\mathcal{S}}$, where again $\mathcal{S}$ is a collection of finite subsets of the index set.*

Proof. To do this, we need to assign a joint distribution to every random vector $(X_{t_1}, \ldots, X_{t_n})$ for all $(t_1, \ldots, t_n)$ in $\mathbb{R}_+^n$ with $t_1 < \cdots < t_n$. Thus, let $(t_1, \ldots, t_n) \in \mathbb{R}_+^n$, with $t_1 < \cdots < t_n$, and μ_0, $\mu_{s,t}$ be the distributions of X_0

and $X_t - X_s$, for every $(s,t) \in \mathbb{R}_+ \times \mathbb{R}_+$, with $s < t$, respectively. We define

$$\begin{aligned} Y_0 &= X_0, \\ Y_1 &= X_{t_1} - X_0, \\ &\cdots \\ Y_n &= X_{t_n} - X_{t_{n-1}}, \end{aligned}$$

where $Y_0, Y_1, \ldots, Y_n$ have the distributions $\mu_0, \mu_{0,t_1}, \ldots, \mu_{t_{n-1},t_n}$, respectively. Moreover, since the Y_i are independent, $(Y_0, \ldots, Y_n)$ have joint distribution $\mu_0 \otimes \mu_{0,t_1} \otimes \cdots \otimes \mu_{t_n,t_{n-1}}$. Let f be a real-valued, $\bigotimes^n \mathcal{B}$-measurable function, and consider the random variable $f(X_{t_1}, \ldots, X_{t_n})$. Then

$$\begin{aligned} &E[f(X_{t_1}, \ldots, X_{t_n})] \\ &= E[f(Y_0 + Y_1, \ldots, Y_0 + \cdots + Y_n)] \\ &= \int f(y_0 + y_1, \ldots, y_0 + \cdots + y_n) d(\mu_0 \otimes \mu_{0,t_1} \otimes \cdots \otimes \mu_{t_{n-1},t_n})(y_0, \ldots, y_n). \end{aligned}$$

In particular, if $f = I_B$, with $B \in \bigotimes^n \mathcal{B}$, we obtain the joint distribution of $X_{t_1}, \ldots, X_{t_n}$:

$$\begin{aligned} &P((X_{t_1}, \ldots, X_{t_n}) \in B) = E[I_B(X_{t_1}, \ldots, X_{t_n})] \\ &= \int I_B(y_0 + y_1, \ldots, y_0 + \cdots + y_n) d(\mu_0 \otimes \mu_{0,t_1} \otimes \cdots \otimes \mu_{t_{n-1},t_n})(y_0, \ldots, y_n). \end{aligned} \tag{2.4}$$

Having obtained P_S, where $S = \{t_1, \ldots, t_n\}$, with $t_1 < \cdots < t_n$, we show that $(P_S)_{S \in \mathcal{S}}$ is a compatible system. Let $S, S' \in \mathcal{S}; S \subset S'$, $S = \{t_1, \ldots, t_n\}$, with $t_1 < \cdots < t_n$ and $S' = \{t_1, \ldots, t_j, s, t_{j+1}, \ldots, t_n\}$, with $t_1 < \cdots < t_j < s < t_{j+1} < \cdots < t_n$. For $B \in \mathcal{B}^S$ and $B' = \pi_{SS'}^{-1}(B)$, we will show that $P_S(B) = P_{S'}(B')$.

We can observe, by the definition of B', that

$$I_{B'}(x_{t_1}, \ldots, x_{t_j}, x_s, x_{t_{j+1}}, \ldots, x_{t_n})$$

does not depend on x_s and is therefore identical to $I_B(x_{t_1}, \ldots, x_{t_n})$. Thus putting $U = X_s - X_{t_j}$ and $V = X_{t_{j+1}} - X_s$, we obtain

$$\begin{aligned} P_{S'}(B') = &\int I_{B'}(y_0 + y_1, \ldots, y_0 + \cdots + y_j, y_0 + \cdots + y_j + u, y_0 + \cdots \\ &+ y_j + u + v, \ldots, y_0 + \cdots + y_n) d(\mu_0 \otimes \mu_{0,t_1} \otimes \cdots \\ &\otimes \mu_{t_j,s} \otimes \mu_{s,t_{j+1}} \otimes \cdots \otimes \mu_{t_{n-1},t_n})(y_0, \ldots, y_j, u, v, y_{j+2}, \ldots, y_n) \\ = &\int I_B(y_0 + y_1, \ldots, y_0 + \cdots + y_j, y_0 + \cdots + y_j + u + v, y_0 + \cdots \\ &+ u + v + y_{j+2}, \ldots, y_0 + \cdots + y_n) d(\mu_0 \otimes \mu_{0,t_1} \otimes \cdots \\ &\otimes \mu_{t_j,s} \otimes \mu_{s,t_{j+1}} \otimes \cdots \otimes \mu_{t_{n-1},t_n})(y_0, \ldots, y_j, u, v, y_{j+2}, \ldots, y_n). \end{aligned}$$

Integrating with respect to all the variables except u and v, after applying Fubini's theorem, we obtain

$$P_{S'}(B') = \int h(u+v)d(\mu_{t_j,s} \otimes \mu_{s,t_{j+1}})(u,v).$$

Letting $y_{j+1} = u + v$ we have

$$P_{S'}(B') = \int h(y_{j+1})d(\mu_{t_j,s} * \mu_{s,t_{j+1}})(y_{j+1}).$$

Moreover, we observe that the definition of $y_{j+1} = u + v$ is compatible with the preceding notation $Y_{j+1} = X_{t_{j+1}} - X_{t_j}$. In fact, we have

$$u + v = x_s - x_{t_j} + x_{t_{j+1}} - x_s = x_{t_{j+1}} - x_{t_j}.$$

Furthermore, for the independence of $(X_{t_{j+1}} - X_s)$ and $(X_s - X_{t_j})$ the sum of random variables

$$X_{t_{j+1}} - X_s + X_s - X_{t_j} = X_{t_{j+1}} - X_{t_j}$$

must have the distribution $\mu_{t_j,s} * \mu_{s,t_{j+1}}$, where $*$ denotes the convolution product. Therefore, having denoted the distribution of $X_{t_{j+1}} - X_{t_j}$ by $\mu_{t_j,t_{j+1}}$, we obtain

$$\mu_{t_j,s} * \mu_{s,t_{j+1}} = \mu_{t_j,t_{j+1}}.$$

As a consequence we have

$$P_{S'}(B') = \int h(y_{j+1})d\mu_{t_j,t_{j+1}}(y_{j+1}).$$

This integral coincides with the one in (2.4), and thus

$$P_S(B') = P((X_{t_1}, \ldots, X_{t_n}) \in B) = P_S(B).$$

If now $S' = S \cup \{s_1, \ldots, s_k\}$, the proof is completed by induction. □

Definition 2.56. A process with independent increments is called *time-homogeneous* if

$$\mu_{s,t} = \mu_{s+h,t+h} \qquad \forall s, t, h \in \mathbb{R}_+, s < t.$$

If $(\Omega, \mathcal{F}, P, (X_t)_{t \in \mathbb{R}_+})$ is a homogeneous process with independent increments, then as a particular case we have

$$\mu_{s,t} = \mu_{0,t-s} \qquad \forall s, t \in \mathbb{R}_+, s < t.$$

Definition 2.57. A family of measures $(\mu_t)_{t \in \mathbb{R}_+}$ that satisfy the condition

$$\mu_{t_1+t_2} = \mu_{t_1} * \mu_{t_2}$$

is called a *convolution semigroup.*

Remark 2.58. A time-homogeneous process with independent increments is completely defined by assigning it a convolution semigroup.

2.6 Martingales

Extension of the concept of continuous-time martingales is mainly due to P.A. Meyer and his coworkers (Meyer (1966)).

Definition 2.59. Let $(X_t)_{t\in\mathbb{R}_+}$ be a real-valued family of random variables defined on the probability space $(\Omega, \mathcal{F}, P)$, and let $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ be a filtration. The stochastic process $(X_t)_{t\in\mathbb{R}_+}$ is said to be *adapted* to the family $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ if, for all $t \in \mathbb{R}_+$, X_t is $\mathcal{F}_t$-measurable.

Definition 2.60. The stochastic process $(X_t)_{t\in\mathbb{R}_+}$, adapted to the filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$, is a *martingale* with respect to this filtration, provided the following conditions hold:

1. X_t is P-integrable for all $t \in \mathbb{R}_+$.
2. For all $(s,t) \in \mathbb{R}_+ \times \mathbb{R}_+, s < t : E[X_t|\mathcal{F}_s] = X_s$ almost surely.

$(X_t)_{t\in\mathbb{R}_+}$ is said to be a *submartingale* (*supermartingale*) with respect to $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ if, in addition to condition 1 and instead of condition 2, we have:

2′. For all $(s,t) \in \mathbb{R}_+ \times \mathbb{R}_+, s < t : E[X_t|\mathcal{F}_s] \geq X_s$ $(E[X_t|\mathcal{F}_s] \leq X_s)$ almost surely.

Remark 2.61. When the filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ is not specified, it is understood to be the increasing σ-algebra generated by the random variables of the process $(\sigma(X_s, 0 \leq s \leq t))_{t\in\mathbb{R}_+}$. In this case we can write $E[X_t|X_r, 0 \leq r \leq s]$, instead of $E[X_t|\mathcal{F}_s]$.

Example 2.62. The evolution of a gambler's wealth in a game of chance, the latter specified by the sequence of real-valued random variables $(X_n)_{n\in\mathbb{N}}$, will serve as a descriptive example of the preceding definitions. Suppose that two players flip a coin and the loser pays the winner (who guessed head or tail correctly) the amount α after every round. If $(X_n)_{n\in\mathbb{N}}$ represents the cumulative fortune of player 1, then after n throws he holds

$$X_n = \sum_{i=0}^{n} \Delta_i.$$

The random variables Δ_i (just like every flip of the coin) are independent and take values α and $-\alpha$ with probabilities p and q, respectively. Therefore, we see that

$$\begin{aligned} E[X_{n+1}|X_0, \ldots, X_n] &= E[\Delta_{n+1} + X_n|X_0, \ldots, X_n] \\ &= X_n + E[\Delta_{n+1}|X_0, \ldots, X_n]. \end{aligned}$$

Since Δ_{n+1} is independent of every $\sum_{i=0}^{k} \Delta_i, k = 0, \ldots, n$, we obtain

$$E[X_{n+1}|X_0, \ldots, X_n] = X_n + E[\Delta_{n+1}] = X_n + \alpha(p - q).$$

- If the game is fair, then $p = q$ and $(X_n)_{n\in\mathbb{N}}$ is a martingale.
- If the game is in player 1's favor, then $p > q$ and $(X_n)_{n\in\mathbb{N}}$ is a submartingale.
- If the game is to the disadvantage of player 1, then $p < q$ and $(X_n)_{n\in\mathbb{N}}$ is a supermartingale.

Example 2.63. Let $(X_t)_{t\in\mathbb{R}_+}$ be (for all $t \in \mathbb{R}_+$) a P-integrable stochastic process on $(\Omega, \mathcal{F}, P)$ with independent increments. Then $(X_t - E[X_t])_{t\in\mathbb{R}_+}$ is a martingale. In fact[4]:

$$E[X_t|\mathcal{F}_s] = E[X_t - X_s|\mathcal{F}_s] + E[X_s|\mathcal{F}_s], \qquad s < t,$$

and recalling that X_s is $\mathcal{F}_s$-measurable and that $(X_t - X_s)$ is independent of $\mathcal{F}_s$, we obtain that

$$E[X_t|\mathcal{F}_s] = E[X_t - X_s] + X_s = X_s, \qquad s < t.$$

Proposition 2.64. *Let $(X_t)_{t\in\mathbb{R}_+}$ be a real-valued martingale. If the function $\phi : \mathbb{R} \to \mathbb{R}$ is both convex and measurable and such that*

$$\forall t \in \mathbb{R}_+ : \qquad E[|\phi(X_t)|] < +\infty,$$

then $(\phi(X_t))_{t\in\mathbb{R}_+}$ is a submartingale.

Proof. Let $(s,t) \in \mathbb{R}_+ \times \mathbb{R}_+, s < t$. Following Jensen's inequality and the properties of the martingale $(X_t)_{t\in\mathbb{R}_+}$, we have that

$$\phi(X_s) = \phi(E[X_t|\mathcal{F}_s]) \leq E[\phi(X_t)|\mathcal{F}_s].$$

Letting

$$\mathcal{V}_s = \sigma(\phi(X_r), 0 \leq r \leq s) \qquad \forall s \in \mathbb{R}_+$$

and with the measurability of ϕ, it is easy to verify that $\mathcal{V}_s \subset \mathcal{F}_s$ for all $s \in \mathbb{R}_+$, and therefore

$$\phi(X_s) = E[\phi(X_s)|\mathcal{V}_s] \leq E[E[\phi(X_t)|\mathcal{F}_s]|\mathcal{V}_s] = E[\phi(X_t)|\mathcal{V}_s].$$

□

Lemma 2.65. *Let X and Y be two positive real random variables defined on $(\Omega, \mathcal{F}, P)$. If $X \in L^p(P)$ $(p > 1)$ and if, for all $\alpha > 0$,*

$$\alpha P(Y \geq \alpha) \leq \int_{\{Y\geq\alpha\}} X dP, \tag{2.5}$$

[4]For simplicity, but without loss of generality, we will assume that $E[X_t] = 0$, for all t. In the case where $E[X_t] \neq 0$, we can always define a variable $Y_t = X_t - E[X_t]$, so that $E[Y_t] = 0$. In that case $(Y_t)_{t\in\mathbb{R}_+}$ will again be a process with independent increments, so that the analysis is analogous.

then $Y \in L^p(P)$ *and* $\|Y\|_p \leq q\|X\|_p$, *where* $\frac{1}{p} + \frac{1}{q} = 1$.

Proof. We have

$$\begin{aligned}
E[Y^p] &= \int_\Omega Y^p(\omega)dP(\omega) = \int_\Omega dP(\omega)p\int_0^{Y(\omega)} \lambda^{p-1}d\lambda \\
&= p\int_\Omega dP(\omega)\int_0^\infty \lambda^{p-1}I_{\{\lambda\leq Y(\omega)\}}(\lambda)d\lambda \\
&= p\int_0^\infty d\lambda\lambda^{p-1}\int_\Omega dP(\omega)I_{\{\lambda\leq Y(\omega)\}}(\lambda) \\
&= p\int_0^\infty d\lambda\lambda^{p-1}P(\lambda\leq Y) = p\int_0^\infty d\lambda\lambda^{p-2}\lambda P(Y\geq\lambda) \\
&\leq p\int_0^\infty d\lambda\lambda^{p-2}\int_{\{Y\geq\lambda\}} XdP \\
&= p\int_\Omega dP(\omega)X(\omega)\int_0^\infty d\lambda\lambda^{p-2}I_{\{Y(\omega)\geq\lambda\}}(\lambda) \\
&= p\int_\Omega dP(\omega)X(\omega)\int_0^{Y(\omega)} d\lambda\lambda^{p-2} = \frac{p}{p-1}\int_\Omega dP(\omega)X(\omega)Y^{p-1}(\omega) \\
&= \frac{p}{p-1}E[Y^{p-1}X],
\end{aligned}$$

where, throughout, λ denotes the Lebesgue measure, and when changing the order of integration we invoke Fubini's theorem. By Hölder's inequality, we obtain

$$E[Y^p] \leq \frac{p}{p-1}E[Y^{p-1}X] \leq \frac{p}{p-1}E[X^p]^{\frac{1}{p}}E[Y^p]^{\frac{p-1}{p}},$$

which, after substitution and rearrangement, gives

$$E[Y^p]^{\frac{1}{p}} \leq qE[X^p]^{\frac{1}{p}},$$

as long as $E[Y^p] < +\infty$ (in such a case we may, in fact, divide the left- and right-hand sides by $E[Y^p]^{\frac{p-1}{p}}$). But in any case we can consider the sequence of random variables $(Y \wedge n)_{n\in\mathbb{N}}$ ($Y \wedge n$ is the random variable defined letting, for all $\omega \in \Omega$, $Y \wedge n(\omega) = \inf\{Y(\omega), n\}$); since, for all n, $Y \wedge n$ satisfies condition (2.5), then we obtain

$$\|Y \wedge n\|_p \leq q\|X\|_p,$$

and in the limit

$$\|Y\|_p = \lim_{n\to\infty}\|Y \wedge n\|_p \leq q\|X\|_p.$$

□

Proposition 2.66. *Let* $(X_n)_{n\in\mathbb{N}^*}$ *be a sequence of real random variables defined on the probability space* $(\Omega, \mathcal{F}, P)$, *and let* X_n^+ *be the positive part of* X_n.

1. *If* $(X_n)_{n\in\mathbb{N}^*}$ *is a submartingale, then*

$$P\left(\max_{1\leq k\leq n} X_k > \lambda\right) \leq \frac{1}{\lambda}E[X_n^+], \qquad \lambda > 0, n \in \mathbb{N}^*.$$

2. *If* $(X_n)_{n\in\mathbb{N}^*}$ *is a martingale and if, for all* $n \in \mathbb{N}^*$, $X \in L^p(P)$, $p > 1$, *then*

$$E\left[\left(\max_{1\leq k\leq n} |X_k|\right)^p\right] \leq \left(\frac{p}{p-1}\right)^p E[|X_n|^p], \qquad n \in \mathbb{N}^*.$$

(Points 1 and 2 are called Doob's inequalities.*)*

Proof.

1. For all $k \in \mathbb{N}^*$ we put $A_k = \bigcap_{j=1}^{k-1}\{X_j \leq \lambda\} \cap \{X_k > \lambda\}$ $(\lambda > 0)$, where all A_k are pairwise disjoint and $A = \{\max_{1\leq k\leq n} X_k > \lambda\}$. Thus it is obvious that $A = \bigcup_{k=1}^n A_k$. Because in A_k, X_k is greater than λ, we have

$$\int_{A_k} X_k dP \geq \lambda \int_{A_k} dP.$$

Therefore,

$$\forall k \in \mathbb{N}^*, \qquad \lambda P(A_k) \leq \int_{A_k} X_k dP,$$

resulting in

$$\begin{aligned}\lambda P(A) = \lambda P\left(\bigcup_{k=1}^n A_k\right) &= \lambda \sum_{k=1}^n P(A_k)\\ &\leq \sum_{k=1}^n \int_{A_k} X_k dP = \sum_{k=1}^n \int_{\Omega} X_k I_{A_k} dP = \sum_{k=1}^n E[X_k I_{A_k}]. \quad (2.6)\end{aligned}$$

Now we have

$$\begin{aligned}E[X_n^+] &= \int_\Omega X_n^+ dP\\ &\geq \int_A X_n^+ dP = \sum_{k=1}^n \int_{A_k} X_n^+ dP = \sum_{k=1}^n \int_\Omega X_n^+ I_{A_k} dP\\ &= \sum_{k=1}^n E[X_n^+ I_{A_k}] = \sum_{k=1}^n E[E[X_n^+ I_{A_k}|X_1,\ldots,X_k]]\\ &= \sum_{k=1}^n E[I_{A_k} E[X_n^+|X_1,\ldots,X_k]] \geq \sum_{k=1}^n E[I_{A_k} E[X_n|X_1,\ldots,X_k]],\end{aligned}$$

where the last row follows from the fact that I_{A_k} is $\sigma(X_1, \ldots, X_k)$-measurable. Moreover, since $(X_n)_{n\in\mathbb{N}^*}$ is a submartingale, we have

$$E[X_n^+] \geq \sum_{k=1}^{n} E[I_{A_k} X_k]. \tag{2.7}$$

By (2.6) and (2.7), $E[X_n^+] \geq \lambda P(A)$, and this completes the proof of 1. We can also observe that

$$\begin{aligned}\sum_{k=1}^{n} E[I_{A_k} X_n^+] &= \sum_{k=1}^{n} E[E[X_n^+ I_{A_k} | X_1, \ldots, X_k]] \\ &\geq \sum_{k=1}^{n} E[I_{A_k} E[X_n | X_1, \ldots, X_k]] \geq \sum_{k=1}^{n} E[I_{A_k} X_k] \geq \lambda P(A),\end{aligned}$$

and therefore

$$\lambda P\left(\max_{1\leq k\leq n} X_k > \lambda\right) \leq \sum_{k=1}^{n} E[I_{A_k} X_n^+]. \tag{2.8}$$

2. Let $(X_n)_{n\in\mathbb{N}^*}$ be a martingale such that $X_n \in L^p(P)$ for all $n \in \mathbb{N}^*$. Since $\phi = |x|$ is a convex function, it follows from Proposition 2.64 that $(|X_n|)_{n\in\mathbb{N}^*}$ is a submartingale. Thus from (2.8) we have

$$\begin{aligned}\lambda P\left(\max_{1\leq k\leq n} |X_k| > \lambda\right) &\leq \sum_{k=1}^{n} E[I_{A_k} |X_n^+|] = \sum_{k=1}^{n} E[I_{A_k} |X_n|] \\ &= \sum_{k=1}^{n} \int_{A_k} |X_n| dP = \int_A |X_n| dP \quad (\lambda > 0, n \in \mathbb{N}^*).\end{aligned}$$

Putting $X = \max_{1\leq k\leq n} |X_k|$ and $Y = |X_n|$, we obtain

$$\lambda P(X > \lambda) \leq \int_A Y dP = \int_{\{X>\lambda\}} Y dP,$$

and from Lemma 2.65 it follows that $\|X\|_p \leq q\|Y\|_p$. Thus $E[X^p] \leq q^p E[Y^p]$, proving 2.

□

Remark 2.67. Because

$$\max_{1\leq k\leq n} |X_k|^p = \left(\max_{1\leq k\leq n} |X_k|\right)^p,$$

by point 2 of Proposition 2.66 it is also true that

$$E\left[\max_{1\leq k\leq n} |X_k|^p\right] \leq \left(\frac{p}{p-1}\right)^p E[|X_n|^p].$$

Corollary 2.68. *If* $(X_n)_{n\in\mathbb{N}^*}$ *is a martingale such that* $X_n \in L^p(P)$ *for all* $n \in \mathbb{N}^*$*, then*

$$P\left(\max_{1\leq k\leq n}|X_k| > \lambda\right) \leq \frac{1}{\lambda^p}E[|X_n|^p], \qquad \lambda > 0.$$

Proof. From Proposition 2.64 we can assert that $(|X_n|^p)_{n\in\mathbb{N}^*}$ is a submartingale. In fact, $\phi(x) = |x|^p, p > 1$, is convex. By point 1 of Proposition 2.66, it follows that

$$P\left(\max_{1\leq k\leq n}|X_k|^p > \lambda^p\right) \leq \frac{1}{\lambda^p}E[|X_n|^p],$$

which is equivalent to

$$P\left(\max_{1\leq k\leq n}|X_k| > \lambda\right) \leq \frac{1}{\lambda^p}E[|X_n|^p].$$

□

Lemma 2.69. *The following statements are true:*

1. *If* $(X_t)_{t\in\mathbb{R}_+}$ *is a martingale, then so is* $(X_t)_{t\in I}$ *for all* $I \subset \mathbb{R}_+$.
2. *If, for all* $I \subset \mathbb{R}_+$ *and* I *finite,* $(X_t)_{t\in I}$ *is a (discrete) martingale, then so is* $(X_t)_{t\in\mathbb{R}_+}$.

Proof.

1. Let $I \subset \mathbb{R}_+, (s,t) \in I^2, s < t$. Because $(X_r)_{r\in\mathbb{R}_+}$ is a martingale,

$$X_s = E[X_t|X_r, 0 \leq r \leq s, r \in \mathbb{R}_+].$$

Observing that

$$\sigma(X_r, 0 \leq r \leq s, r \in I) \subset \sigma(X_r, 0 \leq r \leq s, r \in \mathbb{R}_+)$$

and remembering that in general

$$E[X|B_1] = E[E[X|B_2]|B_1], \qquad B_1 \subset B_2 \subset \mathcal{F},$$

we obtain

$$\begin{aligned} &E[X_t|X_r, 0 \leq r \leq s, r \in I] \\ &\quad = E[E[X_t|X_r, 0 \leq r \leq s, r \in \mathbb{R}_+]|X_r, 0 \leq r \leq s, r \in I] \\ &\quad = E[X_s|X_r, 0 \leq r \leq s, r \in I] \\ &\quad = X_s. \end{aligned}$$

The last equality holds because X_s is measurable with respect to $\sigma(X_r, 0 \leq r \leq s, r \in I)$.

2. See, e.g., Doob (1953).

□

Proposition 2.70. *Let* $(X_t)_{t\in\mathbb{R}_+}$ *be a stochastic process on* $(\Omega, \mathcal{F}, P)$ *valued in* $\mathbb{R}$.

1. *If* $(X_t)_{t\in\mathbb{R}_+}$ *is a submartingale, then*

$$P\left(\sup_{0\leq s\leq t} X_s > \lambda\right) \leq \frac{1}{\lambda} E[X_t^+], \qquad \lambda > 0, t \geq 0.$$

2. *If* $(X_t)_{t\in\mathbb{R}_+}$ *is a martingale such that, for all* $t \geq 0$, $X_t \in L^p(P)$, $p > 1$, *then*

$$E\left[\sup_{0\leq s\leq t} |X_s|^p\right] \leq \left(\frac{p}{p-1}\right)^p E[|X_t|^p].$$

Proof. See, e.g., Doob (1953). □

Definition 2.71. A subset H of $L^1(\Omega, \mathcal{F}, P)$ is *uniformly integrable* if

$$\lim_{c\to\infty} \sup_{Y\in H} \int_{\{|Y|>c\}} |Y| dP = 0.$$

Theorem 2.72. *A martingale is uniformly integrable if and only if it is of the form* $M_n = E[Y|\mathcal{F}_n]$, *where* $Y \in L^1(\Omega, \mathcal{F}, P)$. *Under these conditions* $\{M_n\}_n$ *converges almost surely and in* L^1.

Proof. See, e.g., Baldi (1984). □

The subsequent proposition specifies the limit of a uniformly integrable martingale.

Proposition 2.73. *Let* $Y \in L^1(\Omega, \mathcal{F}, P)$, $\{\mathcal{F}_n\}_n$ *be a filtration and* $\mathcal{F}_\infty = \bigcup_n \mathcal{F}_n$ *the* σ*-algebra generated by* $\{\mathcal{F}_n\}_n$. *Then*

$$\lim_{n\to\infty} E[Y|\mathcal{F}_n] = E[Y|\mathcal{F}_\infty] \text{ almost surely and in } L^1.$$

Proof. See, e.g., Baldi (1984). □

Doob–Meyer Decomposition

In the sequel, whenever not explicitly specified we will refer to the natural filtration of a process, suitably completed.

Proposition 2.74. *Every martingale has a right-continuous version.*

Theorem 2.75. *Let* X_t *be a supermartingale. Then the mapping* $t \to E[X_t]$ *is right-continuous if and only if there exists an RCLL modification of* X_t. *This modification is unique.*

Proof. See, e.g., Protter (1990). □

Definition 2.76. Consider the set $\mathcal{S}$ of stopping times T, with $P(T < \infty) = 1$, of the filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$. The right-continuous adapted process $(X_t)_{t\in\mathbb{R}_+}$ is said to be of *class D* if the family $(X_T)_{T\in\mathcal{S}}$ is uniformly integrable. Instead, if $\mathcal{S}_a$ is the set of stopping times with $P(T \leq a) = 1$, for a finite $a > 0$, and the family $(X_T)_{T\in\mathcal{S}_a}$ is uniformly integrable, then it is said to be of *class DL*.

Proposition 2.77. *Let $(X_t)_{t\in\mathbb{R}_+}$ be a right-continuous submartingale. Then X_t is of class DL under either of the following two conditions:*

1. *$X_t \geq 0$ almost surely for every $t \geq 0$.*
2. *X_t has the form*

$$X_t = M_t + A_t, \qquad t \in \mathbb{R}_+,$$

where $(M_t)_{t\in\mathbb{R}_+}$ is a martingale and $(A_t)_{t\in\mathbb{R}_+}$ an adapted increasing process.

Lemma 2.78. *If $(X_t)_{t\in\mathbb{R}_+}$ is a uniformly integrable martingale, then it is of class D.*

If $(X_t)_{t\in\mathbb{R}_+}$ is a martingale, or it is bounded from below, then it is of class DL.

Definition 2.79. Let $(X_t)_{t\in\mathbb{R}_+}$ be an adapted stochastic process with RCLL trajectories. It is said to be *decomposable* if it can be written as

$$X_t = X_0 + M_t + Z_t,$$

where $M_0 = Z_0 = 0$, M_t is a locally square-integrable martingale and Z_t has RCLL-adapted trajectories of bounded variation.

Theorem 2.80 (Doob–Meyer). *Let $(X_t)_{t\in\mathbb{R}_+}$ be an adapted right-continuous process. It is a submartingale of class D, with $X_0 = 0$ almost surely if and only if it can be decomposed as*

$$\forall t \in \mathbb{R}_+, \qquad X_t = M_t + A_t \, a.s.,$$

where M_t is a uniformly integrable martingale with $M_0 = 0$ and $A_t \in L^1(P)$ is an increasing predictable process with $A_0 = 0$. The decomposition is unique and if, in addition, X_t is bounded, then M_t is uniformly integrable and A_t is integrable.

Proof. See, e.g., Ethier and Kurtz (1986). □

Corollary 2.81. *Let $X = (X_t)_{t\in\mathbb{R}_+}$ be an adapted right-continuous submartingale of class DL; then there exists a unique (up to indistinguishability)* *right-continuous increasing predictable process A adapted to the same filtration as X, with* $A_0 = 0$ *almost surely, such that*

$$M_t = X_t - A_t, \quad t \in \mathbb{R}_+$$

is a martingale, adapted to the same filtration as X.

Definition 2.82. Resorting to the notation of Theorem 2.80, the process $(A_t)_{t\in\mathbb{R}_+}$ is called the *compensator* of X_t.

Proposition 2.83. *Under the assumptions of Theorem 2.80, the compensator* A_t *of* X_t *is continuous if and only if* X_t *is regular in the sense that for every predictable finite stopping time T we have that* $E[X_T] = E[X_{T-}]$.

Definition 2.84. A stochastic process $(M_t)_{t\in\mathbb{R}_+}$ is a *local martingale* with respect to the filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ if there exists a "localizing" sequence $(T_n)_{n\in\mathbb{N}}$ such that for each $n \in \mathbb{N}$, $(M_{t\wedge T_n})_{t\in\mathbb{R}_+}$ is an $\mathcal{F}_t$-martingale.

Definition 2.85. Let $(X_t)_{t\in\mathbb{R}_+}$ be a stochastic process. *Property* $\mathcal{P}$ is said to hold locally if

1. There exists $(T_n)_{n\in\mathbb{N}}$, a sequence of stopping times, with $T_n < T_{n+1}$.
2. $\lim_n T_n = +\infty$ almost surely.

such that $X_{T_n} I_{\{T_n>0\}}$ has property $\mathcal{P}$ for $n \in \mathbb{N}^*$.

Theorem 2.86. *Let* $(M_t)_{t\in\mathbb{R}_+}$ *be an adapted and RCLL stochastic process, and let* $(T_n)_{n\in\mathbb{N}}$ *be as in the preceding definition. If* $M_{T_n} I_{\{T_n>0\}}$ *is a martingale for each* $n \in \mathbb{N}^*$, *then* M_t *is a local martingale.*

Lemma 2.87. *Any martingale is a local martingale.*

Proof. Simply take $T_n = n$ for all $n \in \mathbb{N}^*$. □

Theorem 2.88 (Local form Doob–Meyer). *Let* $(X_t)_{t\in\mathbb{R}_+}$ *be a nonnegative right-continuous* $\mathcal{F}_t$*-local submartingale with* $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ *a right-continuous filtration. Then there exists a unique increasing right-continuous predictable process* $(A_t)_{t\in\mathbb{R}_+}$ *such that* $A_0 = 0$ *almost surely and* $P(A_t < \infty) = 1$ *for all* $t > 0$, *so that* $X_t - A_t$ *is a right-continuous local martingale.*

Definition 2.89. A martingale $M = (M_t)_{t\in\mathbb{R}_+}$ is square-integrable if, for all $t \in \mathbb{R}_+$, $E[|M_t|^2] < +\infty$.

We will denote by $\mathcal{M}$ the family of all right-continuous square-integrable martingales.

Remark 2.90. If $M \in \mathcal{M}$, then M^2 satisfies the conditions of Corollary 2.81; let $< M >$ be the increasing process given by the theorem with $X = M^2$. Then $\langle M_0\rangle = 0$, and $M_t^2 - \langle M_t\rangle$ is a martingale.

Definition 2.91. For two martingales M and N, in $\mathcal{M}$ the process

$$\langle M, N\rangle = \frac{1}{4}(\langle M+N\rangle - \langle M-N\rangle)$$

is called the *predictable covariation* of M and N. Evidently $\langle M, M\rangle = \langle M\rangle$, and so it is called the *predictable variation* of M.

Remark 2.92. Hence $\langle M, N\rangle$ is the unique finite variation predictable RCLL process such that $\langle M, N\rangle_0 = 0$ and $MN - \langle M, N\rangle$ is a martingale. Furthermore, if $\langle M, N\rangle = 0$, then the two martingales are said to be *orthogonal*. Thus M and N are orthogonal if and only if MN is a martingale.

Definition 2.93. A martingale M is said to be a *purely discontinuous martingale* if and only if $M_0 = 0$ and it is orthogonal to any continuous martingale.

Definition 2.94. Two local martingales M and N are said to be *orthogonal* if and only if MN is a local martingale.

Definition 2.95. A local martingale M is said to be a *purely discontinuous local martingale* if and only if $M_0 = 0$ and it is orthogonal to any continuous local martingale.

Having denoted by $\mathcal{M}$ the family of all right-continuous square-integrable martingales, let $\mathcal{M}_c \subset \mathcal{M}$ denote the family of all continuous square integrable martingales, and let $\mathcal{M}_d \subset \mathcal{M}$ denote the family of all purely discontinuous square-integrable martingales.

Theorem 2.96. *Any local martingale M admits a unique (up to indistinguishability) decomposition*

$$M = M_0 + M^c + M^d,$$

where M_c is a continuous local martingale and M_d is a purely discontinuous local martingale, with $M_0^c = M_0^d = 0$.

Proof. See, e.g., Jacod and Shiryaev (1987, p. 43). □

Remark 2.97. The reader has to be cautious about the meaning of the term "purely discontinuous"; it is indeed referring just to an orthogonality property with respect to the continuous case, but is does refer to the kind of discontinuities of its trajectories (e.g., Jacod and Shiryaev 1987, p. 40).

Proposition 2.98. *Let $X = (X_t)_{t\in\mathbb{R}_+}$ be a right-continuous martingale. Then there exists a right-continuous increasing process, denoted by $[X]$, such that for each $t \in \mathbb{R}_+$, and each sequence of partitions $(t_k^{(n)})_{n\in\mathbb{N}, 0\le k\le n}$ of $[0, t]$,* *with* $\max_k(t_{k+1}^{(n)} - t_k^{(n)}) \xrightarrow{n} \infty$*:*

$$\sum_k (X(t_{k+1}^{(n)}) - X(t_k^{(n)}))^2 \xrightarrow[n\to\infty]{P} [X](t). \tag{2.9}$$

If $X \in \mathcal{M}$*, then the convergence in (2.9) is in* L^1*. If* $X \in \mathcal{M}_c$*, then* $[X]$ *can be taken to be continuous.*

Proof. See, e.g., Ethier and Kurtz (1986). □

Definition 2.99. The process $[X]$ introduced above is known as the *quadratic variation* process associated with X.

Proposition 2.100. *If* $M \in \mathcal{M}$*, then* $M^2 - [M]$ *is a martingale.*

Remark 2.101. If $M \in \mathcal{M}_c$, then, by Proposition 2.9, $[M]$ is continuous, and Proposition 2.100 implies, by uniqueness, that $[M] =< M >$, up to indistinguishability.

Proposition 2.102. *Let* $M \in \mathcal{M}_c$*. Then* $< M >= 0$ *if and only if* M *is constant, i.e.,* $M_t = M_0$, a.s., *for any* $t \in \mathbb{R}_+$.

Proof. See, e.g., Revuz-Yor (1991, p. 119). □

2.7 Markov Processes

Definition 2.103. Let $(X_t)_{t\in\mathbb{R}_+}$ be a stochastic process on a probability space, valued in $(E, \mathcal{B})$ and adapted to the increasing family $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ of σ-algebras of subsets of $\mathcal{F}$. $(X_t)_{t\in\mathbb{R}_+}$ is a *Markov process* with respect to $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ if the following condition is satisfied:

$$\forall B \in \mathcal{B}, \forall (s,t) \in \mathbb{R}_+ \times \mathbb{R}_+, s < t: \qquad P(X_t \in B|\mathcal{F}_s) = P(X_t \in B|X_s) \text{ a.s.} \tag{2.10}$$

Remark 2.104. If, for all $t \in \mathbb{R}_+$, $\mathcal{F}_t = \sigma(X_r, 0 \le r \le t)$, then condition (2.10) becomes

$$P(X_t \in B|X_r, 0 \le r \le s) = P(X_t \in B|X_s) \text{ a.s.}$$

for all $B \in \mathcal{B}$, for all $(s,t) \in \mathbb{R}_+ \times \mathbb{R}_+$, and $s < t$.

Proposition 2.105. *Under the assumptions of Definition 2.103, the following two statements are equivalent:*

1. *For all* $B \in \mathcal{B}$ *and all* $(s,t) \in \mathbb{R}_+ \times \mathbb{R}_+, s < t$: $P(X_t \in B|\mathcal{F}_s) = P(X_t \in B|X_s)$ *almost surely.*
2. *For all* $g : E \to \mathbb{R}$, $\mathcal{B}$-$\mathcal{B}_\mathbb{R}$*-measurable such that* $g(X_t) \in L^1(P)$ *for all* t*, for all* $(s,t) \in \mathbb{R}_+^2, s < t$: $E[g(X_t)|\mathcal{F}_s] = E[g(X_t)|X_s]$ *almost surely.*

Proof. The proof is left to the reader as an exercise. □

Lemma 2.106. *If $(Y_k)_{k\in\mathbb{N}^*}$ is a sequence of real, independent random variables, then, putting*

$$X_n = \sum_{k=1}^{n} Y_k \qquad \forall n \in \mathbb{N}^*,$$

the new sequence $(X_n)_{n\in\mathbb{N}^*}$ is Markovian with respect to the family of σ-algebras $(\sigma(Y_1,\ldots,Y_n))_{n\in\mathbb{N}^*}$.

Proof. From the definition of X_k it is obvious that

$$\sigma(Y_1,\ldots,Y_n) = \sigma(X_1,\ldots,X_n) \qquad \forall n \in \mathbb{N}^*.$$

We thus first prove that, for all $C, D \in \mathcal{B}_{\mathbb{R}}$, for all $n \in \mathbb{N}^*$:

$$\begin{aligned} &P(X_{n-1} \in C, Y_n \in D|Y_1,\ldots,Y_{n-1}) \\ &= P(X_{n-1} \in C, Y_n \in D|X_{n-1}) \qquad \text{a.s.} \end{aligned} \tag{2.11}$$

To do this we fix $C, D \in \mathcal{B}_{\mathbb{R}}$ and $n \in \mathbb{N}^*$ and separately look at the left- and right-hand sides of (2.11). We get

$$\begin{aligned} &P(X_{n-1} \in C, Y_n \in D|Y_1,\ldots,Y_{n-1}) = E[I_C(X_{n-1})I_D(Y_n)|Y_1,\ldots,Y_{n-1}] \\ &= I_C(X_{n-1})E[I_D(Y_n)|Y_1,\ldots,Y_{n-1}] = I_C(X_{n-1})E[I_D(Y_n)] \text{ a.s.,} \end{aligned} \tag{2.12}$$

where the second equality of (2.12) holds because $I_C(X_{n-1})$ is $\sigma(Y_1,\ldots,Y_{n-1})$-measurable, and for the last one we use the fact that $I_D(Y_n)$ is independent of $Y_1,\ldots,Y_{n-1}$. On the other hand, we obtain that

$$\begin{aligned} P(X_{n-1} \in C, Y_n \in D|X_{n-1}) &= E[I_C(X_{n-1})I_D(Y_n)|X_{n-1}] \\ &= I_C(X_{n-1})E[I_D(Y_n)] \text{ a.s.} \end{aligned} \tag{2.13}$$

In fact, $I_C(X_{n-1})$ is $\sigma(X_{n-1})$-measurable and $I_D(Y_n)$ is independent of $X_{n-1} = \sum_{k=1}^{n-1} Y_k$. For (2.12) and (2.13), (2.11) follows and hence

$$\begin{aligned} &P((X_{n-1}, Y_n) \in C \times D|Y_1,\ldots,Y_{n-1}) \\ &= P((X_{n-1}, Y_n) \in C \times D|X_{n-1}) \qquad \text{a.s.} \end{aligned} \tag{2.14}$$

As (2.14) holds for the rectangles of $\mathcal{B}_{\mathbb{R}^2}$ $(= \mathcal{B}_{\mathbb{R}} \otimes \mathcal{B}_{\mathbb{R}})$, by the measure extension theorem (e.g., Bauer 1981), it follows that (2.14) is also true for every $B \in \mathcal{B}_{\mathbb{R}^2}$. If now $A \in \mathcal{B}_{\mathbb{R}}$, then the two events

$$\{X_{n-1} + Y_n \in A\} = \{(X_{n-1}, Y_n) \in B\},$$

where $B \in \mathcal{B}_{\mathbb{R}^2}$ is the inverse image of A for a generic mapping $+ : \mathbb{R}^2 \to \mathbb{R}$ (which is continuous and hence measurable), are identical. Applying (2.14) to B, we obtain

$$P(X_{n-1} + Y_n \in A|Y_1,\ldots,Y_{n-1}) = P(X_{n-1} + Y_n \in A|X_{n-1}) \text{ a.s.,}$$

and thus

$$P(X_{n-1}+Y_n \in A|X_1,\ldots,X_{n-1}) = P(X_{n-1}+Y_n \in A|X_{n-1}) \text{ a.s.},$$

and then

$$P(X_n \in A|X_1,\ldots,X_{n-1}) = P(X_n \in A|X_{n-1}) \text{ a.s.}$$

Therefore, $(X_n)_{n\in\mathbb{N}^*}$ is Markovian with respect to $(\sigma(X_1,\ldots,X_n))_{n\in\mathbb{N}^*}$ or, equivalently, with respect to $(\sigma(Y_1,\ldots,Y_n))_{n\in\mathbb{N}^*}$. □

Proposition 2.107. *Let $(X_t)_{t\in\mathbb{R}_+}$ be a real stochastic process defined on the probability space $(\Omega,\mathcal{F},P)$. The following two statements are true:*

1. *If $(X_t)_{t\in\mathbb{R}_+}$ is a Markov process, then so is $(X_t)_{t\in J}$ for all $J\subset\mathbb{R}_+$.*
2. *If for all $J\subset\mathbb{R}_+$, J finite: $(X_t)_{t\in J}$ is a Markov process, then so is $(X_t)_{t\in\mathbb{R}_+}$.*

Proof. See, e.g., Ash and Gardner (1975). □

Theorem 2.108. *Every real stochastic process $(X_t)_{t\in\mathbb{R}_+}$ with independent increments is a Markov process.*

Proof. We define $(t_1,\ldots,t_n)\in\mathbb{R}_+^n$ such that $0<t_1<\cdots<t_n$ and $t_0=0$. If, for simplicity, we further suppose that $X_0=0$, then $X_{t_n}=\sum_{k=1}^n(X_{t_k}-X_{t_{k-1}})$. Putting $Y_k=X_{t_k}-X_{t_{k-1}}$, then, for all $k=1,\ldots,n$, the Y_k are independent (because the process $(X_t)_{t\in\mathbb{R}_+}$ has independent increments) and we have that

$$X_{t_n}=\sum_{k=1}^n Y_k.$$

From Lemma 2.106 we can assert that

$$\forall B\in\mathcal{B}_{\mathbb{R}}: \qquad P(X_{t_n}\in B|X_{t_1},\ldots,X_{t_{n-1}}) = P(X_{t_n}\in B|X_{t_{n-1}}) \text{ a.s.}$$

Thus $\forall J\subset\mathbb{R}_+, J$ finite, $(X_t)_{t\in J}$ is Markovian. The theorem then follows by point 2 of Proposition 2.107. □

Proposition 2.109. *Let $(E,\mathcal{B}_E)$ be a Polish space endowed with the σ-algebra $\mathcal{B}_E$ of its Borel sets. For $t_0,T\in\mathbb{R}$, with $t_0<T$, let $(X_t)_{t\in[t_0,T]}$ be an E-valued Markov process, with respect to its natural filtration.*

The function

$$\begin{aligned}(s,t)\in[t_0,T]\times[t_0,T],\ s\le t;\ x\in E;\ A\in\mathcal{B}_E\mapsto \\ p(s,x;t,A):=P(X_t\in A|X_s=x)\in[0,1]\end{aligned} \tag{2.15}$$

satisfies the following properties:

(i) For all $(s,t)\in[t_0,T]\times[t_0,T]$, $s\le t$, and for all $A\in\mathcal{B}_E$, the function $x\in E\mapsto p(s,x,t,A)$ is $\mathcal{B}_E-\mathcal{B}_{\mathbb{R}}$-measurable.

(ii) *For all* $(s,t) \in [t_0,T] \times [t_0,T]$, $s \leq t$, *and for all* $x \in E$, *the function* $A \in \mathcal{B}_E \mapsto p(s,x,t,A)$ *is a probability measure on* $\mathcal{B}_E$ *such that*

$$p(s,x,s,A) = \begin{cases} 1, \text{ if } x \in A, \\ 0, \text{ if } x \notin A. \end{cases}$$

(iii) *The function* p *defined in (2.15) satisfies the so-called Chapman –Kolmogorov equation, i.e., for all* $x \in E$, *for all* $(s,r,t) \in [t_0,T] \times [t_0,T] \times [t_0,T]$, $s \leq r \leq t$, *and for all* $A \in \mathcal{B}_E$

$$p(s,x,t,A) = \int_{\mathbb{R}} p(s,x,r,dy)p(r,y,t,A) \text{ a.s.}$$

Proof. The proofs of properties (i) and (ii) are trivial consequences of the definitions (e.g., Ash and Gardner 1975). On the other hand the proof of (iii) is left to later analysis, after the introduction of the semigroup associated with the Markov process. □

Definition 2.110. Any nonnegative function $p(s,x,t,A)$ defined for $t_0 \leq s \leq t \leq T, x \in E, A \in \mathcal{B}_E$ that satisfies conditions (i), (ii), and (iii) is called a *Markov transition probability function.*

Definition 2.111. If $(X_t)_{t\in[t_0,T]}$ is a Markov process, then the distribution P_0 of $X(t_0)$ is the *initial distribution* of the process.

Theorem 2.112. *An* $(E,\mathcal{B}_E)$*-valued process* $(X_t)_{t\in[t_0,T]}$ *is a Markov process, with transition probability function* $p(r,x,s,A), t_0 \leq r < s \leq T, x \in E, A \in \mathcal{B}_E$ *and initial distribution* P_0, *if and only if, for any* $t_0 < t_1 < \cdots < t_k$, $k \in \mathbb{N}^*$, *and for any family* f_i, $i = 0,1,\ldots,k$ *of nonnegative Borel measurable real-valued functions*

$$E\left[\prod_{i=0}^{k} f_i(X_{t_i})\right] = \int_E P_0(dx_0)f_0(x_0)\int_E p(0,x_0,t_1,dx_1)f_1(x_1)\cdots$$
$$\cdots\int_E p(t_{k-1},x_{k-1},t_k,dx_k)f_k(x_k).$$

Proof. See, e.g., Revuz-Yor (1991, p. 76). □

Theorem 2.113. *Let* E *be a Polish space endowed with the* σ*-algebra* $\mathcal{B}_E$ *of its Borel sets,* P_0 *a probability measure on* $\mathcal{B}_E$, *and* $p(r,x,s,A), t_0 \leq r < s \leq T, x \in E, A \in \mathcal{B}_E$ *a Markov transition probability function. Then there exists a unique (in the sense of equivalence) Markov process* $(X_t)_{t\in[t_0,T]}$ *valued in* E, *with* P_0 *as its initial distribution and* p *as its transition probability.*

Proof. See, e.g., Ash and Gardner (1975), Dynkin (1965), Applebaum (2004, p. 124), and Kallenberg (1997, p. 120). □

Remark 2.114. From Theorem 2.113 we can deduce that

$$p(s,x,t,A) = P(X_t \in A|X_s = x), \text{a.s.} \qquad t_0 \leq s \leq t \leq T, x \in E, A \in \mathcal{B}_E.$$

Semigroups Associated with Markov Transition Probability Functions

In this section we will consider the case $E = \mathbb{R}$ as a technical simplification.

Let $BC(\mathbb{R})$ be the space of all continuous and bounded functions on $\mathbb{R}$, endowed with the norm $\|f\| = \sup_{x\in\mathbb{R}} |f(x)|(< \infty)$, and let $p(s, x, t, A)$ be a transition probability function ($t_0 \leq s < t \leq T, x \in \mathbb{R}, A \in \mathcal{B}_{\mathbb{R}}$). We consider the operator

$$T_{s,t} : BC(\mathbb{R}) \to BC(\mathbb{R}), \qquad t_0 \leq s < t \leq T,$$

defined by assigning, for all $f \in BC(\mathbb{R})$,

$$(T_{s,t}f)(x) = \int_{\mathbb{R}} f(y)p(s, x, t, dy) = E[f(X(t))|X(s) = x].$$

Proposition 2.115. *The family* $\{T_{s,t}\}_{t_0\leq s\leq t\leq T}$ *associated with the transition probability function* $p(s, x, t, A)$ *(or with its corresponding Markov process) is a* semigroup of linear operators *on* $BC(\mathbb{R})$, *i.e., it satisfies the following properties.*

1. *For any* $t_0 \leq s \leq t \leq T$, $T_{s,t}$ *is a linear operator on* $BC(\mathbb{R})$.
2. *For any* $t_0 \leq s \leq T$, $T_{s,s} = I$ *(the identity operator).*
3. *For any* $t_0 \leq s \leq t \leq T$, $T_{s,t}\, 1 = 1$.
4. *For any* $t_0 \leq s \leq t \leq T$, $\|T_{s,t}\| \leq 1$ *(contraction semigroup).*
5. *For any* $t_0 \leq s \leq t \leq T$, *and* $f \in BC(\mathbb{R})$, $f \geq 0$ *implies* $T_{s,t}f \geq 0$.
6. *For any* $t_0 \leq r \leq s \leq t \leq T$, $T_{r,s}T_{s,t} = T_{r,t}$ *(Chapman–Kolmogorov).*

Proof. All the preceding statements, apart from 4 and 6, are a direct consequence of the definitions that we are going to prove.

Proof of 4: Let $t_0 \leq s \leq t \leq T$, and $f \in BC(\mathbb{R})$;

$$\begin{aligned}\|T_{s,t}f\| &= \sup_{x\in\mathbb{R}} |E(f(X(t))|X(s) = x)| \\ &\leq \sup_{x\in\mathbb{R}} E(|f(X(t))|\ |X(s) = x) \\ &\leq \sup_{x\in\mathbb{R}} |f(x)| \sup_{x\in\mathbb{R}} E(1|X(s) = x) \\ &= \|f\|\, 1 = \|f\|,\end{aligned}$$

as stated. This fact lets us claim in particular that indeed $T_{s,t} : BC(\mathbb{R}) \to BC(\mathbb{R})$, for $t_0 \leq s < t \leq T$.

Proof of 6: Let $t_0 \leq r \leq s \leq t \leq T$, and $f \in BC(\mathbb{R})$; for any $x \in \mathbb{R}$

$$
\begin{aligned}
(T_{r,t}f)(x) &= E[f(X(t))|X(r)=x] \\
\text{(by the tower property)} &= E[E[f(X(t))|\mathcal{F}_s]|X(r)=x] \\
\text{(since } \mathcal{F}_r \subset \mathcal{F}_s) &= E[E[f(X(t))|X(s)]|X(r)=x] \\
&= E[(T_{s,t}f)(X(s))|X(r)=x] \\
&= (T_{r,s}(T_{s,t}f))(x),
\end{aligned}
$$

as stated. □

As the transition probability function $p(s, x, t, A)$ defines the semigroup $\{T_{s,t}\}_{t_0 \leq s \leq t \leq T}$ associated with it, conversely we may obtain the transition probability function from the semigroup, since we may easily recognize that

$$p(s, x, t, A) = P(X_t \in A | X_s = x) = (T_{s,t} I_A)(x) \text{ a.s.}$$

for $t_0 \leq s \leq t \leq T, x \in \mathbb{R}, A \in \mathcal{B}_{\mathbb{R}}$.

We may now finally prove the following proposition.

Proposition 2.116. *Let X be a real-valued Markov process, indexed in $\mathbb{R}$; the function p defined in (2.15) satisfies the so-called Chapman–Kolmogorov equation, i.e., for all $x \in \mathbb{R}$, for all $(s, r, t) \in [t_0, T] \times [t_0, T] \times [t_0, T]$, $s \leq r \leq t$, and for all $A \in \mathcal{B}_{\mathbb{R}}$*

$$p(s, x, t, A) = \int_{\mathbb{R}} p(s, x, r, dy) p(r, y, t, A) \text{ a.s.}$$

Proof. From definitions and Proposition 2.115 we easily obtain

$$
\begin{aligned}
p(s, x, t, A) &= (T_{s,t} I_A)(x) = (T_{s,r}(T_{r,t} I_A))(x) \\
&= \int_{\mathbb{R}} (T_{r,t} I_A)(y) p(s, x, r, dy) \\
&= \int_{\mathbb{R}} p(s, x, r, dy) p(r, y, t, A) \text{ a.s.}
\end{aligned}
$$

□

Definition 2.117. Let $(X_t)_{t \in \mathbb{R}_+}$ be a Markov process with transition probability function $p(s, x, t, A)$, and let $\{T_{s,t}\}$ $(s, t \in \mathbb{R}_+, s \leq t)$ be its associated semigroup. If, for all $f \in BC(\mathbb{R})$, the function

$$(t, x) \in \mathbb{R}_+ \times \mathbb{R} \to (T_{t,t+\lambda} f)(x) = \int_{\mathbb{R}} p(t, x, t+\lambda, dy) f(y) \in \mathbb{R}$$

is continuous for all $\lambda > 0$, then we say that the process satisfies the *Feller property*.

Theorem 2.118. *If $(X_t)_{t\in\mathbb{R}_+}$ is a Markov process with right-continuous trajectories satisfying the Feller property, then, for all $t \in \mathbb{R}_+$, $\mathcal{F}_t = \mathcal{F}_{t^+}$, where $\mathcal{F}_{t^+} = \bigcap_{t'>t} \sigma(X(s), 0 \leq s \leq t')$, and the filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ is right-continuous.*

Proof. See, e.g., Friedman (1975). □

Remark 2.119. It can be shown that $\mathcal{F}_{t^+}$ is a σ-algebra.

Example 2.120. Examples of processes with the Feller property, or simply *Feller processes*, include Wiener processes (Brownian motions), Poisson processes, and all Lévy processes (see later sections).

Definition 2.121. If $(X_t)_{t\in\mathbb{R}_+}$ is a Markov process with transition probability function p and associated semigroup $\{T_{s,t}\}$, then the operator

$$\mathcal{A}_s f = \lim_{h\downarrow 0} \frac{T_{s,s+h}f - f}{h}, \qquad s \geq 0, f \in BC(\mathbb{R})$$

is called the *infinitesimal generator of the Markov process* $(X_t)_{t\geq 0}$. Its domain $\mathcal{D}_{\mathcal{A}_s}$ consists of all $f \in BC(\mathbb{R})$ for which the preceding limit exists uniformly (and therefore in the norm of $BC(\mathbb{R})$) (see e.g., Feller 1971).

Remark 2.122. From the preceding definition we observe that

$$(\mathcal{A}_s f)(x) = \lim_{h\downarrow 0} \frac{1}{h} \int_{\mathbb{R}} [f(y) - f(x)] p(s, x, s+h, dy).$$

Remark 2.123. Up to this point we have been referring to the space $BC(\mathbb{R}^d)$ of bounded and continuous functions on $\mathbb{R}^d$. Actually, a more accurate analysis would require us to refer to its subspace $C_0(\mathbb{R}^d)$ of continuous functions, which tend to zero at infinity. This one is still a Banach space with the sup norm. In such a space it can be shown that a Feller semigroup is completely characterized by its infinitesimal generator (e.g., Kallenberg 1997, p. 317).

Examples of Stopping Times

Let $(\mathbf{X}_t)_{t\in\mathbb{R}_+}$ be a continuous Markov process taking values in $\mathbb{R}^v$, and suppose that the filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$, generated by the process, is right-continuous. Let $B \in \mathcal{B}_{\mathbb{R}^v} \setminus \{\emptyset\}$, and we define $T : \Omega \to \bar{\mathbb{R}}_+$ as

$$\forall \omega \in \Omega, \qquad T(\omega) = \begin{cases} \inf \{t \geq 0 | \mathbf{X}(t,\omega) \in B\} & \text{if the set is} \neq \emptyset, \\ +\infty & \text{if the set is} = \emptyset. \end{cases}$$

This gives rise to the following theorem.

Theorem 2.124. *If B is an open or closed subset of $\mathbb{R}^v$, then T is a stopping time.*

Proof. For B open, let $t \in \mathbb{R}_+$. In this case it can be shown that

$$\{T < t\} = \bigcup_{r<t, r\in\mathbb{Q}^+} \{\omega | \mathbf{X}(r,\omega) \in B\}.$$

Since $\mathbf{X}(r)$ is $\mathcal{F}$-measurable,

$$\{\omega | \mathbf{X}(r,\omega) \in B\} \in \mathcal{F}_r \subset \mathcal{F}_t \qquad \forall r < t, r \in \mathbb{Q}^+,$$

and therefore the (countable) union of such events will be an element of $\mathcal{F}_t$ as well, and thus $\{T < t\} \in \mathcal{F}_t$. Now, fixing $\delta > 0$ and $N \in \mathbb{N}$ such that $\delta > \frac{1}{N}$, we have that

$$\forall n \in \mathbb{N}, n \geq N\colon \qquad \left\{T < t + \frac{1}{n}\right\} \in \mathcal{F}_{t+\delta}.$$

Hence

$$\{T \leq t\} = \bigcap_{n=N}^{\infty} \left\{T < t + \frac{1}{n}\right\} \in \mathcal{F}_{t+\delta}$$

and, due to the arbitrary choice of δ, this results in

$$\{T \leq t\} \in \bigcap_{\delta>0} \mathcal{F}_{t+\delta} = \mathcal{F}_t^+ = \mathcal{F}_t.$$

For B closed, for all $n \in \mathbb{N}$, we define $V_n = \left\{\mathbf{x} \in \mathbb{R}^v | d(\mathbf{x}, B) < \frac{1}{n}\right\}$ and

$$T_n = \begin{cases} \inf\{t \geq 0 | \mathbf{X}(t,\omega) \in V_n\} & \text{if the set is } \neq \emptyset, \\ +\infty & \text{if the set is } = \emptyset. \end{cases}$$

It can be shown that $B = \bigcap_{n\in\mathbb{N}} V_n$ and $\{T \leq t\} = \bigcap_{n\in\mathbb{N}} \{T_n < t\}$, and, since (with V_n open) $\{T_n < t\} \in \mathcal{F}_{t^+}$, we finally get that $\{T \leq t\} \in \mathcal{F}_{t^+} = \mathcal{F}_t$. $\square$

Definition 2.125. The stopping time T is the *first hitting time* of B or, equivalently, the *first exit time* from $\mathbb{R}^v \setminus B$.

Definition 2.126. A Markov process $(X_t)_{t\in\mathbb{R}_+}$ with transition probability function $p(s, x, t, A)$ is said to have the *strong Markov property* if, for any stopping time T of the process and for all $A \in \mathcal{B}_\mathbb{R}$,

$$P(X(T + t) \in A | \mathcal{F}_T) = p(T, X(T), T + t, A) \qquad \text{a.s.} \tag{2.16}$$

Remark 2.127. Equation (2.16) is formally analogous to the Markov property

$$P(X(t) \in A | \mathcal{F}_s) = p(s, X(s), t, A) \text{ for } s < t,$$

with which it coincides when $T = s$ (constant).

Proposition 2.128. *Equation (2.16) is equivalent to the assertion that for all $f : \mathbb{R} \to \mathbb{R}$, measurable, bounded,*

$$E[f(X(T+t))|\mathcal{F}_T] = E[f(X(T+t))|X(T)] \qquad a.s.$$

Proof. See, e.g., Ash and Gardner (1975). □

Remark 2.129. By Proposition 2.43 and Theorem 2.49, if $(X_t)_{t\in\mathbb{R}_+}$ is right-continuous and if T is a finite stopping time of the process, then $X(T)$ is $\mathcal{F}_T$-measurable.

Lemma 2.130. *Every Markov process $(X_t)_{t\in\mathbb{R}_+}$ that satisfies the Feller property has the strong Markov property, at least for a discrete stopping time T.*

Proof. Let T be a discrete stopping time of the process $(X_t)_{t\in\mathbb{R}_+}$ and $\{t_j\}_{j\in\mathbb{N}}$ its codomain. Fixing a $j \in \mathbb{N}$ we have $\{T \leq t_j\} \in \mathcal{F}_{t_j}$ and $\{T < t_j\} = \bigcup_{t_l<t_j} \{T \leq t_l\} \in \mathcal{F}_{t_j}$. Therefore,

$$G_j \equiv \{T = t_j\} = \{T \leq t_j\} \setminus \{T < t_j\} \in \mathcal{F}_{t_j}$$

and

$$\forall t \in \mathbb{R}_+, \qquad G_j \cap \{T \leq t\} = \begin{cases} \emptyset & \text{for } t_j > t, \\ G_j & \text{for } t \geq t_j. \end{cases}$$

From this we obtain, for all $t \in \mathbb{R}_+$, $G_j \cap \{T \leq t\} \in \mathcal{F}_t$, that is, $G_j \in \mathcal{F}_T$. Proving (2.16) is equivalent to showing that if $t \in \mathbb{R}_+, A \in \mathcal{B}_\mathbb{R}$, then:

1. $p(T, X(T), T+t, A)$ is $\mathcal{F}_T$-measurable.
2. For all $E \in \mathcal{F}_T$, $P((X(T+t) \in A) \cap E) = \int_E p(T, X(T), T+t, A)dP$.

Before proving point 1, we will show that 2 holds. Let $E \in \mathcal{F}_T$; then, by $\Omega = \bigcup_{j\in\mathbb{N}} G_j$, it follows that

$$\begin{aligned} P((X(T+t) \in A) \cap E) &= \sum_{j\in\mathbb{N}} P((X(T+t) \in A) \cap E \cap G_j) \\ &= \sum_{j\in\mathbb{N}} P((X(T+t) \in A) \cap E \cap (T = t_j)) \\ &= \sum_{j\in\mathbb{N}} P((X(t+t_j) \in A) \cap E \cap (T = t_j)) \\ &= \sum_{j\in\mathbb{N}} P((X(t+t_j) \in A) \cap E \cap G_j). \end{aligned} \qquad (2.17)$$

But

$$E \cap G_j = E \cap (\{T \leq t_j\} \setminus \{T < t_j\}) \in \mathcal{F}_{t_j}$$

(in fact, $E \cap \{T \leq t_j\} \in \mathcal{F}_{t_j}$ following point 4 of Theorem 2.48), and therefore

$$P((X(t+t_j) \in A) \cap E \cap G_j) = \int_{E \cap G_j} P(X(t+t_j) \in A | \mathcal{F}_{t_j}) dP.$$

Moreover, by the Markov property,

$$P(X(t+t_j) \in A | \mathcal{F}_{t_j}) = p(t_j, X(t_j), t_j + t, A) \text{ a.s.} \tag{2.18}$$

Using (2.17) and (2.18), we obtain

$$\begin{aligned} P((X(T+t) \in A) \cap E) &= \bigcup_{j \in \mathbb{N}} \int_{E \cap G_j} p(t_j, X(t_j), t_j + t, A) dP \\ &= \bigcup_{j \in \mathbb{N}} \int_{E \cap \{T = t_j\}} p(t_j, X(t_j), t_j + t, A) dP \\ &= \bigcup_{j \in \mathbb{N}} \int_{E \cap \{T = t_j\}} p(T, X(T), T + t, A) dP \\ &= \int_E p(T, X(T), T + t, A) dP. \end{aligned}$$

For the proof of 1, we now observe that, by the Feller property, the mapping

$$(r, z) \in \mathbb{R}_+ \times \mathbb{R} \to \int_{\mathbb{R}} p(r, z, r + t, dy) f(y) \in \mathbb{R}$$

is continuous [for $f \in BC(\mathbb{R})$]. Furthermore, T and $X(T)$ are $\mathcal{F}_T$-measurable, and therefore the mapping

$$\omega \in \Omega \to (T(\omega), X(T(\omega), \omega))$$

is $\mathcal{F}_T$-measurable. Hence the composite of the two mappings

$$\omega \in \Omega \to \int_{\mathbb{R}} p(T, X(T), T + t, dy) f(y) \in \mathbb{R}$$

is $\mathcal{F}_T$-measurable [for $f \in BC(\mathbb{R})$]. Now let $(f_m)_{m \in \mathbb{N}} \in (BC(\mathbb{R}))^{\mathbb{N}}$ be a sequence of uniformly bounded functions such that $\lim_{m \to \infty} f_m = I_A$. Then, from our previous observations,

$$\forall m \in \mathbb{N}, \qquad \int_{\mathbb{R}} p(T, X(T), T + t, dy) f_m(y)$$

is $\mathcal{F}_T$-measurable and, following Lebesgue's theorem on integral limits, we get

$$\int_{\mathbb{R}} p(T, X(T), T + t, dy) I_A(y) = \lim_{m \to \infty} \int_{\mathbb{R}} p(T, X(T), T + t, dy) f_m(y),$$

and thus

$$p(T, X(T), T+t, A) = \int_{\mathbb{R}} p(T, X(T), T+t, dy) I_A(y)$$

is $\mathcal{F}_T$-measurable. □

Before generalizing Lemma 2.130, we assert the following lemma.

Lemma 2.131. *If T is a stopping time of the stochastic process $(X_t)_{t\in\mathbb{R}_+}$, then there exists a sequence of stopping times $(T_n)_{n\in\mathbb{N}}$ such that:*

1. *For all $n \in \mathbb{N}$, T_n has a codomain that is at most countable.*
2. *For all $n \in \mathbb{N}$, $T_n \geq T$.*
3. *$T_n \downarrow T$ almost surely for $n \to \infty$.*

Moreover, $\{T_n = \infty\} = \{T = \infty\}$ for every n.

Proof. See, e.g., Friedman (1975). □

Theorem 2.132. *If $(X_t)_{t\in\mathbb{R}_+}$ is a right-continuous Markov process that satisfies the Feller property, then it satisfies the strong Markov property.*

Proof. Let T be a finite stopping time of the process $(X_t)_{t\in\mathbb{R}_+}$ and $(T_n)_{n\in\mathbb{N}}$ a sequence of stopping times satisfying properties 1–3 of Lemma 2.131 with respect to T. We observe that, for all $n \in \mathbb{N}$, $\mathcal{F}_T \subset \mathcal{F}_{T_n}$. In fact, if $A \in \mathcal{F}_T$, then

$$\forall t \in \mathbb{R}_+, \qquad A \cap \{T_n \leq t\} = (A \cap \{T \leq t\}) \cap \{T_n \leq t\} \in \mathcal{F}_t,$$

provided that $A \cap \{T \leq t\} \in \mathcal{F}_t, \{T_n \leq t\} \in \mathcal{F}_t$. Just like for Lemma 2.130, we will need to show that points 1 and 2 of its proof hold in this present case. Following Proposition 2.128, point 2 is equivalent to asserting that for all $E \in \mathcal{F}_T$ and all $f \in BC(\mathbb{R})$:

$$\int_E f(X(T+t))dP = \int_E dP \int_{\mathbb{R}} p(T, X(T), T+t, dy) f(y). \tag{2.19}$$

Then, by Proposition 2.43, for all $n \in \mathbb{N}$, we have that for all $E \in \mathcal{F}_{T_n}$ and all $f \in BC(\mathbb{R})$

$$\int_E f(X(T_n+t))dP = \int_E dP \int_{\mathbb{R}} p(T_n, X(T_n), T_n+t, dy) f(y).$$

Moreover, since $T_n \downarrow T$ for $n \to \infty$ and by the right-continuity of process X, it follows that

$$X(T_n) \to X(T) \text{ for } n \to \infty.$$

From the continuity[5] of the mapping

$$(\lambda, x) \in \mathbb{R}_+ \times \mathbb{R} \to \int_{\mathbb{R}} p(\lambda, x, \lambda + t, \lambda y) f(y) \text{ for } f \in BC(\mathbb{R}),$$

we have that, for $n \to \infty$,

$$\int_{\mathbb{R}} p(T_n, X(T_n), T_n + t, dy) f(y) \to \int_{\mathbb{R}} p(T, X(T), T + t, dy) f(y). \tag{2.20}$$

On the other hand, if f is continuous, then we also get

$$f(X(T_n + t)) \to f(X(T + t)) \text{ for } n \to \infty. \tag{2.21}$$

Therefore, if $E \in \mathcal{F}_T$ and $f \in BC(\mathbb{R})$, then $E \in \mathcal{F}_{T_n}$ for all n, and we have

$$\lim_{n\to\infty} \int_E f(X(T_n + t)) dP = \lim_{n\to\infty} \int_E dP \int_{\mathbb{R}} p(T_n, X(T_n), T_n + t, dy) f(y).$$

Since f and p are bounded, following Lebesgue's theorem, we can take the limit of the integral and then (2.19) follows from (2.20) and (2.21). The proof of point 1 is entirely analogous to the proof of Lemma 2.131. □

The preceding results may be extended to more general, possibly uncountable, state spaces. In particular, we will assume that E is a subset of $\mathbb{R}^d$ for $d \in \mathbb{N}^*$.

Time-Homogeneous Markov Processes

An important class of Markov processes is the time-homogeneous case.

Definition 2.133. A Markov process $(X_t)_{t\in[t_0,T]}$ is said to be *time-homogeneous* if the transition probability functions $p(s, x, t, A)$ depend on t and s only through their difference $t - s$. Therefore, for all $(s, t) \in [t_0, T]^2$, $s < t$, for all $u \in [0, T - t]$, for all $A \in \mathcal{B}_{\mathbb{R}}$, and for all $x \in \mathbb{R}$:

$$p(s, x, t, A) = p(s + u, x, t + u, A) \qquad \text{a.s.}$$

Remark 2.134. If $(X_t)_{t\in[t_0,T]}$ is a homogeneous Markov process with transition probability function p, then, for all $(s, t) \in [t_0, T]^2, s < t$, for all $A \in \mathcal{B}_{\mathbb{R}}$, and for all $x \in \mathbb{R}$, we obtain

$$p(t_0, x, t_0 + t - s, A) = p(s, x, t, A) \text{ a.s.},$$

where $p(t_0, x, t_0 + t - s, A)$ is denoted by $p(\bar{t}, x, A)$, with $\bar{t} = (t - s) \in [0, T - t_0], x \in \mathbb{R}, A \in \mathcal{B}_{\mathbb{R}}$.

[5] By the Feller property.

If we consider the time-homogeneous case, a Markov process $(X_t)_{t\in\mathbb{R}_+}$ on $(E,\mathcal{B}_E)$ will be defined in terms of a transition kernel $p(t,x,B)$ for $t\in\mathbb{R}_+$, $x\in E$, $B\in\mathcal{B}_E$, such that

$$p(h,X_t,B)=P(X_{t+h}\in B|\mathcal{F}_t)\qquad \forall t,h\in\mathbb{R}_+, B\in\mathcal{B}_E,$$

given that $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ is the natural filtration of the process. Equivalently, if we denote by $BC(E)$ the Banach space of all continuous and bounded functions on E, endowed with the *sup norm*, then

$$E[g(X_{t+h})|\mathcal{F}_t]=\int_E g(y)p(h,X_t,dy)\qquad \forall t,h\in\mathbb{R}_+, g\in BC(E).$$

In this case the transition semigroup of the process is such that

$$T_{s,t}=T_{0,t-s}=:T(t-s)$$

for any $s,t\in\mathbb{R}_+,\ s\le t$, which defines a one-parameter contraction semigroup $(T(t),t\in\mathbb{R}_+)$ on $BC(E)$; it is then such that

$$T(t)g(x):=\int_E g(y)p(t,x,dy)=E[g(X_t)|X_0=x],\qquad x\in E,$$

for any $g\in BC(E)$.

Hence in this case the semigroup property reduces to

$$T(s+t)=T(s)T(t)=T(t)T(s)$$

for any $s,t\in\mathbb{R}_+$.

Up to now we have referred to $BC(\mathbb{R})$, i.e., the family of bounded and continuous functions on $\mathbb{R}$. For various reasons, as the reader may see later, it is more convenient to refer to its Banach subspace $C_0(\mathbb{R})$, the family of continuous functions vanishing at infinity, since it has nicer analytical properties.

Definition 2.135. Let $(T(t))_{t\in\mathbb{R}_+}$ be the transition semigroup associated with a time-homogeneous Markov process $X=(X_t)_{t\in\mathbb{R}_+}$. We say that X is a *Feller process* if the following statements hold:

(i) $T(t)(C_0(\mathbb{R}))\subset C_0(\mathbb{R})$ for all $t\in\mathbb{R}_+$.
(ii) $\lim_{t\to 0}\|T(t)f-f\|=0$ for all $f\in C_0(\mathbb{R})$.

In this case we say that the semigroup $(T(t))_{t\in\mathbb{R}_+}$ is a *Feller semigroup.*

Proposition 2.136. *For any Feller semigroup on $C_0(\mathbb{R})$ there exists a unique time-homogeneous transition probability measure $p(t,x,B)$ on $(\mathbb{R},\mathcal{B}_\mathbb{R})$ such that, for all $f\in C_0(\mathbb{R})$,*

$$T(t)f(x)=\int_\mathbb{R} f(y)p(t,x,dy),\quad x\in\mathbb{R},\quad t\in\mathbb{R}_+.$$

Proof. See, e.g., Revuz-Yor (1991, p. 83). □

Definition 2.137. A time-homogeneous transition probability measure associated to a Feller semigroup is called a *Feller transition function.*

For time-homogeneous Markov processes the infinitesimal generator will be time independent. It is defined as

$$\mathcal{A}g = \lim_{t \to 0+} \frac{1}{t}(T(t)g - g)$$

for $g \in \mathcal{D}(\mathcal{A})$, the subset of $BC(E)$ for which the preceding limit exists, in $BC(E)$, with respect to the sup norm. Given the preceding definitions, it is obvious that for all $g \in \mathcal{D}(\mathcal{A})$,

$$\mathcal{A}g(x) = \lim_{t \to 0+} \frac{1}{t} E[g(X_t) - g(X_0)|X_0 = x], \qquad x \in E.$$

If $(T(t), t \in \mathbb{R}_+)$ is the contraction semigroup associated with a Markov process, it is not difficult to show that the mapping $t \to T(t)g$ is right-continuous in $t \in \mathbb{R}_+$ provided that $g \in BC(E)$ is such that the mapping $t \to T(t)g$ is right-continuous in $t = 0$.

The following properties hold, by considering Riemann integrals and strong derivatives (Applebaum, 2004, p. 129).

1. For any $t \geq 0: \ T(t)\mathcal{D}(\mathcal{A}) \subset \mathcal{D}(\mathcal{A})$.
2. For any $t \geq 0$ and for any $g \in \mathcal{D}(\mathcal{A})$: $T(t)\mathcal{A}g = \mathcal{A}T(t)g$.
3. For any $t \geq 0$ and for any $g \in \mathcal{D}(\mathcal{A})$: $\int_0^t T(s)g ds \in \mathcal{D}(\mathcal{A})$.
4. For any $t \geq 0$ and for any $g \in \mathcal{D}(\mathcal{A})$:
$$T(t)g - g = \mathcal{A}\int_0^t T(s)gds = \int_0^t \mathcal{A}T(s)gds = \int_0^t T(s)\mathcal{A}gds.$$
5. For any $t \geq 0$ and for any $g \in \mathcal{D}(\mathcal{A})$:
$$\frac{d}{dt}[T(t)g] = \mathcal{A}[T(t)g] = T(t)[\mathcal{A}g].$$
6. For any $g \in \mathcal{D}(\mathcal{A})$, the function $t \in \mathbb{R}_+ \mapsto T(t)g \equiv u(t) \in \mathcal{D}(\mathcal{A})$ is a solution of the following initial value problem in the Banach space $BC(\mathbb{R}^d)$:
$$\begin{cases} \frac{d}{dt}u(t) = \mathcal{A}u(t), \\ u(0) \quad = g. \end{cases}$$

These results justify the notation $T(t) = e^{t\mathcal{A}}$.

The following so-called *Dynkin's formula* establishes a fundamental link between Markov processes and martingales (Rogers and Williams 1994, p. 253).

Given a process $(X_t)_{t\in\mathbb{R}_+}$, we will denote by $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ its natural filtration.

Theorem 2.138. *Assume $(X_t)_{t\in\mathbb{R}_+}$ is a Markov process on $(E,\mathcal{B}_E)$, with transition kernel $p(t,x,B)$, $t\in\mathbb{R}_+$, $x\in E$, $B\in\mathcal{B}_E$. Let $(T(t),t\in\mathbb{R}_+)$ denote its transition semigroup and $\mathcal{A}$ its infinitesimal generator. Then, for any $g\in\mathcal{D}(\mathcal{A})$, the stochastic process*

$$M(t) := g(X_t) - g(X_0) - \int_0^t \mathcal{A}g(X_s)ds$$

is an $\mathcal{F}_t$-martingale (indeed a zero-mean martingale).

Proof. The following equations hold:

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

Now, thanks to the Markov property,

$$E\left[g(X_{t+h})|\,\mathcal{F}_t\right] = E\left[g(X_{t+h})|\,X_t\right] = T(h)g(X_t),$$

and

$$\begin{aligned}
E\left[\int_t^{t+h}\mathcal{A}g(X_s)ds|\mathcal{F}_t\right] &= \int_t^{t+h} ds\, E[\mathcal{A}g(X_s)|\mathcal{F}_t]\\
&= \int_t^{t+h} ds\, E[\mathcal{A}g(X_s)|X_t] = \int_t^{t+h} ds\, T(s-t)\mathcal{A}g(X_t)\\
&= \int_0^h ds\, T(s)\mathcal{A}g(X_t) = \int_0^h ds\, \mathcal{A}T(s)g(X_t)\\
&= \int_0^h d[T(s)g(X_t)] = T(h)g(X_t) - T(0)g(X_t)\\
&= T(h)g(X_t) - g(X_t).
\end{aligned}$$

As a consequence

$$\begin{aligned}
&E[M(t+h)|\mathcal{F}_t] + g(X_0)\\
&= T(h)g(X_t) - T(h)g(X_t) + g(X_t) - \int_0^t \mathcal{A}g(X_s)ds\\
&= g(X_t) - \int_0^t \mathcal{A}g(X_s)ds = M(t) + g(X_0).
\end{aligned}$$

□

The next proposition shows that a Markov process is indeed characterized by its infinitesimal generator via a martingale problem (e.g., Rogers and Williams 1994, p. 253).

Theorem 2.139 (Martingale problem for Markov processes). *If an RCLL Markov process $(X_t)_{t\in\mathbb{R}_+}$ is such that*

$$g(X_t) - g(X_0) - \int_0^t \mathcal{A}g(X_s)ds$$

is an $\mathcal{F}_t$-martingale for any function $g \in \mathcal{D}(\mathcal{A})$, where $\mathcal{A}$ is the infinitesimal generator of a contraction semigroup on E, then X_t is equivalent to a Markov process having $\mathcal{A}$ as its infinitesimal generator.

Remark 2.140. Note that, from

$$M(t) := g(X_t) - g(X_0) - \int_0^t \mathcal{A}g(X_s)ds$$

one may derive

$$g(X_t) - g(X_0) = \int_0^t \mathcal{A}g(X_s)ds + M(t).$$

Formally, by a suitable definition of differential of a martingale, this may be rewritten as

$$dg(X_t) = \mathcal{A}g(X_t) + dM(t).$$

Hence, apart from the "noise" $M(t)$, the evolution of any function $g(X_t)$ of a Markov process $\{X_t, t \in \mathbb{R}_+\}$ is determined by its infinitesimal generator.

Theorem 2.141. *Let $\{X_t, t \in \mathbb{R}_+\}$ be a Feller process on $\mathbb{R}$ having infinitesimal generator $\mathcal{A}$ with domain $\mathcal{D}_{\mathcal{A}}$. If $g \in C_0(\mathbb{R})$ and there exists an $f \in C_0(\mathbb{R})$ such that*

$$M(t) := g(X_t) - g(X_0) - \int_0^t f(X_s)ds, \ t \in \mathbb{R}_+,$$

is an $\mathcal{F}_t$-martingale, then $g \in \mathcal{D}_{\mathcal{A}}$, and $f = Ag$.

Proof. See, e.g., Revuz-Yor (1991, p. 262) □

The preceding results lead to an extension of the concept of infinitesimal generator called an *extended infinitesimal generator* (Revuz-Yor 1991, p. 263).

In what follows we shall denote by $C_K^p(\mathbb{R}),\ p \in \overline{\mathbb{R}}_+^*$, the set of real functions with compact support, which are continuous with their derivatives up to the order p.

Theorem 2.142. *Let $\{X_t, t \in \mathbb{R}_+\}$ be a Feller process on $\mathbb{R}$, having infinitesimal generator $\mathcal{A}$ with domain $\mathcal{D}_{\mathcal{A}}$, such that $C_K^\infty(\mathbb{R}) \subset \mathcal{D}_{\mathcal{A}}$. Then* *(i)* $C_K^2(\mathbb{R}) \subset \mathcal{D}_\mathcal{A}$.

(ii) *For any relatively compact open set* U *there exist functions* a, b, *and* $c \leq 0$ *on* U, *and a kernel measure* $N(x, B)$, $x \in U$, $B \in \mathbb{R} - \{0\}$, *which is a Radon measure for any* $x \in U$ *such that, for* $f \in C_K^2(\mathbb{R})$ *and* $x \in U$,

$$(\mathcal{A}f)(x) = \frac{1}{2}b^2(x)\frac{\partial^2 f}{\partial x^2} + a(x)\frac{\partial f}{\partial x} + c(x) + \int_{\mathbb{R}-\{0\}} \left[f(y) - f(x) - I_U(y)(y-x)\frac{\partial f}{\partial x}\right] N(x, dy).$$

If the process $\{X_t, t \in \mathbb{R}_+\}$ *has continuous paths, then*

$$(\mathcal{A}f)(x) = \frac{1}{2}b^2(x)\frac{\partial^2 f}{\partial x^2} + a(x)\frac{\partial f}{\partial x} + c(x).$$

Proof. See, e.g., Revuz-Yor (1991, p. 267) □

Example 2.143. A Poisson process (see the following section for more details) is an integer-valued Markov process $(N_t)_{t \in \mathbb{R}_+}$. If its intensity parameter is $\lambda > 0$, then the process $(X_t)_{t \in \mathbb{R}_+}$, defined by $X_t = N_t - \lambda t$, is a stationary Markov process with independent increments. The transition kernel of X_t is

$$p(h, x, B) = \sum_{k=0}^{\infty} \frac{(\lambda h)^k}{k!} e^{-\lambda h} I_{\{x+k-\lambda h \in B\}} \text{ for } x \in \mathbb{N}, h \in \mathbb{R}_+, B \subset \mathbb{N}.$$

Its transition semigroup is then

$$T(h)g(x) = \sum_{k=0}^{\infty} \frac{(\lambda h)^k}{k!} e^{-\lambda h} g(x + k - \lambda h) \quad \text{for } x \in \mathbb{N}, g \in BC(\mathbb{R}).$$

The infinitesimal generator is then

$$\mathcal{A}g(x) = \lambda(g(x+1) - g(x)) - \lambda g'(x+).$$

According to previous theorems,

$$M(t) = g(X_t) - \int_0^t ds(\lambda(g(X_s + 1) - g(X_s)) - \lambda g'(X_s+)$$

is a martingale for any $g \in BC(\mathbb{R})$ (where $g(0) = 0$).

Holding Times for a Markov Process

Suppose that a Markov process $(X_t)_{t \in \mathbb{R}_+}$ on $\mathbb{R}$ starts at a point x. We wish to evaluate the probability distribution of the *holding time* at x, i.e., the time it spends in that state before leaving it:

$$\tau_x := \inf\{t \in \mathbb{R}_+ \mid X_s = x, X_{t+s} \neq x\}$$

for any given $s \in \mathbb{R}_+$.

The following proposition holds.

Proposition 2.144. *For any right-continuous time-homogeneous Markov process we have*

$$F_x(t) = P(\tau_x \leq t) = 1 - \exp\{-c_x t\}, \quad t \in \mathbb{R}_+,$$

for some $c_x \in [0, +\infty]$.

Proof. See, e.g., Lamperti (1977, p. 195). □

Markov Diffusion Processes

Definition 2.145. A Markov process on $\mathbb{R}$ with transition probability function $p(s, x, t, A)$ is called a *diffusion process* if

1. It has a.s. continuous trajectories.
2. For all $\epsilon > 0$, for all $t \geq 0$, and for all $x \in \mathbb{R}$: $\lim_{h\downarrow 0} \frac{1}{h} \int_{|x-y|>\epsilon} p(t, x, t+h, dy) = 0$.
3. There exist $a(t, x)$ and $b(t, x)$ such that, for all $\epsilon > 0$, for all $t \geq 0$, and for all $x \in \mathbb{R}$,

$$\lim_{h\downarrow 0} \frac{1}{h} \int_{|x-y|<\epsilon} (y-x) p(t, x, t+h, dy) = a(t, x),$$
$$\lim_{h\downarrow 0} \frac{1}{h} \int_{|x-y|<\epsilon} (y-x)^2 p(t, x, t+h, dy) = b(t, x),$$

where $a(t, x)$ is the *drift coefficient* and $b(t, x)$ the *diffusion coefficient* of the process.

Lemma 2.146. *Conditions 1 and 2 of Definition 2.145 are satisfied if*

1.* *There exists a* $\delta > 0$ *such that, for all* $t \geq 0$ *and for all* $x \in \mathbb{R}$, $\lim_{h\downarrow 0} \frac{1}{h} \int_{\mathbb{R}} |x-y|^{2+\delta} p(t, x, t+h, dy) = 0$.

2.* *For all* $t \geq 0$ *and for all* $x \in \mathbb{R}$,

$$\lim_{h\downarrow 0} \frac{1}{h} \int_{\mathbb{R}} (y-x) p(t, x, t+h, dy) = a(t, x),$$
$$\lim_{h\downarrow 0} \frac{1}{h} \int_{\mathbb{R}} (y-x)^2 p(t, x, t+h, dy) = b(t, x).$$

Proof. We fix $\epsilon > 0$, $x \in \mathbb{R}$, $|x-y| > \epsilon \Rightarrow \dfrac{|y-x|^{2+\delta}}{\epsilon^{2+\delta}} \geq 1$, and hence

$$\begin{aligned}\frac{1}{h} \int_{|x-y|>\epsilon} p(t, x, t+h, dy) &\leq \frac{1}{h\epsilon^{2+\delta}} \int_{|x-y|>\epsilon} |y-x|^{2+\delta} p(t, x, t+h, dy) \\ &\leq \frac{1}{h\epsilon^{2+\delta}} \int_{\mathbb{R}} |y-x|^{2+\delta} p(t, x, t+h, dy).\end{aligned}$$

From this, due to 1^*, point 1 of Definition 2.145 follows. Analogously, for $j = 1, 2$,

$$\frac{1}{h}\int_{|x-y|>\epsilon}|y-x|^j p(t,x,t+h,dy) \leq \frac{1}{h\epsilon^{2+\delta-j}}\int_{\mathbb{R}}|y-x|^{2+\delta}p(t,x,t+h,dy),$$

and again from 1^* we obtain

$$\lim_{h\downarrow 0}\frac{1}{h}\int_{|x-y|>\epsilon}|y-x|^j p(t,x,t+h,dy) = 0.$$

Moreover,

$$\lim_{h\downarrow 0}\frac{1}{h}\int_{\mathbb{R}}|y-x|^j p(t,x,t+h,dy) = \lim_{h\downarrow 0}\frac{1}{h}\left(\int_{|x-y|>\epsilon}|y-x|^j p(t,x,t+h,dy) + \int_{|x-y|<\epsilon}|y-x|^j p(t,x,t+h,dy)\right),$$

which, along with 2^*, gives point 2 of Definition 2.145. □

Proposition 2.147. *If $(X_t)_{t\in\mathbb{R}_+}$ is a diffusion process with transition probability function p and drift and diffusion coefficients $a(x,t)$ and $b(x,t)$, respectively, and if $\mathcal{A}_s$ is the infinitesimal generator associated with p, then we have that*

$$(\mathcal{A}_s f)(x) = a(s,x)\frac{\partial f}{\partial x} + \frac{1}{2}b(s,x)\frac{\partial^2 f}{\partial x^2},$$

provided that f is bounded and twice continuously differentiable.

Proof. Let $f \in BC(\mathbb{R}) \cap C^2(\mathbb{R})$. From Taylor's formula we obtain

$$f(y) - f(x) = f'(x)(y-x) + \frac{1}{2}f''(x)(y-x)^2 + o(|y-x|^2) \tag{2.22}$$

for $|y-x| < \delta$ (which is in a suitable neighborhood of x), and thus

$$\begin{aligned}(\mathcal{A}_s f)(x) &= \lim_{h\downarrow 0}\frac{1}{h}\int_{\mathbb{R}}[f(y)-f(x)]p(s,x,s+h,dy)\\ &= \lim_{h\downarrow 0}\frac{1}{h}\int_{|y-x|<\delta}f'(x)(y-x)p(s,x,s+h,dy)\\ &\quad +\frac{1}{2}\lim_{h\downarrow 0}\frac{1}{h}\int_{|y-x|<\delta}f''(x)(y-x)^2 p(s,x,s+h,dy)\\ &\quad +\lim_{h\downarrow 0}\frac{1}{h}\int_{|y-x|<\delta}o(|y-x|^2)p(s,x,s+h,dy)\\ &\quad +\lim_{h\downarrow 0}\frac{1}{h}\int_{|y-x|\geq\delta}[f(y)-f(x)]p(s,x,s+h,dy).\end{aligned}$$

Because $f \in BC(\mathbb{R})$,

$$\lim_{h\downarrow 0} \frac{1}{h} \int_{|y-x|\geq\delta} [f(y) - f(x)]p(s, x, s+h, dy)$$
$$\leq \lim_{h\downarrow 0} \frac{1}{h} c \int_{|y-x|\geq\delta} p(s, x, s+h, dy) = 0,$$

by point 1 of Definition 2.145, where c is a constant. By point 2 of the same definition:

$$\lim_{h\downarrow 0} \frac{1}{h} \int_{|y-x|<\delta} f'(x)(y-x)p(s, x, s+h, dy)$$
$$= f'(x) \lim_{h\downarrow 0} \frac{1}{h} \int_{|y-x|<\delta} (y-x)p(s, x, s+h, dy)$$
$$= f'(x)a(t, x),$$

as well as

$$\frac{1}{2} \lim_{h\downarrow 0} \frac{1}{h} \int_{|y-x|<\delta} f''(x)(y-x)^2 p(s, x, s+h, dy) = \frac{1}{2} f''(x)b(x, t).$$

Fixing $\epsilon > 0$, we finally observe that if we choose δ such that Taylor's formula (2.22) holds, so that

$$|y-x| < \delta \Rightarrow \frac{o(|y-x|^2)}{|y-x|^2} < \epsilon,$$

we get

$$\lim_{h\downarrow 0} \frac{1}{h} \int_{|y-x|<\delta} o(|y-x|^2)p(s, x, s+h, dy)$$
$$\leq \lim_{h\downarrow 0} \frac{1}{h} \int_{|y-x|<\delta} \epsilon |y-x|^2 p(s, x, s+h, dy)$$
$$= \epsilon b(t, x)$$

and, from the fact that ϵ is arbitrary, we conclude that

$$\lim_{h\downarrow 0} \frac{1}{h} \int_{|y-x|<\delta} o(|y-x|^2)p(s, x, s+h, dy) = 0.$$

□

A detailed account of conditions that ensure a.s. path continuity of the trajectories of Markov processes can be found in Lamperti (1977, p. 188).

Markov Jump Processes

Consider a Markov process $(X_t)_{t\in\mathbb{R}_+}$ valued in a countable set E (say, $\mathbb{N}$ or $\mathbb{Z}$). In such a case it is sufficient (with respect to Theorem 2.113) to provide the so-called *one-point* transition probability function

$$p_{ij}(s,t) := p(s,i,t,j) := P(X_t = j|X_s = i)$$

for $t_0 \leq s < t$, $i,j \in E$. It follows from the general structure of Markov processes that the one-point transition probabilities satisfy the following relations:

(a) $p_{ij}(s,t) \geq 0$
(b) $\sum_{j\in E} p_{ij}(s,t) = 1$
(c) $p_{ij}(s,t) = \sum_{k\in E} p_{ik}(s,r)p_{kj}(r,t)$

provided $t_0 \leq s \leq r \leq t$, in $\mathbb{R}_+$, and $i,j \in E$. To these three conditions we need to add

(d)
$$\lim_{t\to s+} p_{ij}(s,t) = p_{ij}(s,s) = \delta_{ij} = \begin{cases} 1 & \text{for } i = j, \\ 0 & \text{for } i \neq j. \end{cases}$$

The time-homogeneous case gives transition probabilities $(\tilde{p}_{ij}(t))_{t\in\mathbb{R}_+}$ such that

$$p_{ij}(s,t) = \tilde{p}_{ij}(t-s), \qquad s \leq t.$$

Henceforth we shall limit our analysis to the time-homogeneous case whose transition probabilities will be denoted by $(p_{ij}(t))_{t\in\mathbb{R}_+}$. The following theorems hold (Gihman and Skorohod 1974, pp. 304–306).

Theorem 2.148. *The transition probabilities $(p_{ij}(t))_{t\in\mathbb{R}_+}$ of a homogeneous Markov process on a countable state space E are uniformly continuous in $t \in \mathbb{R}_+$ for any fixed $i,j \in E$.*

Theorem 2.149. *The limit*

$$q_i = \lim_{h\to 0+} \frac{1 - p_{ii}(h)}{h} \leq +\infty$$

always exists (finite or not), and for arbitrary $t > 0$:

$$\frac{1 - p_{ii}(t)}{t} \leq q_i.$$

If $q_i < +\infty$, then for all $t > 0$ the derivatives $p'_{ij}(t)$ exist for any $i,j \in E$ and are continuous. They satisfy the following relations:

1. $p'_{ij}(t+s) = \sum_{k\in E} p'_{ik}(t)p_{kj}(s)$
2. $\sum_{j\in E} p'_{ij}(t) = 0$
3. $\sum_{j\in E} |p'_{ij}(t)| \leq 2q_i$

In the following theorem the condition $q_i < +\infty$ is not required.

Theorem 2.150. *The limits*

$$\lim_{t\to 0+} \frac{p_{ij}(t)}{t} = p'_{ij}(0) =: q_{ij} < +\infty$$

always exist (finite) for any $i \neq j$.

As a consequence of Theorems 2.149 and 2.150, provided $q_i < +\infty$, we obtain evolution equations for $p_{ij}(t)$:

$$p'_{ij}(t) = \sum_{k\in E} q_{ik} p_{kj}(t),$$

with $q_{ii} = -q_i$. These equations are known as *Kolmogorov backward equations*. Consider the family of matrices $(P(t))_{t\in\mathbb{R}_+}$, with entries $(p_{ij}(t))_{t\in\mathbb{R}_+}$, for $i, j \in E$. We may rewrite conditions (c) and (d) in matrix form as follows:

(c′) $P(s+t) = P(s)P(t)$ for any $s, t \geq 0$
(d′) $\lim_{h\to 0+} P(h) = P(0) = I$

A family of stochastic matrices fulfilling conditions (c′) and (d′) is called a *matrix transition function*. If a matrix transition function satisfies the condition

$$\sum_{j\neq i} q_{ij} = -q_{ii} \equiv q_i < +\infty$$

for any $i \in E$, it is called *conservative*. The matrix $Q = (q_{ij})_{i,j\in E}$ is called the *intensity matrix*. The Kolmogorov backward equations can be rewritten in matrix form as

$$P'(t) = QP(t), \qquad t > 0,$$

subject to

$$P(0) = I.$$

If Q is a finite-dimensional matrix, then the function $exp\{tQ\}$ for $t > 0$ is well defined.

Theorem 2.151 (Karlin and Taylor 1975, p. 152). *If E is finite, then the matrix transition function can be represented in terms of its intensity matrix Q via*

$$P(t) = e^{tQ}, \qquad t \geq 0.$$

Given an intensity matrix Q of a conservative Markov jump process with stationary (time-homogeneous) transition probabilities, we have that (Doob 1953)

$$P(X_u = i \;\forall u \in]s, s+t] | X_s = i) = e^{-q_i t}$$

for every $s, t \in \mathbb{R}_+$, and state $i \in E$. This shows that the sojourn time in state i is exponentially distributed with parameter q_i. This is independent of the initial time $s \geq 0$.

Furthermore, let π_{ij}, $i \neq j$, be the conditional probability of a jump to state j, given that a jump from state i has occurred. It can be shown (Doob 1953) that

$$\pi_{ij} = \frac{q_{ij}}{q_i},$$

provided that $q_i > 0$. For $q_i = 0$, state i is *absorbing*, which obviously means that once state i is entered, the process remains there permanently. Indeed,

$$P(X_u = i, \text{ for all } u \in]s, s+t] | X_s = i) = e^{-q_i t} = 1$$

for all $t \geq 0$. A state i for which $q_i = +\infty$ is called an *instantaneous state*. The expected sojourn time in such a state is zero. A state i for which $0 \leq q_i < +\infty$ is called a *stable state*.

Example 2.152. If $(X_t)_{t\in\mathbb{R}_+}$ is a homogeneous Poisson process with intensity $\lambda > 0$, then

$$p_{ij}(t) = \begin{cases} e^{-\lambda t}\frac{(\lambda t)^{j-i}}{(j-i)!} & \text{for } j > i, \\ 0 & \text{otherwise.} \end{cases}$$

This implies that

$$q_{ij} = p'_{ij}(0) \begin{cases} \lambda & \text{for } j = i+1, \\ -\lambda & \text{for } j = i, \\ 0 & \text{otherwise.} \end{cases}$$

For the following result we refer again to Doob (1953).

Theorem 2.153. *For any $x \in E$ there exists a unique RCLL Markov process associated with a given intensity matrix Q and such that $P(X(0) = x) = 1$.*

Consider a time-homogeneous Markov jump process on a countable state space E with intensity matrix $Q = (q_{ij})_{i,j\in E}$. The matrix Q can be seen as a functional operator on E as follows. For any $f : E \to \mathbb{R}_+$ define

$$Q : f \to Q(f) = \sum_{j\in E} q_{ij} f(j) = \sum_{j\neq i} q_{ij}(f(j) - f(i)).$$

For f bounded in E we may define, for any $x \in E$,

$$\begin{aligned} &E_x[f(X(t+s))] - E_x[f(X(t))] \\ &= E_x[E_{X(t)}[f(X(s)) - f(X(0))]] \\ &= \sum_{j\neq i}(f(j) - f(i))P(X(s) = j|X(0) = i)P_x(X(t) = i). \end{aligned}$$

Assume we may interchange the derivative and sum of the series

$$\frac{d}{dt}E_x[f(X(t))] = \sum_{j \neq i} q_{ij}(f(j) - f(i))P_x(X(s) = i),$$

which can be written as

$$\frac{d}{dt}E_x[f(X(t))] = E_x[Q(f)(X(t))].$$

By returning to the integral formulation

$$E_x[f(X(t))] - E_x[f(X(0))] = \int_0^t E_x[Q(f)(X(s))]ds, \tag{2.23}$$

the preceding formula can be seen as a Dynkin formula for Markov jump processes in terms of the intensity matrix Q. Indeed, from Rogers and Williams (1994, pp. 30–37) we obtain the following theorem.

Theorem 2.154. *For any function $g \in C^{1,0}(\mathbb{R}_+ \times E)$ such that the mapping*

$$t \to \frac{\partial}{\partial t} g(t, x)$$

is continuous for all $x \in E$, the process

$$\left(g(t, X(t)) - g(0, X(0)) - \int_0^t \left(\frac{\partial g}{\partial t} + Q(g(s, \cdot)) \right) (s, X(s)) ds \right)_{t \in \mathbb{R}_+}$$

is a local martingale.

Corollary 2.155. *For any real function f defined on E, the process*

$$\left(f(X(t)) - f(X(0)) - \int_0^t Q(f)X(s)ds \right)_{t \in \mathbb{R}_+} \tag{2.24}$$

is a local martingale. Whenever the local martingale is a martingale, we recover (2.23).

Proposition 2.156 (Martingale problem for Markov jump processes). *Given an intensity matrix Q, if an RCLL Markov process $X \equiv (X(t))_{t \in \mathbb{R}_+}$ on E is such that the process (2.24) is a local martingale, then Q is the intensity matrix of the Markov process X.*

Further discussions on this topic may be found in Doob (1953) and Karlin and Taylor (1981) [an additional and updated source regarding discrete-space continuous-time Markov chains is Anderson (1991)]. For applications, see, for example, Robert (2003).

2.8 Brownian Motion and the Wiener Process

A small particle (e.g., a pollen corn) suspended in a liquid is subject to infinitely many collisions with atoms, and therefore it is impossible to observe its exact trajectory. With the help of a microscope it is only possible to confirm that the movement of the particle is entirely chaotic. This type of movement, discovered under similar circumstances by the botanist Robert Brown, is called Brownian motion. As its mathematical inventor Einstein had already observed, it is necessary to make approximations in order to describe the process. The formalized mathematical model defined on the basis of these facts is called a Wiener process. Henceforth, we will limit ourselves to the study of the one-dimensional Wiener process in $\mathbb{R}$, under the assumption that the three components determining its motion in space are independent.

Definition 2.157. The real-valued process $(W_t)_{t\in\mathbb{R}_+}$ is a *Wiener process* if it satisfies the following conditions:

1. $W_0 = 0$ almost surely.
2. $(W_t)_{t\in\mathbb{R}_+}$ is a process with independent increments.
3. $W_t - W_s$ is normally distributed with $N(0, t-s)$, $(0 \leq s < t)$.

Remark 2.158. From point 3 of Definition 2.157 it becomes obvious that every Wiener process is homogeneous.

Proposition 2.159. *If $(W_t)_{t\in\mathbb{R}_+}$ is a Wiener process, then*

1. $E[W_t] = 0 \qquad$ *for all* $t \in \mathbb{R}_+$
2. $K(s,t) = Cov[W_t, W_s] = \min\{s,t\}, \qquad s,t \in \mathbb{R}_+$

Proof.

1. By fixing $t \in \mathbb{R}$, we observe that $W_t = W_0 + (W_t - W_0)$ and, thus, $E[W_t] = E[W_0] + E[W_t - W_0] = 0$. The latter is given by the fact that $E[W_0] = 0$ (by point 1 of Definition 2.157) and $E[W_t - W_0] = 0$ (by point 3 of Definition 2.157).
2. Let $s,t \in \mathbb{R}_+$ and $Cov[W_t, W_s] = E[W_tW_s] - E[W_t]E[W_s]$, which (by point 1) gives $Cov[W_t, W_s] = E[W_tW_s]$. For simplicity, if we suppose that $s < t$, then

$$E[W_tW_s] = E[W_s(W_s + (W_t - W_s))] = E[W_s^2] + E[W_s(W_t - W_s)].$$

Since $(W_t)_{t\in\mathbb{R}_+}$ has independent increments, we obtain

$$E[W_s(W_t - W_s)] = E[W_s]E[W_t - W_s],$$

and by point 1 of Proposition 2.136 (or point 3 of Definition 2.157) it follows that this is equal to zero, thus

$$Cov[W_t, W_s] = E[W_s^2] = Var[W_s].$$

If we now observe that $W_s = W_0 + (W_s - W_0)$ and hence $Var[W_s] = Var[W_0 + (W_s - W_0)]$, then, by the independence of the increments of the process, we get

$$Var[W_0 + (W_s - W_0)] = Var[W_0] + Var[W_s - W_0].$$

Therefore, by points 1 and 3 of Definition 2.157 it follows that

$$Var[W_s] = s = \inf\{s, t\},$$

which completes the proof.

□

Proposition 2.160. *The Wiener process is a Gaussian process.*

Proof. In fact, if $n \in \mathbb{N}^*, (t_1, \ldots, t_n) \in \mathbb{R}_+^n$ with $0 = t_0 < t_1 < \ldots < t_n$ and $(a_1, \ldots, a_n) \in \mathbb{R}^n, (b_1, \ldots, b_n) \in \mathbb{R}^n$, such that $a_i \leq b_i, i = 1, 2, \ldots, n$, then it can be shown that

$$\begin{aligned} &P(a_1 \leq W_{t_1} \leq b_1, \ldots, a_n \leq W_{t_n} \leq b_n) \\ &= \int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} g(0|x_1, t_1) g(x_1|x_2, t_2 - t_1) \cdots g(x_{n-1}|x_n, t_n - t_{n-1}) dx_n \cdots dx_1, \end{aligned} \tag{2.25}$$

where

$$g(x|y, t) = \frac{e^{-\frac{|x-y|^2}{2t}}}{\sqrt{2\pi t}}.$$

In order to prove that the density of $(W_{t_1}, \ldots, W_{t_n})$ is given by the integrand of (2.25), by the uniqueness of the characteristic function, it is sufficient to show that the characteristic function ϕ' of the n-dimensional real-valued random vector, whose density is given by the integrand of (2.25), is identical to the characteristic function ϕ of $(W_{t_1}, \ldots, W_{t_n})$. Thus, let $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^n$. Then

$$\begin{aligned} \phi(\boldsymbol{\lambda}) &= E\left[e^{i(\lambda_1 W_{t_1} + \cdots + \lambda_n W_{t_n})}\right] \\ &= E\left[e^{i(\lambda_n (W_{t_n} - W_{t_{n-1}}) + (\lambda_n + \lambda_{n-1})(W_{t_{n-1}} - W_{t_{n-2}}) + \cdots + (\lambda_1 + \cdots + \lambda_n) W_{t_1})}\right] \\ &= E\left[e^{i\lambda_n (W_{t_n} - W_{t_{n-1}})}\right] E\left[e^{i(\lambda_n + \lambda_{n-1})(W_{t_{n-1}} - W_{t_{n-2}})}\right] \cdots \\ &\quad \cdots E\left[e^{i(\lambda_1 + \cdots + \lambda_n) W_{t_1}}\right], \end{aligned}$$

where we exploit the independence of the random variables $W_{t_i} - W_{t_{i-1}}$, $i = 1, \ldots, n$. Furthermore, because $(W_{t_i} - W_{t_{i-1}})$ is $N(0, t_i - t_{i-1})$, $i = 1, \ldots, n$, we get

$$\phi(\boldsymbol{\lambda}) = e^{\frac{-\lambda_n^2}{2}(t_n - t_{n-1})} e^{\frac{-(\lambda_n+\lambda_{n-1})^2}{2}(t_{n-1}-t_{n-2})} \cdots e^{\frac{-(\lambda_1+\cdots+\lambda_n)^2}{2}t_1}.$$

We continue by calculating the characteristic function ϕ':

$$\begin{aligned}\phi'(\boldsymbol{\lambda}) &= \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} e^{i\boldsymbol{\lambda}\cdot\mathbf{x}} g(0|x_1, t_1) \cdots g(x_{n-1}|x_n, t_n - t_{n-1}) dx_n \cdots dx_1 \\ &= \int_{-\infty}^{+\infty} \cdots \left(\int_{-\infty}^{+\infty} e^{i\lambda_n x_n} g(x_{n-1}|x_n, t_n - t_{n-1}) dx_n \right) \cdots dx_1.\end{aligned}$$

Because

$$\int_{-\infty}^{+\infty} e^{i\lambda x} \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-|x-m|^2}{2\sigma^2}} dx = e^{im\lambda - \frac{\lambda^2\sigma^2}{2}}, \tag{2.26}$$

we obtain

$$\begin{aligned}\phi'(\boldsymbol{\lambda}) &= \int_{-\infty}^{+\infty} \cdots \left(e^{i\lambda_n x_{n-1} - \frac{\lambda_n^2}{2}(t_n - t_{n-1})} \right) \cdots dx_1 \\ &= e^{-\frac{\lambda_n^2}{2}(t_n - t_{n-1})} \int_{-\infty}^{+\infty} \cdots \\ &\quad \left(\int_{-\infty}^{+\infty} e^{i(\lambda_n+\lambda_{n-1})x_{n-1}} g(x_{n-2}|x_{n-1}, t_{n-1} - t_{n-2}) dx_{n-1} \right) \cdots dx_1.\end{aligned}$$

Recalling (2.26) and applying it to each variable, we obtain

$$\phi'(\boldsymbol{\lambda}) = e^{\frac{-\lambda_n^2}{2}(t_n - t_{n-1})} e^{\frac{-(\lambda_n+\lambda_{n-1})^2}{2}(t_{n-1}-t_{n-2})} \cdots e^{\frac{-(\lambda_1+\ldots+\lambda_n)^2}{2}t_1},$$

and hence $\phi'(\boldsymbol{\lambda}) = \phi(\boldsymbol{\lambda})$. We now show that $g(0|x_1, t_1) \cdots g(x_{n-1}|x_n, t_n - t_{n-1})$ is of the form

$$\frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{\det K}} e^{-\frac{1}{2}(\mathbf{x}-\mathbf{m})' K^{-1}(\mathbf{x}-\mathbf{m})}.$$

We will only show it for the case where $n = 2$; then

$$\begin{aligned}g(0|x_1, t_1) g(x_1|x_2, t_2 - t_1) &= \frac{1}{2\pi\sqrt{t_1(t_2 - t_1)}} e^{-\frac{1}{2}\left[\frac{x_1^2}{t_1} + \frac{(x_2 - x_1)^2}{t_2 - t_1}\right]} \\ &= \frac{1}{2\pi\sqrt{t_1(t_2 - t_1)}} e^{-\frac{1}{2}\left[\frac{x_1^2(t_2-t_1) + (x_2-x_1)^2 t_1}{t_1(t_2-t_1)}\right]}.\end{aligned}$$

If we put

$$K = \begin{pmatrix} t_1 & t_1 \\ t_1 & t_2 \end{pmatrix} \qquad (\text{where } K_{ij} = Cov[W_{t_i}, W_{t_j}]; i,j = 1,2),$$

then

$$K^{-1} = \begin{pmatrix} \frac{t_2}{t_1(t_2-t_1)} & -\frac{1}{t_2-t_1} \\ -\frac{1}{t_2-t_1} & \frac{1}{t_2-t_1} \end{pmatrix},$$

resulting in

$$g(0|x_1,t_1)g(x_1|x_2,t_2-t_1) = \frac{1}{2\pi\sqrt{\det K}} e^{-\frac{1}{2}(\mathbf{x}-\mathbf{m})' K^{-1}(\mathbf{x}-\mathbf{m})},$$

where $m_1 = E[W_{t_1}] = 0, m_2 = E[W_{t_2}] = 0$. Thus

$$g(0|x_1,t_1)g(x_1|x_2,t_2-t_1) = \frac{1}{2\pi\sqrt{\det K}} e^{-\frac{1}{2}\mathbf{x}' K^{-1}\mathbf{x}},$$

completing the proof. □

Remark 2.161. By point 1 of Definition 2.157, it follows, for all $t \in \mathbb{R}_+$, that $W_t = W_t - W_0$ almost surely and, by point 3 of the same definition, that W_t is distributed as $N(0,t)$. Thus

$$P(a \leq W_t \leq b) = \frac{1}{\sqrt{2\pi t}} \int_a^b e^{-\frac{x^2}{2t}} dx, \qquad a \leq b.$$

Proposition 2.162. *If $(W_t)_{t\in\mathbb{R}_+}$ is a Wiener process, then it is a martingale.*

Proof. The proposition follows from Example 2.63 because $(W_t)_{t\in\mathbb{R}_+}$ is a centered process with independent increments. □

Theorem 2.163 (Kolmogorov's continuity theorem). *Let $(X_t)_{t\in\mathbb{R}_+}$ be a separable real-valued stochastic process. If there exist positive real numbers r, c, ϵ, δ such that*

$$\forall h < \delta, \forall t \in \mathbb{R}_+, \qquad E[|X_{t+h} - X_t|^r] \leq ch^{1+\epsilon}, \tag{2.27}$$

then, for almost every $\omega \in \Omega$, the trajectories are continuous in $\mathbb{R}_+$.

Proof. For simplicity, we will only consider the interval $I =]0,1[$, instead of $\mathbb{R}_+$, so that $(X_t)_{t\in]0,1[}$. Let $t \in]0,1[$ and $0 < h < \delta$ such that $t+h \in]0,1[$.

Then by the Markov inequality and by (2.27) we obtain

$$P(|X_{t+h} - X_t| > h^k) \leq h^{-rk} E[|X_{t+h} - X_t|^r] \leq ch^{1+\epsilon-rk} \tag{2.28}$$

for $k > 0$ and $\epsilon - rk > 0$. Therefore,

$$\lim_{h\to 0} P\left(|X_{t+h} - X_t| > h^k\right) = 0;$$

namely, the process is continuous in probability and, by hypothesis, separable. Under these two conditions, it can be shown that any arbitrary countable dense subset T_0 of $]0,1[$ can be regarded as a separating set. Thus we define

$$T_0 = \left\{ \frac{j}{2^n} \middle| j = 1, \ldots, 2^n - 1; n \in \mathbb{N}^* \right\}$$

and observe that, by (2.28),

$$P\left(\max_{1\le j\le 2^n-2} \left|X_{\frac{j+1}{2^n}} - X_{\frac{j}{2^n}}\right| \ge \frac{1}{2^{nk}}\right) \le \sum_{j=1}^{2^n-2} P\left(\left|X_{\frac{j+1}{2^n}} - X_{\frac{j}{2^n}}\right| \ge \frac{1}{2^{nk}}\right)$$
$$\le 2^n c 2^{-n(1+\epsilon-rk)} = c2^{-n(\epsilon-rk)}.$$

Because $(\epsilon - rk) > 0$ and $\sum_n 2^{-n(\epsilon-rk)} < \infty$, we can apply the Borel–Cantelli Lemma 1.161 to the sets

$$F_n = \left\{\max_{0\le j\le 2^n-1} \left|X_{\frac{j+1}{2^n}} - X_{\frac{j}{2^n}}\right| \ge \frac{1}{2^{nk}}\right\},$$

yielding $P(B) = 0$, where $B = \limsup_n F_n = \bigcap_n \bigcup_{k\le n} F_k$. As a consequence, if $\omega \notin B$, then $\omega \in \Omega \setminus (\bigcap_n \bigcup_{k\ge n} F_k)$, i.e., there exists an $N = N(\omega) \in \mathbb{N}^*$ such that, for all $n \ge N$,

$$\left|X_{\frac{j+1}{2^n}}(\omega) - X_{\frac{j}{2^n}}(\omega)\right| < \frac{1}{2^{nk}}, \qquad j = 0, \ldots, 2^n - 1. \tag{2.29}$$

Now, let $\omega \notin B$ and s be a rational number such that

$$s = j2^{-n} + a_1 2^{-(n+1)} + \cdots + a_m 2^{-(n+m)}, \qquad s \in [j2^{-n}, (j+1)2^{-n}[,$$

where either $a_j = 0$ or $a_j = 1$ and $m \in \mathbb{N}^*$. If we put

$$b_r = j2^{-n} + a_1 2^{-(n+1)} + \cdots + a_r 2^{-(n+r)},$$

with $b_0 = j2^{-n}$ and $b_m = s$ for $r = 0, \ldots, m$, then

$$|X_s(\omega) - X_{j2^{-n}}(\omega)| \le \sum_{r=0}^{m-1} |X_{b_{r+1}}(\omega) - X_{b_r}(\omega)|.$$

If $a_{r+1} = 0$, then $[b_r, b_{r+1}[= \emptyset$; if $a_{r+1} = 1$, then $[b_r, b_{r+1}[$ is of the form $[l2^{-(n+r+1)}, (l+1)2^{-(n+r+1)}[$. Hence from (2.29) it follows that

$$|X_s(\omega) - X_{j2^{-n}}(\omega)| \le \sum_{r=0}^{m-1} 2^{-(n+r+1)k} \le 2^{-nk} \sum_{r=0}^{\infty} 2^{-(r+1)k} \le M2^{-nk}, \tag{2.30}$$

with $M \geq 1$. Fixing $\epsilon > 0$, there exists an $N_1 > 0$ such that, for all $n \geq N_1$, $M2^{-nk} < \frac{\epsilon}{3}$, and from the fact that $M \geq 1$ it also follows that, for all $n \geq N_1$, $2^{-nk} < \frac{\epsilon}{3}$. Let t_1, t_2 be elements of T_0 (separating set) such that $|t_1 - t_2| < \min\left\{2^{-N_1}, 2^{-N(\omega)}\right\}$. If $n = \max\{N_1, N(\omega)\}$, then there is at most one rational number of the form $\frac{j+1}{2^n}$ $(j = 1, \ldots, 2^n - 1)$ between t_1 and t_2. Therefore, by (2.29) and (2.30), it follows that

$$\begin{aligned} &|X_{t_1}(\omega) - X_{t_2}(\omega)| \\ &\leq \left|X_{t_1}(\omega) - X_{\frac{j}{2^n}}(\omega)\right| + \left|X_{\frac{j+1}{2^n}}(\omega) - X_{\frac{j}{2^n}}(\omega)\right| + \left|X_{t_2}(\omega) - X_{\frac{j+1}{2^n}}(\omega)\right| \\ &< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon. \end{aligned}$$

Hence the trajectory is uniformly continuous almost everywhere in T_0 and has a continuous extension in $[0, 1]$. By Theorem 2.27, the extension coincides with the original trajectory. Therefore, the trajectory is continuous almost everywhere in $]0, 1[$. □

Theorem 2.164. *If $(W_t)_{t \in \mathbb{R}_+}$ is a real-valued Wiener process, then it has continuous trajectories almost surely.*

Proof. Let $t \in \mathbb{R}_+$ and $h > 0$. Because $W_{t+h} - W_t$ is normally distributed as $N(0, h)$, if we put $Z_{t,h} = \frac{W_{t+h} - W_t}{\sqrt{h}}$, then $Z_{t,h}$ has a standard normal distribution. Therefore, it is clear that there exists an $r > 2$ such that $E[|Z_{t,h}|^r] > 0$, and thus $E[|W_{t+h} - W_t|^r] = E[|Z_{t,h}|^r] h^{\frac{r}{2}}$. If we write $r = 2(1+\epsilon)$, then we obtain $E[|W_{t+h} - W_t|^r] = ch^{1+\epsilon}$, with $c = E[|Z_{t,h}|^r]$. The assertion then follows by Kolmogorov's continuity theorem. □

Remark 2.165. Since Brownian motion is continuous in probability, then by Theorem 2.34, it admits a separable and progressively measurable modification.

Theorem 2.166. *Every Wiener process $(W_t)_{t \in \mathbb{R}_+}$ is a Markov diffusion process. Its transition density is*

$$g(x, y; t) = \frac{e^{-\frac{(x-y)^2}{2t}}}{\sqrt{2\pi t}}, \quad \text{for} \quad x, y \in \mathbb{R},\ t \in \mathbb{R}_+^*.$$

Its infinitesimal generator is

$$\mathcal{A} = \frac{1}{2}\frac{\partial^2}{\partial x^2}, \quad \text{with domain} \quad \mathcal{D}_{\mathcal{A}} = C^2(\mathbb{R}).$$

Proof. The theorem follows directly by Theorem 2.108. See also Lamperti (1977, p. 170) and Revuz-Yor (1991, p. 264) □

Theorem 2.167 (Lévy characterization of Brownian motion). *Let $(X_t)_{t \in \mathbb{R}_+}$ be a real-valued continuous random process on a probability space $(\Omega, \mathcal{F}, P)$. Then the following two statements are equivalent:*

1. *$(X_t)_{t\in\mathbb{R}_+}$ is a P-Brownian motion.*
2. *$(X_t)_{t\in\mathbb{R}_+}$ and $(X_t^2 - t)_{t\in\mathbb{R}_+}$ are P-martingales (with respect to their respective natural filtrations).*

Proof. (For example, Ikeda and Watanabe 1989). Here we shall only prove that statement 1 implies statement 2.

The Wiener process $(W_t)_{t\in\mathbb{R}_+}$ is a continuous square-integrable martingale, with $W_t - W_s \sim N(0, t-s)$, for all $0 \leq s < t$. To show that $W_t^2 - t$ is also a martingale, we need to show that either

$$E[W_t^2 - t|\mathcal{F}_s] = W_s^2 - s \qquad \forall 0 \leq s < t$$

or, equivalently, that

$$E[W_t^2 - W_s^2|\mathcal{F}_s] = t - s \qquad \forall 0 \leq s < t.$$

In fact,

$$E[W_t^2 - W_s^2|\mathcal{F}_s] = E\left[(W_t - W_s)^2\Big|\,\mathcal{F}_s\right] = Var[W_t - W_s] = t - s.$$

Because of uniqueness, we can say that $\langle W_t\rangle = t$ for all $t \geq 0$ by indistinguishability. □

Additional characterizations are offered by the following proposition.

Proposition 2.168. *Let $(X_t)_{t\in\mathbb{R}_+}$ be a real-valued continuous process starting at 0 at time 0, and let $\mathcal{F}^X$ denote its natural filtration. It is a Wiener process if and only if either of the following statements applies:*

(i) For any real number λ, the process $\left(\exp\left\{\lambda X_t - \frac{\lambda^2}{2}t\right\}\right)_{t\in\mathbb{R}_+}$ is an $\mathcal{F}^X$ -local martingale.

(ii) For any real number λ, the process $\left(\exp\left\{i\lambda X_t + \frac{\lambda^2}{2}t\right\}\right)_{t\in\mathbb{R}_+}$ is an $\mathcal{F}^X$ -local martingale.

Proof. See, e.g., Revuz-Yor (1991). □

We may state the converse of Proposition 2.160 as follows.

Proposition 2.169. *Let $(X_t)_{t\in\mathbb{R}_+}$ be a real-valued continuous process starting at 0 at time 0. If the process is a Gaussian process satisfying*

1. *$E[X_t] = 0$ for all $t \in \mathbb{R}_+$*
2. *$K(s,t) = Cov[X_t, X_s] = \min\{s,t\}, \qquad s,t \in \mathbb{R}_+$*

then it is a Wiener process.

Proof. See, e.g., Revuz-Yor (1991, p. 35). □

Lemma 2.170. *Let* $(W_t)_{t\in\mathbb{R}_+}$ *be a real-valued Wiener process. If* $a > 0$*, then*

$$P\left(\max_{0\leq s\leq t} W_s > a\right) = 2P(W_t > a).$$

Proof. We employ the *reflection principle* by defining the process $(\tilde{W}_t)_{t\in\mathbb{R}}$ as

$$\begin{cases} \tilde{W}_t = W_t & \text{if } W_s < a, \forall s < t, \\ \tilde{W}_t = 2a - W_t & \text{if } \exists s < t \text{ such that } W_s = a. \end{cases}$$

The name arises because once $W_s = a$, then $\tilde{W}_s$ becomes a reflection of W_s about the *barrier* a. It is obvious that $(\tilde{W}_t)_{t\in\mathbb{R}}$ is a Wiener process as well. Moreover, we can observe that

$$\max_{0\leq s\leq t} W_s > a$$

if and only if either $W_t > a$ or $\tilde{W}_t > a$. These two events are mutually exclusive and thus their probabilities are additive. As they are both Wiener processes, it is obvious that the two events have the same probability, and thus

$$P\left(\max_{0\leq s\leq t} W_s > a\right) = P(W_t > a) + P(\tilde{W}_t > a) = 2P(W_t > a),$$

completing the proof. For a more general case, see (B.8). □

Theorem 2.171. *If* $(W_t)_{t\in\mathbb{R}_+}$ *is a real-valued Wiener process, then*

1. $P(\sup_{t\in\mathbb{R}_+} W_t = +\infty) = 1$
2. $P(\inf_{t\in\mathbb{R}_+} W_t = -\infty) = 1$

Proof. For $a > 0$,

$$P\left(\sup_{t\in\mathbb{R}_+} W_t > a\right) \geq P\left(\sup_{0\leq s\leq t} W_s > a\right) = P\left(\max_{0\leq s\leq t} W_s > a\right),$$

where the last equality follows by continuity of trajectories. By Lemma 2.170:

$$P\left(\sup_{t\in\mathbb{R}_+} W_t > a\right) \geq 2P\left(W_t > a\right) = 2P\left(\frac{W_t}{\sqrt{t}} > \frac{a}{\sqrt{t}}\right), \text{ for } t > 0.$$

Because W_t is normally distributed as $N(0,t)$, $\frac{W_t}{\sqrt{t}}$ is standard normal and, denoting by Φ its cumulative distribution, we get

$$2P\left(\frac{W_t}{\sqrt{t}} > \frac{a}{\sqrt{t}}\right) = 2\left(1 - \Phi\left(\frac{a}{\sqrt{t}}\right)\right).$$

By $\lim_{t\to\infty} \Phi(\frac{a}{\sqrt{t}}) = \frac{1}{2}$, it follows that

$$\lim_{t\to\infty} 2P\left(\frac{W_t}{\sqrt{t}} > \frac{a}{\sqrt{t}}\right) = 1,$$

and because

$$\left\{\sup_{t\in\mathbb{R}_+} W_t = +\infty\right\} = \bigcap_{a=1}^{\infty}\left\{\sup_{t\in\mathbb{R}_+} W_t > a\right\},$$

we obtain 1.

Point 2 follows directly from point 1, by the observation that if $(W_t)_{t\in\mathbb{R}_+}$ is a real-valued Wiener process, then so is $(-W_t)_{t\in\mathbb{R}_+}$. □

Theorem 2.172. *If $(W_t)_{t\in\mathbb{R}_+}$ is a real-valued Wiener process, then*

$$\forall h > 0, \qquad P\left(\max_{0\leq s\leq h} W_s > 0\right) = P\left(\min_{0\leq s\leq h} W_s < 0\right) = 1.$$

Moreover, for almost every $\omega \in \Omega$ the process $(W_t)_{t\in\mathbb{R}_+}$ has a zero (i.e., crosses the spatial axis) in $]0,h]$ for all $h > 0$.

Proof. If $h > 0$ and $a > 0$, then it is obvious that

$$P\left(\max_{0\leq s\leq h} W_s > 0\right) \geq P\left(\max_{0\leq s\leq h} W_s > a\right).$$

Then, by Lemma 2.170,

$$P\left(\max_{0\leq s\leq h} W_s > a\right) = 2P(W_h > a) = 2P\left(\frac{W_h}{\sqrt{h}} > \frac{a}{\sqrt{h}}\right) = 2\left(1 - \Phi\left(\frac{a}{\sqrt{h}}\right)\right).$$

For $a \to 0$, $2(1 - \Phi(\frac{a}{\sqrt{h}})) \to 1$, and thus $P(\max_{0\leq s\leq h} W_s > 0) = 1$. Furthermore,

$$P\left(\min_{0\leq s\leq h} W_s < 0\right) = P\left(\max_{0\leq s\leq h} (-W_s) > 0\right) = 1.$$

Now we can observe that

$$P\left(\max_{0\leq s\leq h} W_s > 0, \forall h > 0\right) \geq P\left(\bigcap_{n=1}^{\infty}\left(\max_{0<s\leq \frac{1}{n}} W_s > 0\right)\right) = 1.$$

Hence

$$P\left(\max_{0\leq s\leq h} W_s > 0, \forall h > 0\right) = 1$$

and, analogously,

$$P\left(\min_{0\leq s\leq h} W_s < 0, \forall h > 0\right) = 1.$$

From this it can be deduced that for almost every $\omega \in \Omega$ the process $(W_t)_{t\in\mathbb{R}_+}$ becomes zero in $]0,h]$ for all $h > 0$. On the other hand, since $(W_t)_{t\in\mathbb{R}_+}$ is a time-homogeneous Markov process with independent increments, it has the same behavior in $]h,2h]$ as in $]0,h]$, and thus it has zeros in every interval. □

Theorem 2.173. *Almost every trajectory of the Wiener process $(W_t)_{t\in\mathbb{R}_+}$ is nowhere differentiable.*

Proof. Let $D = \{\omega \in \Omega | W_t(\omega)$be differentiable for at least one $t \in \mathbb{R}_+\}$. We will show that $D \subset G$, with $P(G) = 0$ (obviously, if P is complete, then $D \in \mathcal{F}$). Let $k > 0$ and

$$A_k = \left\{\omega \left| \limsup_{h\downarrow 0} \frac{|W_{t+h}(\omega) - W_t(\omega)|}{h} < k \text{ for at least one } t \in [0,1[\right.\right\}.$$

Then, if $\omega \in A_k$, we can choose $m \in \mathbb{N}$ sufficiently large such that $\frac{j-1}{m} \leq t < \frac{j}{m}$ for $j \in \{1,\ldots,m\}$, and for $t \leq s \leq \frac{j+3}{m}$, $W(s,\omega)$ is enveloped by a cone with slope k. Then, for an integer $j \in \{1,\ldots,m\}$, we get

$$\begin{aligned}\left|W_{\frac{j+1}{m}}(\omega) - W_{\frac{j}{m}}(\omega)\right| &\leq \left|W_{\frac{j+1}{m}}(\omega) - W_t(\omega)\right| + \left|-W_t(\omega) + W_{\frac{j}{m}}(\omega)\right| \\ &< \left(\frac{j+1}{m} - \frac{j-1}{m}\right)k + \left(\frac{j}{m} - \frac{j-1}{m}\right)k \\ &= \frac{2k}{m} + \frac{k}{m} = \frac{3k}{m}. \end{aligned} \tag{2.31}$$

Analogously, we obtain that

$$\left|W_{\frac{j+2}{m}}(\omega) - W_{\frac{j+1}{m}}(\omega)\right| \leq \frac{5k}{m} \tag{2.32}$$

and

$$\left|W_{\frac{j+3}{m}}(\omega) - W_{\frac{j+2}{m}}(\omega)\right| \leq \frac{7k}{m}. \tag{2.33}$$

Because $\frac{W_{t+h}-W_t}{\sqrt{h}}$ is distributed as $N(0,1)$, it follows that

$$\begin{aligned} P(|W_{t+h} - W_t| < a) &= P\left(\frac{|W_{t+h} - W_t|}{\sqrt{h}} < \frac{a}{\sqrt{h}}\right) \\ &= \int_{-\frac{a}{\sqrt{h}}}^{\frac{a}{\sqrt{h}}} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2}\right\} dx \\ &\leq \frac{1}{\sqrt{2\pi}} 2\frac{a}{\sqrt{h}} = \frac{2a}{\sqrt{2\pi h}}. \end{aligned}$$

Putting $A_{m,j} = \{\omega | (2.31),(2.32),(2.33) \text{ are true}\}$, because the process has independent increments, we obtain

$$P(A_{m,j}) = P(\{\omega|(2.31) \text{ is true}\})P(\{\omega|(2.32) \text{ is true}\})P(\{\omega|(2.33) \text{ is true}\})$$
$$\leq 8\left(\frac{2\pi}{m}\right)^{-\frac{3}{2}} \frac{3k}{m}\frac{5k}{m}\frac{7k}{m},$$

and thus $P(A_{m,j}) \leq cm^{-\frac{3}{2}}$, $j = 1, \ldots, m$. Putting $A_m = \bigcup_{j=1}^m A_{m,j}$, then

$$P(A_m) \leq \sum_{j=1}^{m} P(A_{m,j}) \leq cm^{-\frac{1}{2}}.$$

Now let $m = n^4$ $(n \in \mathbb{N}^*)$; we obtain $P(A_{n^4}) \leq cn^{-2} = \frac{c}{n^2}$ and thus

$$\sum_n P(A_{n^4}) \leq c \sum_n \frac{1}{n^2} < \infty.$$

Therefore, by the Borel–Cantelli Lemma 1.161,

$$P\left(\limsup_n A_{n^4}\right) = 0.$$

It can now be shown that

$$A_k \subset \liminf_m A_m \equiv \bigcup_m \bigcap_{i \geq m} A_i \subset \liminf_n A_{n^4} \subset \limsup_n A_{n^4},$$

hence $A_k \subset A''_{n^4}$ and $P(A''_{n^4}) = 0$. Let

$$D_0 = \{\omega | W(\cdot, \omega) \text{ is differentiable in at least one } t \in [0,1[\}.$$

Then $D_0 \subset \bigcup_{k=1}^{\infty} A_k = G_0$, which means that D_0 is contained in a set of probability zero, namely, $D_0 \subset G_0$ and $P(G_0) = 0$. Decomposing $\mathbb{R}_+ = \bigcup_n [n, n+1[$, since the motion is Brownian and of independent increments,

$$D_n = \{\omega | W(\cdot, \omega) \text{ is differentiable in at least one } t \in [n, n+1[\},$$

analogously to D_0, will be contained in a set of probability zero, i.e., $D_n \subset G_n$ and $P(G_n) = 0$. But $D \subset \bigcup_n D_n \subset \bigcup_n G_n$, thus completing the proof. □

A trivial consequence of the preceding theorem is the following corollary.

Corollary 2.174. *Almost every trajectory of a Wiener process* $(W_t)_{t \in \mathbb{R}_+}$ *is of unbounded variation on any finite interval.*

An important property of the trajectories of a Brownian motion is their Hölder continuity. We may recall that a real function f defined on a real line satisfies a Hölder condition, or is Hölder continuous, when there are nonnegative real constants C, α such that

$$|f(y) - f(x)| \leq C|y - x|^\alpha$$

for all x and y in the domain of f. The number α is called the exponent or order of the Hölder condition.

The following theorem holds.

Theorem 2.175. *Almost every trajectory of a Wiener process* $(W_t)_{t\in\mathbb{R}_+}$ *is Hölder continuous for any order* $\alpha < \frac{1}{2}$.

Almost every trajectory of a Wiener process $(W_t)_{t\in\mathbb{R}_+}$ *is not Hölder continuous for any order* $\alpha \geq \frac{1}{2}$.

Proof. See, e.g., http://math.nyu.edu/faculty/varadhan/processes.html. □

Proposition 2.176. *Let* $(W_t)_{t\in\mathbb{R}_+}$ *be a Wiener process; then the following properties hold:*

(i) (Symmetry) The process $(-W_t)_{t\in\mathbb{R}_+}$ *is a Wiener process.*

(ii) (Time scaling) The time-scaled process $(\tilde{W}_t)_{t\in\mathbb{R}_+}$ *defined by*

$$\tilde{W}_t = tW_{1/t},\ t > 0, \quad \tilde{W}_0 = 0$$

is also a Wiener process.

(iii) (Space scaling) For any $c > 0$, *the space-scaled process* $(\tilde{W}_t)_{t\in\mathbb{R}_+}$ *defined by*

$$\tilde{W}_t = cW_{t/c^2},\ t > 0, \quad \tilde{W}_0 = 0,$$

is also a Wiener process.

Proof. See, e.g., Karlin and Taylor (1975). □

Proposition 2.177. *If* $(W_t)_{t\in\mathbb{R}_+}$ *is a Wiener process, then the process*

$$X_t = W_t - tW_1, \quad t \in [0, 1]$$

is a Brownian bridge.

Proof. See, e.g., Revuz-Yor (1991, p. 35). □

We may observe that $X_0 = X_1 = 0$, from which the name follows.

Proposition 2.178 (Strong law of large numbers). *Let* $(W_t)_{t\in\mathbb{R}_+}$ *be a Wiener process. Then*

$$\frac{W_t}{t} \to 0, \qquad as \qquad t \to +\infty, \qquad a.s.$$

Proof. See, e.g., Karlin and Taylor (1975). □

Proposition 2.179 (Law of iterated logarithms). *Let* $(W_t)_{t\in\mathbb{R}_+}$ *be a Wiener process. Then*

$$\limsup_{t\to+\infty} \frac{W_t}{\sqrt{2t\ln\ln t}} = 1, \qquad a.s.,$$
$$\liminf_{t\to+\infty} \frac{W_t}{\sqrt{2t\ln\ln t}} = -1, \qquad a.s.$$

As a consequence, for any $\epsilon > 0$ *there exists a* $t_0 > 0$ *such that for any* $t > t_0$ *we have*

$$-(1+\epsilon)\sqrt{2t\ln\ln t} \leq W_t \leq (1+\epsilon)\sqrt{2t\ln\ln t}, \qquad a.s.$$

Moreover,

$$P(W_t \geq (1+\epsilon)\sqrt{2t\ln\ln t},\ \text{i.o.}) = 0;$$

while

$$P(W_t \geq (1-\epsilon)\sqrt{2t\ln\ln t},\ \text{i.o.}) = 1.$$

Proof. See, e.g., Breiman (1968, p. 266). □

Proposition 2.180. *For almost every* $\omega \in \Omega$ *the trajectory* $(W_t(\omega))_{t\in\mathbb{R}_+}$ *of a Wiener process is locally Hölder continuous with exponent* δ *if* $\delta \in (0, \frac{1}{2})$. *But for almost every* $\omega \in \Omega$ *it is nowhere Hölder continuous with exponent* δ *if* $\delta > \frac{1}{2}$.

Wiener Process Started at x

Let $(W_t)_{t\in\mathbb{R}_+}$ be a Wiener process. For any $x \in \mathbb{R}$ the process $(W_t^x)_{t\in\mathbb{R}_+}$, defined by

$$W_t^x := x + W_t \quad t \in \mathbb{R}_+,$$

is called a *Wiener process started at* x. It is such that for any $t \in \mathbb{R}_+$ and any $B \in \mathcal{B}_\mathbb{R}$

$$P(W_t^x \in B) = \frac{1}{\sqrt{2\pi t}} \int_B e^{-\frac{(y-x)^2}{2t}} dy.$$

Reflected Brownian Motion

If $(W_t)_{t\in\mathbb{R}_+}$ is a Wiener process, the process $(|W_t|)_{t\in\mathbb{R}_+}$ is valued in $\mathbb{R}_+$; its transition density is

$$g(x,y;t) = \frac{1}{\sqrt{2\pi t}}\left[\exp\left\{-\frac{(y-x)^2}{2t}\right\} + \exp\left\{-\frac{(y+x)^2}{2t}\right\}\right],$$

for $x, y \in \mathbb{R}_+,\ t \in \mathbb{R}_+^*$.

It is known as *reflected Brownian motion*.

Its infinitesimal generator is

$$\mathcal{A} = \frac{1}{2}\frac{\partial^2}{\partial x^2}, \quad \text{with domain} \quad \mathcal{D}_\mathcal{A} = \left\{ f \in C^2(\mathbb{R}_+) \,|\, f'(0) = 0 \right\}$$

(Lamperti 1977, pp. 126 and 173).

Absorbed Brownian Motion

Let $(W_t)_{t\in\mathbb{R}_+}$ be a Wiener process; for a given $a \in \mathbb{R}$ let τ_a denote the first passage time of the process started at $W_0 = 0$. The stopped process $(X_t)_{t\in\mathbb{R}_+}$ defined by

$$\begin{aligned} X_t &= W_t, \quad \text{for} \quad 0 \le t \le \tau_a \\ X_t &= a, \qquad \text{for} \quad t \ge \tau_a, \end{aligned}$$

is called *absorbed Brownian motion.*

Its cumulative probability distribution is given by

$$P(X_t \le y) = \begin{cases} \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{y} e^{-\frac{z^2}{2t}}\, dz - \int_{2a-y}^{+\infty} e^{-\frac{z^2}{2t}}\, dz & \text{for} \quad y < a, \\ 1 & \text{for} \quad y \ge a, \end{cases}$$

for any $t \in \mathbb{R}_+$ and any $y \in \mathbb{R}$.

Its infinitesimal generator is

$$\mathcal{A} = \frac{1}{2}\frac{\partial^2}{\partial x^2},$$

with domain $\mathcal{D}_\mathcal{A} = \left\{ f \in C^2(\mathbb{R}_+) \,|\, f(x) = 0, \quad \text{for } x \ge a \right\}$ (Schuss 2010, p. 58).

Brownian Motion After a Stopping Time

Let $(W(t))_{t\in\mathbb{R}_+}$ be a Wiener process with a finite stopping time T and $\mathcal{F}_T$ the σ-algebra of events preceding T. By Remark 2.165 and Theorem 2.49, $W(T)$ is $\mathcal{F}_T$-measurable and, hence, measurable.

Remark 2.181. Brownian motion is endowed with the Feller property and therefore also with the strong Markov property (This can be shown using the representation of the semigroup associated with $(W(t))_{t\in\mathbb{R}_+}$.).

Theorem 2.182. *Resorting to the previous notation, we have that*

1. *The process* $y(t) = W(T+t) - W(T), t \geq 0$, *is again a Brownian motion.*
2. $\sigma(y(t), t \geq 0)$ *is independent of* $\mathcal{F}_T$.

(Thus a Brownian motion remains a Brownian motion after a stopping time.)

Proof. If $T = s$ (s constant), then the assertion is obvious. We now suppose that T has a countable codomain $(s_j)_{j\in\mathbb{N}}$ and that $B \in \mathcal{F}_T$. If we consider further that $0 \leq t_1 < \cdots < t_n$ and that $A_1, \ldots, A_n$ are Borel sets of $\mathbb{R}$, then

$$\begin{aligned}
&P(y(t_1) \in A_1, \ldots, y(t_n) \in A_n, B)\\
&= \sum_{j\in\mathbb{N}} P(y(t_1) \in A_1, \ldots, y(t_n) \in A_n, B, T = s_j)\\
&= \sum_{j\in\mathbb{N}} P((W(t_1 + s_j) - W(s_j)) \in A_1, \ldots\\
&\ldots, (W(t_n + s_j) - W(s_j)) \in A_n, B, T = s_j).
\end{aligned}$$

Moreover, $(T = s_j) \cap B = (B \cap (T \leq s_j)) \cap (T = s_j) \in \mathcal{F}_{s_j}$ (as observed in the proof of Theorem 2.49), and since a Wiener process has independent increments, the events $((W(t_1+s_j)-W(s_j)) \in A_1, \ldots, (W(t_n+s_j)-W(s_j)) \in A_n)$ and $(B, T = s_j)$ are independent; therefore,

$$\begin{aligned}
&P(y(t_1) \in A_1, \ldots, y(t_n) \in A_n, B)\\
&= \sum_{j\in\mathbb{N}} P((W(t_1 + s_j) - W(s_j)) \in A_1, \ldots\\
&\ldots, (W(t_n + s_j) - W(s_j)) \in A_n) P(B, T = s_j)\\
&= \sum_{j\in\mathbb{N}} P(W(t_1) \in A_1, \ldots, W(t_n) \in A_n) P(B, T = s_j)\\
&= P(W(t_1) \in A_1, \ldots, W(t_n) \in A_n) P(B),
\end{aligned}$$

where we note that $W(t_k + s_j) - W(s_j)$ has the same distribution as $W(t_k)$. From these equations (having factorized) follows point 2. Furthermore, if we take $B = \Omega$, we obtain

$$P(y(t_1) \in A_1, \ldots, y(t_n) \in A_n) = P(W(t_1) \in A_1, \ldots, W(t_n) \in A_n).$$

This shows that the finite-dimensional distributions of the process $(y(t))_{t\geq 0}$ coincide with those of W. Therefore, by the Kolmogorov–Bochner theorem, the proof of 1 is complete.

Let T be a generic finite stopping time of the Wiener process $(W_t)_{t\geq 0}$ and (as in Lemma 2.131) $(T_n)_{n\in\mathbb{N}}$ a sequence of stopping times such that $T_n \geq T, T_n \downarrow T$ as $n \to \infty$ and T_n has an at most countable codomain. We put, for all $n \in \mathbb{N}$, $y_n(t) = W(T_n + t) - W(T_n)$ and let $B \in \mathcal{F}_T, 0 \leq t_1 \leq \cdots \leq t_k$. Then, because for all $n \in \mathbb{N}$, $\mathcal{F}_T \subset \mathcal{F}_{T_n}$ (see the proof of Theorem 2.132) and for all $n \in \mathbb{N}$, the theorem holds for T_n (as already shown above), we have

$$P(y_n(t_1) \leq x_1, \ldots, y_n(t_k) \leq x_k, B) = P(W(t_1) \leq x_1, \ldots, W(t_k) \leq x_k)P(B).$$

Moreover, since W is continuous, from $T_n \downarrow T$ as $n \to \infty$, it follows that $y_n(t) \to y(t)$ a.s. for all $t \geq 0$. Thus, if $(x_1, \ldots, x_k)$ is a point of continuity of the k-dimensional distribution F_k of $(W(t_1), \ldots, W(t_k))$, we get by Lévy's continuity Theorem 1.177

$$\begin{aligned} &P(y(t_1) \leq x_1, \ldots, y(t_k) \leq x_k, B) \\ &= P(W(t_1) \leq x_1, \ldots, W(t_k) \leq x_k)P(B). \end{aligned} \tag{2.34}$$

Since F_k is continuous almost everywhere (given that Gaussian distributions are absolutely continuous with respect to the Lebesgue measure and thus have density), (2.34) holds for every $x_1, \ldots, x_k$. Therefore, for every Borel set $A_1, \ldots, A_k$ of $\mathbb{R}$, we have that

$$P(y(t_1) \in A_1, \ldots, y(t_k) \in A_k, B) = P(W(t_1) \in A_1, \ldots, W(t_n) \in A_k)P(B),$$

completing the proof. □

Definition 2.183. The real-valued process $(W_1(t), \ldots, W_n(t))'_{t \geq 0}$ is said to be an *n-dimensional Wiener process (or Brownian motion)* if

1. For all $i \in \{1, \ldots, n\}$, $(W_i(t))_{t \geq 0}$ is a Wiener process
2. The processes $(W_i(t))_{t \geq 0}$, $i = 1, \ldots, n$, are independent

(thus the σ-algebras $\sigma(W_i(t), t \geq 0)$, $i = 1, \ldots, n$, are independent).

Proposition 2.184. *If $(W_1(t), \ldots, W_n(t))'_{t \geq 0}$ is an n-dimensional Brownian motion, then it can be shown that*

1. $(W_1(0), \ldots, W_n(0)) = (0, \ldots, 0)$ *almost surely.*
2. $(W_1(t), \ldots, W_n(t))'_{t \geq 0}$ *has independent increments.*
3. $(W_1(t), \ldots, W_n(t))' - (W_1(s), \ldots, W_n(s))', 0 \leq s < t$, *has multivariate normal distribution* $N(\mathbf{0}, (t-s)I)$ *(where* $\mathbf{0}$ *is the null vector of order* n *and* I *is the* $n \times n$ *identity matrix).*

Proof. The proof follows from Definition 2.183. □

2.9 Counting, and Poisson Processes

Whereas Brownian motion and the Wiener process are continuous in space and time, there exists a family of processes that are continuous in time, but discontinuous in space, admitting jumps. The simplest of these is a counting process, of which the Poisson process is a special case. The latter also allows many explicit results. The most general process admitting both continuous and discontinuous movements is the Lévy process, which contains both Brownian motion and the Poisson process. Finally, a stable process is a particular type of Lévy process, which reproduces itself under addition.

Definition 2.185. Let $(\tau_i)_{i\in\mathbb{N}^*}$ be a strictly increasing sequence of positive random variables on the space $(\Omega, \mathcal{F}, P)$, with $\tau_0 \equiv 0$. Then the process $(N_t)_{t\in\bar{\mathbb{R}}_+}$ given by

$$N_t = \sum_{i\in\mathbb{N}^*} I_{[\tau_i,+\infty]}(t), \qquad t \in \bar{\mathbb{R}}_+,$$

valued in $\bar{\mathbb{N}}$, is called a *counting process* associated with the sequence $(\tau_i)_{i\in\mathbb{N}^*}$. Moreover, the random variable $\tau = \sup_i \tau_i$ is the *explosion time* of the process. If $\tau = \infty$ almost surely, then N_t is *nonexplosive*.

We may easily notice that, due to the following equality, which holds for any $t_1, t_2, \ldots, t_n \in \mathbb{R}_+$,

$$P(\tau_1 \leq t_1, \tau_2 \leq t_2, \ldots, \tau_n \leq t_n) = P(N(t_1) \geq 1, N(t_2) \geq 2, \ldots, N(t_n) \geq n),$$

we may claim that it is equivalent to knowledge of the probability law of $(N_t)_{t\in\mathbb{R}_+}$ and that of $(\tau_n)_{n\in\mathbb{N}^*}$.

Theorem 2.186. *Let $(\mathcal{F}_t)_{t\in\bar{\mathbb{R}}_+}$ be a filtration that satisfies the usual hypotheses (Definition 2.35). A counting process $(N_t)_{t\in\bar{\mathbb{R}}_+}$ is adapted to $(\mathcal{F}_t)_{t\in\bar{\mathbb{R}}_+}$ if and only if its associated random variables $(\tau_i)_{i\in\mathbb{N}^*}$ are stopping times.*

Proof. See, e.g., Protter (1990, p. 13). □

The following proposition holds (Protter 1990, p. 16).

Theorem 2.187. *Let $(N_t)_{t\in\mathbb{R}_+}$ be a counting process. Then its natural filtration is right-continuous.*

Hence, by a suitable extension, we may consider as underlying filtered space the given probability space $(\Omega, \mathcal{F}, P)$ endowed with the natural filtration $\mathcal{F}_t = \sigma\{N_s | s \leq t\}$.

With respect to the natural filtration, the jump times τ_n for $n \in \mathbb{N}^*$ are stopping times.

Remark 2.188. A nonexplosive counting process is RCLL. Its trajectories are right-continuous step functions with upward jumps of magnitude 1 and $N_0 = 0$ almost surely.

Proposition 2.189. *An RCLL process may admit at most jump discontinuities.*

Definition 2.190. We say that a process $(X_t)_{t\in\mathbb{R}_+}$ has a *fixed jump* at a time t if $P(X_t \neq X_{t^-}) > 0$.

Poisson Process

Definition 2.191. A counting process $(N_t)_{t\in\mathbb{R}_+}$ is a *Poisson process* if it is a process with time-homogeneous independent increments.

Theorem 2.192(Çynlar 1975, p. 71; Protter 1990, p. 13). *Let $(N_t)_{t\in\mathbb{R}_+}$ be a Poisson process. Then a $\lambda > 0$ exists such that, for any $t \in \mathbb{R}_+$, N_t has a Poisson distribution with parameter λt, i.e.,*

$$P(N_t = n) = e^{-\lambda t}\frac{(\lambda t)^n}{n!}, \qquad n \in \mathbb{N}.$$

Moreover $(N_t)_{t\in\mathbb{R}_+}$ is continuous in probability and does not have explosions.

Proposition 2.193 (Chung 1974). *Let $(N_t)_{t\in\mathbb{R}_+}$ be a Poisson process. Then*

$$P(\tau_\infty = \infty) = 1,$$

namely, almost all sample functions are step functions.

The following theorem specifies the distribution of the random variable N_t, $t \in \mathbb{R}_+$.

Theorem 2.194. *Let $(N_t)_{t\in\mathbb{R}_+}$ be a Poisson process of intensity $\lambda > 0$. Then for any $t \in \mathbb{R}_+$, $E[N_t] = \lambda t$, $Var[N_t] = \lambda t$, its characteristic function is*

$$\phi_{N_t}(u) = E\left[e^{iuN_t}\right] = e^{-\lambda t(1-\exp\{iu\})},$$

and its probability-generating function is

$$g_{N_t}(u) = E\left[u^{N_t}\right] = e^{\lambda t(u-1)}, \qquad u \in \mathbb{R}^*_+.$$

Proof. All formulas are a consequence of the Poisson distribution of N_t for any $t \in \mathbb{R}_+$:

$$\begin{aligned}
E[N_t] &= \sum_{n=0}^{\infty} n\frac{(\lambda t)^n}{n!}e^{-\lambda t} = \lambda t\sum_{n=0}^{\infty}\frac{(\lambda t)^{n-1}}{(n-1)!}e^{-\lambda t} = \lambda t,\\
E\left[N_t^2\right] &= \sum_{n=0}^{\infty} n^2\frac{(\lambda t)^n}{n!}e^{-\lambda t} = \lambda t\sum_{n=0}^{\infty}((n-1)+1)\frac{(\lambda t)^{n-1}}{(n-1)!}e^{-\lambda t}\\
&= (\lambda t)^2 + \lambda t,\\
Var[N_t] &= E\left[N_t^2\right] - (E[N_t])^2,\\
E\left[e^{iuN_t}\right] &= \sum_{n=0}^{\infty} e^{iun}\frac{(\lambda t)^n}{n!}e^{-\lambda t} = e^{-\lambda t(1-\exp\{iu\})}\sum_{n=0}^{\infty}\frac{\left(\lambda t e^{iu}\right)^n}{n!}e^{-\lambda t\exp\{iu\}}\\
&= e^{-\lambda t(1-\exp\{iu\})},\\
E\left[u^{N_t}\right] &= \sum_{n=0}^{\infty} u^n\frac{(\lambda t)^n}{n!}e^{-\lambda t} = e^{\lambda t(u-1)}\sum_{n=0}^{\infty}\frac{(u\lambda t)^n}{n!}e^{-u\lambda t} = e^{\lambda t(u-1)}.
\end{aligned}$$

□

Due to the independence of the increments, the following theorem holds.

Theorem 2.195. *A Poisson process* $(N_t)_{t\in\mathbb{R}_+}$ *is an RCLL Markov process.*

Proposition 2.196 (Rolski et al. 1999, p. 157; Billingsley 1986, p. 307). *Let* $(N_t)_{t\in\mathbb{R}_+}$ *be a counting process. From the definition,* $\tau_n = \inf\{t \in \mathbb{R}_+ : N_t \geq n\}$*; we denote by* $T_n = \tau_n - \tau_{n-1}$*, for* $n \in \mathbb{N} \setminus \{0\}$*, the interarrival times. The following statements are all equivalent:*

P^1 *:* $(N_t)_{t\in\mathbb{R}_+}$ *is a Poisson process with intensity parameter* $\lambda > 0$.

P^2 *:* T_n *are independent exponentially distributed random variables with parameter* λ.

P^3 *: For any* $t \in \mathbb{R}_+$*, and for any* $n \in \mathbb{N} - \{0\}$*, the joint conditional distribution of* $(T_1, \ldots, T_n)$*, given* $\{N_t = n\}$*, has density*

$$\frac{n!}{t^n}\mathbf{1}_{\{0<t_1<\cdots<t_n\}}$$

with respect to the Lebesgue measure, i.e., it has the same distribution of the order statistics of n independent real random variables having uniform law on $[0, t]$.

P^4 *: For any* $0 < t_1 < \cdots < t_k$ *the increments* $N_{t_2} - N_{t_1}, \ldots, N_{t_k} - N_{t_{k-1}}$ *are independent and each of them is Poisson distributed:*

$$N_{t_i} - N_{t_{i-1}} \sim P(\lambda(t_i - t_{i-1})).$$

P^5 *:* $(N_t)_{t\in\mathbb{R}_+}$ *has time-homogeneous independent increments and, as* $h \downarrow 0$,

$$P(N_h = 1) = \lambda h + o(h),$$
$$P(N_h \geq 2) = o(h);$$

moreover, $(N_t)_{t\in\mathbb{R}_+}$ *has no fixed jumps.*

Theorem 2.197. *A process* $(N_t)_{t\in\mathbb{R}_+}$ *with stationary increments has a version in which it is constant on all sample paths except for upward jumps of magnitude* 1 *if and only if there exists a parameter* $\lambda > 0$ *so that its characteristic function*

$$\phi_{N_t}(u) = E\left[e^{iuN_t}\right] = e^{-\lambda t(1-\exp\{iu\})}$$

or, equivalently, $N_t \sim P(\lambda t)$.

Proof. See, e.g., Breiman (1968). □

Remark 2.198. Let us consider a Poisson process $(N_t)_{t\in\mathbb{R}_+}$ with intensity parameter $\lambda > 0$; for any $a, b \in \mathbb{R}_+,\ a < b$ we denote

$$N((a, b]) = N_b - N_a. \tag{2.35}$$

Due to the fact that $(N_t)_{t\in\mathbb{R}_+}$ is nondecreasing and *càdlàg*, with $N(0) = 0$, by means of (2.35) it will generate a random measure on $\mathcal{B}_{\mathbb{R}_+}$ in the usual way. It is such that, for any $B \in \mathcal{B}_{\mathbb{R}_+}$,

$$N(B) = \sharp \{n \in \mathbb{N}^* | \tau_n \in B\}.$$

It is not difficult to show that $N(B)$ is a Poisson random variable with parameter $\lambda\nu^1(B)$, where ν^1 denotes the usual Lebesgue measure on $\mathcal{B}_{\mathbb{R}_+}$.

A particular consequence of this is the fact that a Poisson process $(N_t)_{t\in\mathbb{R}_+}$ cannot have a fixed jump at any $t_0 \in \mathbb{R}_+$ since

$$P(N_{t_0} - N_{t_0-}) = P(N(\{t_0\} > 0) = 1 - e^{-\lambda\nu^1(\{t_0\})} = 1 - 1 = 0.$$

Theorem 2.199. *Let $(N_t)_{t\in\mathbb{R}_+}$ be a Poisson process of intensity λ. Then $(N_t - \lambda t)_{t\in\mathbb{R}_+}$ and $((N_t - \lambda t)^2 - \lambda t)_{t\in\mathbb{R}_+}$ are martingales.*

Remark 2.200. Because $M_t = (N_t - \lambda t)^2 - \lambda t$ is a martingale, by uniqueness, the process $(\lambda t)_{t\in\mathbb{R}_+}$ is the predictable compensator of $(N_t - \lambda t)^2$, i.e., $\langle (N_t - \lambda t)^2 \rangle = \lambda t$, for all $t \in \mathbb{R}_+$, as well as the compensator of the Poisson process $(N_t)_{t\in\mathbb{R}_+}$ directly.

The following theorem, known as the *Watanabe characterization*, provides the converse of Theorem 2.199 (e.g., Bremaud 1981, p. 25).

Theorem 2.201. *Let $(N_t)_{t\in\mathbb{R}_+}$ be a counting process. Suppose that a deterministic $\lambda \in \mathbb{R}^*_+$ exists such that $(N_t - \lambda t)_{t\in\mathbb{R}_+}$ is a martingale with respect to the natural filtration of the process; then $(N_t)_{t\in\mathbb{R}_+}$ is a Poisson process.*

Corollary 2.202 (Çynlar 1975, p. 76). *Let $(N_t)_{t\in\mathbb{R}_+}$ be an integer-valued stochastic process such that its trajectory almost surely satisfies the following statements:*

1. *It is nondecreasing.*
2. *It increases by jumps only.*
3. *It is right continuous.*
4. $N_0 = 0$.

Then $(N_t)_{t\in\mathbb{R}_+}$ is a Poisson process with (deterministic) intensity $\lambda \in \mathbb{R}^*_+$ if and only if

(a) *Almost surely each jump of the process is of unit magnitude.*
(b) *For any $s, t \in \mathbb{R}_+$*

$$E[N_{t+s} - N_t | \mathcal{F}_t] = \lambda s \quad \text{a.s.}$$

Theorem 2.201 can be extended to the nonhomogeneous case.

Theorem 2.203. *Let $(N_t)_{t\in\mathbb{R}_+}$ be a counting process, and let $\lambda : \mathbb{R}_+ \to \mathbb{R}_+$ be a locally integrable function such that $(N_t - \int_0^t \lambda(s)ds)_{t\in\mathbb{R}_+}$ is a martingale with respect to the natural filtration of the process; then $(N_t)_{t\in\mathbb{R}_+}$ is a Poisson* *process, with nonhomogeneous intensity* $\lambda(t)$*, i.e., for all* $0 \le s \le t$*,* $N_t - N_s$ *is a Poisson random variable with parameter* $\int_s^t \lambda(\tau)d\tau$*, independent of* $\mathcal{F}_s$*.*

Definition 2.204. A counting process $(N_t)_{t\in\mathbb{R}_+}$ is *simple* if

$$P(N_t - N_{t^-} \in \{0,1\} \quad \text{for any} \quad t \in \mathbb{R}_+) = 1.$$

Definition 2.205. A counting process $(N_t)_{t\in\mathbb{R}_+}$ is *orderly* if

$$\lim_{h\downarrow 0} \frac{1}{h} P(N_t \ge 2) = 0.$$

Proposition 2.206. *A Poisson process with intensity* $\lambda > 0$ *is orderly and simple.*

Proof. A Poisson process has time-homogeneous increments; since it is orderly, it is also simple by Proposition 3.3.VI in Daley and Vere-Jones (1988, p. 48). Further, since λ is finite, simplicity implies orderliness by Dobrushin's lemma (e.g., Daley and Vere-Jones 1988, p. 48). □

Theorem 2.207. *Let* $(N_t)_{t\in\mathbb{R}_+}$ *be a simple counting process on* $\mathbb{R}_+$ *adapted to* $\mathcal{F}_t$*. If the* $\mathcal{F}_t$*-compensator* $(A_t)_{t\in\mathbb{R}_+}$ *of* $(N_t)_{t\in\mathbb{R}_+}$ *is a continuous and* $\mathcal{F}_0$*-measurable random process, then* $(N_t)_{t\in\mathbb{R}_+}$ *is a doubly stochastic Poisson process (with stochastic intensity), directed by* A_t*, also known as a Cox process.*

Proof. For $u \in \mathbb{R}$ let

$$M_t(u) = e^{iuN_t - (\exp\{iu\}-1)A_t}.$$

Then, using the properties of stochastic integrals, it can be shown that

$$E[M_t(u)|\mathcal{F}_0] = E\left[e^{iuN_t - (\exp\{iu\}-1)A_t} \middle| \mathcal{F}_0 \right] = 1.$$

Because A_t is assumed to be $\mathcal{F}_0$-measurable,

$$E\left[e^{iuN_t}|\mathcal{F}_0\right] = e^{(\exp\{iu\}-1)A_t},$$

representing the characteristic function of a Poisson distribution with (stochastic) intensity A_t. □

2.10 Marked Point Processes

2.10.1 Random Measures

Consider a Polish space $(E, \mathcal{B}_E)$; we denote by $\mathcal{N}$ the family of all σ-finite integer-valued measures on $(E, \mathcal{B}_E)$; we define the measurable space $(\mathcal{N}, \mathcal{B}_{\mathcal{N}})$ by assigning $\mathcal{B}_{\mathcal{N}}$ as the smallest σ-algebra on $\mathcal{N}$ with respect to which all maps

$$\mu \in \mathcal{N} \mapsto \left\{ \mu(B) \in \overline{\mathcal{N}}, B \in \mathcal{B}_E \right\}$$

are measurable.

Definition 2.208. Given a probability space $(\Omega, \mathcal{F}, P)$, a *random (point) measure* is any measurable function

$$N : (\Omega, \mathcal{F}) \to (\mathcal{N}, \mathcal{B}_{\mathcal{N}}).$$

Definition 2.209. Given a probability space $(\Omega, \mathcal{F}, P)$, a *Poisson random measure* is a random (point) measure N such that

(i) For any $B \in \mathcal{B}_E$, $N(B)$ is an integer valued random variable on $(\Omega, \mathcal{F}, P)$, admitting a Poisson distribution, i.e.,

$$P(N(B) = k) = e^{-\Lambda(B)} \frac{(\Lambda(B))^k}{k!}, \quad k \in \mathbb{N},$$

where Λ is a deterministic σ-finite measure on $\mathcal{B}_E$. An obvious consequence is that

$$\Lambda(B) = E[N(B)], \quad B \in \mathcal{B}_E;$$

(ii) For any $B_1, B_2 \in \mathcal{B}_E$, such that $B_1 \cap B_2 = \emptyset$, the random variables $N(B_1)$ and $N(B_2)$ are independent.

Theorem 2.210. *Given a deterministic σ-finite measure Λ on a Polish space $(E, \mathcal{B}_E)$, there exists a Poisson random measure N on $(E, \mathcal{B}_E)$ such that for any $B \in \mathcal{B}_E$*

$$\Lambda(B) = E[N(B)].$$

Proof. See, e.g., Ikeda and Watanabe (1989, p. 42). □

For a more detailed updated account on random measures and point processes the reader may refer to Daley and Vere-Jones (2008).

2.10.2 Stochastic Intensities

We will now generalize the notion of a compensator (Definition 2.82) to a larger class of counting processes, including the so-called marked point processes. For this we will commence with a point process on $\mathbb{R}_+$,

$$N = \sum_{n \in \mathbb{N}^*} \epsilon_{\tau_n},$$

defined by the sequence of random times $(\tau_n)_{n \in \mathbb{N}^*}$ on the underlying probability space $(\Omega, \mathcal{F}, P)$. Here ϵ_t is the Dirac measure (also called point mass) on $\mathbb{R}_+$, i.e.,

$$\forall A \in \mathcal{B}_{\mathbb{R}_+}: \qquad \epsilon_t(A) = \begin{cases} 1 & \text{if } t \in A, \\ 0 & \text{if } t \notin A. \end{cases}$$

The corresponding definition of the same process as a counting process was given in Definition 2.185.

Definition 2.211. (A^*): Let $\mathcal{F}_t = \sigma(N_s, 0 \leq s \leq t)$, $t \in \mathbb{R}_+$, be the natural filtration of the counting process $(N_t)_{t\in\mathbb{R}_+}$. We assume that

1. The filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, P)$ satisfies the usual hypotheses (Definition 2.35).
2. $E[N_t] < \infty$ for all $t \in \mathbb{R}_+$, i.e., avoiding the problem of exploding martingales in the Doob–Meyer decomposition (Theorem 2.88).

Proposition 2.212. *Under assumption (A^*) of Definition 2.211, there exists a unique increasing right-continuous predictable process $(A_t)_{t\in\mathbb{R}_+}$ such that*

1. $A_0 = 0$.
2. $P(A_t < \infty) = 1$ *for any* $t > 0$.
3. *The process $(M_t)_{t\in\mathbb{R}_+}$ defined as $M_t = N_t - A_t$ is a right-continuous zero-mean martingale.*

The process $(A_t)_{t\in\mathbb{R}_+}$ is called the compensator of the process $(N_t)_{t\in\mathbb{R}_+}$.

Proposition 2.213 (Bremaud 1981; Karr 1986). *For every nonnegative $\mathcal{F}_t$-predictable process $(C_t)_{t\in\mathbb{R}_+}$, by Proposition 2.212, we have that*

$$E\left[\int_0^\infty C_t dN_t\right] = E\left[\int_0^\infty C_t dA_t\right]. \tag{2.36}$$

Theorem 2.214. *Given a point (or counting) process $(N_t)_{t\in\mathbb{R}_+}$ satisfying assumption (A^*) of Definition 2.211 and a predictable random process $(A_t)_{t\in\mathbb{R}_+}$, the following two statements are equivalent:*

1. *$(A_t)_{t\in\mathbb{R}_+}$ is the compensator of $(N_t)_{t\in\mathbb{R}_+}$.*
2. *The process $M_t = N_t - A_t$ is a zero-mean martingale.*

Remark 2.215. In infinitesimal form, (2.36) provides the heuristic expression

$$dA_t = E[dN_t|\mathcal{F}_{t-}],$$

giving a dynamical interpretation to the compensator. In fact, the increment $dM_t = dN_t - dA_t$ is the unpredictable part of dN_t over $[0, t[$, also therefore known as the *innovation martingale* of $(N_t)_{t\in\mathbb{R}_+}$.

In the case where the innovation martingale M_t is bounded in L^2, we may apply Theorem 2.90 and introduce the predictable variation process $\langle M\rangle_t$, with $\langle M\rangle_0 = 0$ and $M_t^2 - \langle M\rangle_t$ being a uniformly integrable martingale. Then the variation process can be compensated in terms of A_t by the following theorem.

Theorem 2.216 (Karr 1986, p. 64). *Let* $(N_t)_{t\in\mathbb{R}_+}$ *be a point process on* $\mathbb{R}_+$ *with compensator* $(A_t)_{t\in\mathbb{R}_+}$*, and let the innovation process* $M_t = N_t - A_t$ *be an* L^2*-martingale. Defining* $\Delta A_t = A_t - A_{t-}$*, then*

$$\langle M\rangle_t = \int_0^t (1 - \Delta A_s) dA_s.$$

Remark 2.217. In particular, if A_t is continuous in t, then $\Delta A_t = 0$, so that $\langle M\rangle_t = A_t$. Formally, in this case we have

$$E\left[(dN_t - E[dN_t|\mathcal{F}_{t-}])^2|\mathcal{F}_{t-}\right] = dA_t = E[dN_t|\mathcal{F}_{t-}],$$

so that the counting process has locally and conditionally the typical behavior of a Poisson process.

Let N be a simple point process on $\mathbb{R}_+$ with a compensator A, satisfying the assumptions of Proposition 2.212.

Definition 2.218. We say that N admits an $\mathcal{F}_t$*-stochastic intensity* if a (nontrivial) nonnegative, predictable process $\lambda = (\lambda_t)_{t\in\mathbb{R}_+}$ exists such that

$$A_t = \int_0^t \lambda_s ds, \qquad t \in \mathbb{R}_+.$$

Remark 2.219. Due to the uniqueness of the compensator, the stochastic intensity, whenever it exists, is unique.

Formally, from

$$dA_t = E[dN_t|\mathcal{F}_{t-}]$$

it follows that

$$\lambda_t dt = E[dN_t|\mathcal{F}_{t-}],$$

i.e.,

$$\lambda_t = \lim_{\Delta t\to 0+} \frac{1}{\Delta t} E[\Delta N_t|\mathcal{F}_{t-}],$$

and, because of the simplicity of the process, we also have

$$\lambda_t = \lim_{\Delta t\to 0+} \frac{1}{\Delta t} P(\Delta N_t = 1|\mathcal{F}_{t-}),$$

meaning that $\lambda_t dt$ is the conditional probability of a new *event* during $[t, t+dt]$, given the history of the process over $[0, t]$.

Example 2.220. (Poisson process). A stochastic intensity does exist for a Poisson process with intensity $(\lambda_t)_{t\in\mathbb{R}_+}$ and, in fact, is identically equal to the latter (hence deterministic).

A direct consequence of Theorem 2.216 and of the previous definitions is the following theorem.

Theorem 2.221 (Karr 1986, p. 64). *Let $(N_t)_{t\in\mathbb{R}_+}$ be a point process satisfying assumption (A^*) of Definition 2.211 and admitting stochastic intensity $(\lambda_t)_{t\in\mathbb{R}_+}$. Assume further that the innovation martingale*

$$M_t = N_t - \int_0^t \lambda_s ds, \qquad t \in \mathbb{R}_+$$

is an L^2-martingale. Then for any $t \in \mathbb{R}_+$:

$$\langle M \rangle_t = \int_0^t \lambda_s ds.$$

An important theorem that further explains the role of the stochastic intensity for counting processes is as follows (Karr 1986, p. 71).

Theorem 2.222. *Let $(\Omega, \mathcal{F}, P)$ be a probability space over which a simple point process with an $\mathcal{F}_t$-stochastic intensity $(\lambda_t)_{t\in\mathbb{R}_+}$ is defined. Suppose that P_0 is another probability measure on $(\Omega, \mathcal{F})$ with respect to which $(N_t)_{t\in\mathbb{R}_+}$ is a stationary Poisson process with rate 1. Then $P << P_0$, and for any $t \in \mathbb{R}_+$ we have*

$$\frac{dP}{dP_0}|_{\mathcal{F}_t} = \exp\left\{\int_0^t (1-\lambda_s)ds + \int_0^t \ln \lambda_s dN_s\right\}. \tag{2.37}$$

Conversely, if P_0 is as above and P a probability measure on $(\Omega, \mathcal{F})$, absolutely continuous with respect to P_0, then there exists a predictable process λ such that N has stochastic intensity λ with respect to P [and (2.37) holds].

Marked Point Processes

We will now consider a generic Polish space endowed with its σ-algebra $(E, \mathcal{E})$ and introduce a sequence of $(E, \mathcal{E})$-valued random variables $(Z_n)_{n\in\mathbb{N}^*}$ in addition to the sequence of random times $(\tau_n)_{n\in\mathbb{N}^*}$, which are $\bar{\mathbb{R}}_+$-valued random variables.

Definition 2.223. The random measure on $\bar{\mathbb{R}}_+ \times E$,

$$N = \sum_{n\in\mathbb{N}^*} \epsilon_{(\tau_n, z_n)},$$

is called a *marked point process* with *mark space* $(E, \mathcal{E})$. z_n is called the *mark* of the event occurring at time τ_n. The process

$$N_t = N([0,t] \times E), \qquad t \in \mathbb{R}_+,$$

is called the *underlying counting process* of the process N. As usual, we assume that the process $(N_t)_{t \in \mathbb{R}_+}$ is simple.

For $B \in \mathcal{E}$ the process

$$N_t(B) := N([0,t] \times B) = \sum_{n \in \mathbb{N}^*} I_{[\tau_n \leq t, Z_n \in B]}(t), \qquad t \in \mathbb{R}_+,$$

represents the counting process of events occurring up to time t with marks in $B \in \mathcal{E}$. The *history* of the process up to time t is denoted as

$$\mathcal{F}_t := \sigma(N_s(B) | 0 \leq s \leq t, B \in \mathcal{E}).$$

We will assume throughout that the filtered space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}_+}, P)$ satisfies the usual hypotheses (Definition 2.35).

Remark 2.224. Note that, for any $n \in \mathbb{N}^*$, while τ_n is $\mathcal{F}_{\tau_n -}$-measurable, Z_n is $\mathcal{F}_{\tau_n}$-measurable but not $\mathcal{F}_{\tau_n -}$-measurable, i.e.,

$$\mathcal{F}_{\tau_n} = \sigma\left((\tau_1, Z_1), \ldots, (\tau_n, Z_n)\right),$$

whereas

$$\mathcal{F}_{\tau_n -} = \sigma\left((\tau_1, Z_1), \ldots, (\tau_{n-1}, Z_{n-1}), \tau_n\right).$$

Hence τ_n is an $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$ stopping time.

By a reasoning similar to that employed for regular conditional probabilities in Chap. 1, the following theorem can be proved, which provides an extension of Theorem 2.214 to marked point processes.

Theorem 2.225 (Bremaud 1981; Karr 1986; Last and Brandt 1995). *Let N be a marked point process such that the underlying counting process $(N_t)_{t \in \mathbb{R}_+}$ satisfies the assumptions of Proposition 2.212. Then there exists a unique random measure Λ on $\mathbb{R}_+ \times E$ such that*

1. *For any $B \in \mathcal{E}$, the process $\Lambda([0,t] \times B)$ is $\mathcal{F}_t$-predictable.*
2. *For any nonnegative $\mathcal{F}_t$-predictable process C on $\mathbb{R}_+ \times E$:*

$$E\left[\int C(t,z) N(dt \times dz)\right] = E\left[\int C(t,z) \Lambda(dt \times dz)\right].$$

The random measure ν introduced in the preceding theorem is called the $\mathcal{F}_t$-compensator of the process N. The preceding theorem again suggests that formally the following holds:

$$\Lambda(dt \times dz) = E\left[N(dt \times dz) | \mathcal{F}_{t-}\right].$$

The following propositions mimic the corresponding results for the unmarked point processes.

Proposition 2.226. *For any $B \in \mathcal{E}$, the process*

$$M_t(B) := N_t(B) - \Lambda([0,t] \times B), \qquad t \in \mathbb{R}_+,$$

is a zero-mean martingale.

We will call the process $M = (M_t(B))_{t\in\mathbb{R}_+, B\in\mathcal{E}}$ the innovation process of N. Henceforth let us denote $A_t(B) := \Lambda([0,t] \times B)$.

Proposition 2.227 (Karr 1986, p. 65). *Let N be a marked point process on $\mathbb{R}_+ \times E$, with compensator Λ, and let B_1 and B_2 be two disjoint sets in $\mathcal{E}$ for which $M_t(B_1)$ and $M_t(B_2)$ are L^2-martingales. Then*

$$\langle M_t(B_1), M_t(B_2)\rangle_t = -\int_0^t \Delta A_s(B_1)\Delta A_s(B_2)ds.$$

Hence, if $(A_t(B))_{t\in\mathbb{R}_+}$ is continuous in t for any $B \in \mathcal{E}$, then the two martingales $M_t(B_1)$ and $M_t(B_2)$ are orthogonal.

Definition 2.228. Let N be a marked point process on $\mathbb{R}_+ \times E$. We say that $(\lambda_t(B))_{t\in\mathbb{R}_+, B\in\mathcal{E}}$ is the $\mathcal{F}_t$*-stochastic intensity* of N provided that

1. For any $t \in \mathbb{R}_+$ the map

$$B \in \mathcal{E} \to \lambda_t(B) \in \mathbb{R}_+$$

is a random measure on $\mathcal{E}$.
2. For any $B \in \mathcal{E}$ the process $(\lambda_t(B))_{t\in\mathbb{R}_+}$ is the stochastic intensity of the counting process

$$N_t(B) = \sum_{n\in\mathbb{N}^*} I_{[\tau_n \le t, Z_n \in B]}(t);$$

i.e., for any $t \in \mathbb{R}_+, B \in \mathcal{E}$:

$$A_t(B) = \int_0^t \lambda_s(B)ds,$$

in which case the process $(A_t(B))_{t\in\mathbb{R}_+, B\in\mathcal{E}}$ is known as the *cumulative stochastic intensity* of N.

In the presence of the absolute continuity (hence the continuity) of the process $A_t(B)$ as a function of t, the following proposition is an obvious consequence of Proposition 2.227.

Proposition 2.229. *Let N be a marked point process on $\mathbb{R}_+$ with mark space $(E, \mathcal{E})$ and stochastic intensity $(\lambda_t(B))_{t\in\mathbb{R}_+, B\in\mathcal{E}}$. Let B_1 and B_2 be two disjoint sets in $\mathcal{E}$ such that the corresponding innovation martingales are bounded in L^2. Then $M(B_1)$ and $M(B_2)$ are orthogonal, i.e.,*

$$\langle M_t(B_1), M_t(B_2)\rangle_t = 0 \qquad \text{for any } t \in \mathbb{R}_+.$$

Representation of Point Process Martingales

Let N be a point process on $\mathbb{R}_+$ with $\mathcal{F}$-compensator A. From the section on martingales we know that if $M = N - A$ is the innovation martingale of N and H is a bounded predictable process, then

$$\tilde{M}_t = \int_0^t H(s) dM_s, \qquad t \in \mathbb{R}_+,$$

is also a martingale. In fact, the converse also holds, as stated by the following theorem, which extends an analogous result for Wiener processes to marked point processes.

Theorem 2.230 (Martingale representation). *Let N be a marked point process on $\mathbb{R}_+$ with mark space $(E, \mathcal{E})$, and let M be its innovation process with respect to the internal history $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$. Suppose the assumptions of Proposition 2.212 are satisfied, and let $(\tilde{M}_t)_{t \in \mathbb{R}_+}$ be a right-continuous and uniformly integrable $\mathcal{F}_t$-martingale. Then there exists a process $(H(t,x))_{t \in \mathbb{R}_+, x \in E}$ such that*

$$\tilde{M}_t = \tilde{M}_0 + \int_{[0,t] \times E} H(s,x) M_s(dx).$$

Proof. See Last and Brandt (1995, p. 342). □

The Marked Poisson Process

A marked Poisson process is a marked point process such that any univariate point process counting its points with a mark in a fixed Borel set is Poisson. It turns out that these processes are necessarily independent whenever the corresponding mark sets are disjoint. Consider a marked point process N on $\mathbb{R}_+ \times E$, and let Λ be a σ-finite deterministic measure on $\mathbb{R}_+ \times E$. Then, formally, we have the following definition.

Definition 2.231. N is a marked Poisson process if, for any $s, t \in \mathbb{R}_+, s < t$ and any $B \in \mathcal{E}$,

$$P(N(]s,t] \times B) = k | \mathcal{F}_s) = \frac{(\Lambda(]s,t] \times B))^k}{k!} \exp\left\{-\Lambda(]s,t] \times B)\right\},$$

for $k \in \mathbb{N}$, almost surely with respect to P.

In the preceding case the intensity measure Λ is such that

$$\Lambda(]s,t] \times B) = E[N(]s,t] \times B)]$$

for any $s, t \in \mathbb{R}_+, s < t$ and any $B \in \mathcal{E}$. It is the (deterministic) compensator of the marked Poisson process, formally:

$$\Lambda(dt \times dx) = E[N(dt \times dx)|\mathcal{F}_{t-}] = E[N(dt \times dx)],$$

thereby confirming the independence of increments for the marked Poisson process. Now the following theorem is a consequence of the definitions.

Theorem 2.232. *Let N be a marked Poisson process and $B_1, \ldots, B_m \in \mathcal{E}$ for $m \in \mathbb{N}^*$ mutually disjoint sets. Then $N(\cdot \times B_1), \ldots, N(\cdot \times B_m)$ are independent Poisson processes with intensity measures $\Lambda(\cdot \times B_1), \ldots, \Lambda(\cdot \times B_m)$, respectively.*

Proof. See Last and Brandt (1995, p. 182). □

The underlying counting process of a marked Poisson process $N(]0, t] \times E)$ is itself a univariate Poisson process with intensity measure $\bar{\Lambda}(]s, t]) = \Lambda(]s, t] \times E)$ for any $s, t \in \mathbb{R}_+, s < t$. The intensity measure may be chosen to be continuous, in which case $\bar{\Lambda}(\{t\}) = 0$, or even absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}_+$, so that

$$\bar{\Lambda}([0, t]) = \int_0^t \lambda(s) ds,$$

where $\lambda \in \mathcal{L}^1(\mathbb{R}_+)$.

Time-Homogenous Marked Poisson Process

A particular case of interest for our subsequent analysis is the following one.

Definition 2.233. A marked Poisson process N on $\mathbb{R}_+ \times E$ is time-homogenous if there exists a σ-finite deterministic measure ν on $\mathcal{E}$ such that the intensity measure Λ of N is given by

$$\Lambda(]s, t] \times B) = E[N(]s, t] \times B)] = (t - s)\, \nu(B)$$

for any $s, t \in \mathbb{R}_+, s < t$ and any $B \in \mathcal{E}$.

Formally:

$$E[N(dt \times dx)|\mathcal{F}_{t-}] = E[N(dt \times dx)] = dt\, \nu(dx).$$

We will have in particular

$$E[N([0, t] \times B)] = t\, \nu(B)$$

for any $t \in \mathbb{R}_+$ and any $B \in \mathcal{E}$ and

$$\nu(B) = E[N_1(B)] = E[N([0, 1] \times B)]$$

for any $B \in \mathcal{E}$.

Proposition 2.234. *For any $B \in \mathcal{E}$, the process*

$$M_t(B) := N_t(B) - t\nu(B), \qquad t \in \mathbb{R}_+$$

is a zero-mean martingale.

The random measure $N(dt \times dx) - dt\,\nu(dx)$ is usually called the *compensated Poisson measure.* The measure ν is called the *characteristic* of the time-homogeneous marked Poisson measure.

As a consequence of the preceding theorems we may state the following.

Proposition 2.235. *If N is a Poisson random measure with intensity measure $\Lambda(]s,t] \times B) = E[N(]s,t] \times B)] = (t-s)\,\nu(B)$, then for any $B \in \mathcal{E}$,*

- $N_t(B) = \int_0^t \int_B N(dt \times dx)$, $t \in \mathbb{R}_+$ *is a Poisson process, with intensity* $t\,\nu(B)$.
- $N_t(B)$ *is independent of* $N_t(B')$ *if* $B \cap B' = \emptyset$ *for any* $t \in \mathbb{R}_+$.

Theorem 2.236. *Given a deterministic σ-finite measure ν on a Polish space $(E, \mathcal{B}_E)$, there exists a time-homogeneous marked Poisson process N on $\mathbb{R}_+ \times E$ having characteristic measure ν.*

Proof. See, e.g., Ikeda and Watanabe (1989, p. 44). □

2.11 Lévy Processes

Definition 2.237. Let $(X_t)_{t \in \mathbb{R}_+}$ be an adapted process with $X_0 = 0$ almost surely. If X_t

1. has independent increments,
2. has stationary increments,
3. is continuous in probability so that $X_s \xrightarrow[s \to t]{P} X_t$,

then it is a *Lévy process.*

Proposition 2.238. *Both the Wiener and the Poisson processes are Lévy processes.*

Theorem 2.239. *Let $(X_t)_{t \in \mathbb{R}_+}$ be a Lévy process. Then it has an RCLL version $(Y_t)_{t \in \mathbb{R}_+}$, which is also a Lévy process.*

Proof. See, e.g., Kallenberg (1997, p. 235). □

For Lévy processes we can invoke examples of filtrations that satisfy the usual hypotheses.

Theorem 2.240. *Let $(X_t)_{t\in\mathbb{R}_+}$ be a Lévy process and $\mathcal{G}_t = \sigma(\mathcal{F}_t, \mathcal{N})$, where $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ is the natural filtration of X_t and $\mathcal{N}$ the family of P-null sets of $\mathcal{F}_t$. Then $(\mathcal{G}_t)_{t\in\mathbb{R}_+}$ is right-continuous.*

Proof. See, e.g., Protter (2004, p. 22). □

Remark 2.241. Because, by Theorem 2.239, every Lévy process has an RCLL version, by Proposition 2.189, the only type of discontinuity it may admit is jumps.

Definition 2.242. Taking the left limit $X_{t^-} = \lim_{s\to t} X_s$, $s < t$, we define

$$\Delta X_t = X_t - X_{t^-}$$

as the jump at t. If $\sup_t |\Delta X_t| \leq c$ almost surely, $c \in \mathbb{R}_+$, constant and nonrandom, then X_t is said to have *bounded jumps*.

Theorem 2.243. *Let $(X_t)_{t\in\mathbb{R}_+}$ be a Lévy process with bounded jumps. Then*

$$E\left[|X_t|^p\right] < \infty, \qquad \text{i.e., } X_t \in \mathcal{L}^p \qquad \text{for any } p \in \mathbb{N}^*.$$

Proof. See, e.g., Protter (2004, p. 25). □

Theorem 2.244. *Let $(X_t)_{t\in\mathbb{R}_+}$ be a Lévy process. Then*

(i) If $X_t \in \mathcal{L}^1$ for some $t \in \mathbb{R}_+$, then $X_t \in \mathcal{L}^1$ for any $t \in \mathbb{R}_+$, and

$$E[X_t] = tE[X_1].$$

(ii) If $X_t \in \mathcal{L}^2$ for some $t \in \mathbb{R}_+$, then $X_t \in \mathcal{L}^2$ for any $t \in \mathbb{R}_+$, and

$$Var[X_t] = tVar[X_1].$$

Proof. See, e.g., Mikosch (2009, p. 338). □

Theorem 2.245. *Let $(X_t)_{t\in\mathbb{R}_+}$ be a Lévy process. Then it has an RCLL version without fixed jumps (Proposition 2.189).*

Proof. See, e.g., Kallenberg (1997). □

We proceed with the general representation theorem of a Lévy process, commencing with the analysis of the structure of its jumps. Along the lines of the definition of counting and Poisson processes, let $\Lambda \in \mathcal{B}_{\mathbb{R}}$, such that 0 is not in $\bar{\Lambda}$, the closure of Λ. For a Lévy process $(X_t)_{t\in\mathbb{R}_+}$ we also, as before, define the random variables

$$\tau_{i+1}^{\Lambda} = \inf\left\{t > \tau_i^{\Lambda} | \Delta X_t \in \Lambda\right\}, \qquad i = 0, \ldots, n; \tau_0^{\Lambda} \equiv 0.$$

Because $(X_t)_{t\in\mathbb{R}_+}$ has RCLL paths and $0 \notin \Lambda$, it is easy to demonstrate that

$$\left\{\tau_n^{\Lambda} \geq t\right\} \in \mathcal{F}_{t+} = \mathcal{F}_t;$$

thus $(\tau_i^{\Lambda})_{i\in\mathbb{N}^*}$ are stopping times, and moreover, $\tau_i^{\Lambda} > 0$ almost surely as well as $\lim_{n\to\infty} \tau_n^{\Lambda} = +\infty$ almost surely. If we now define

$$N_t^{\Lambda} = \sum_{0<s\leq t} I_{\{\Lambda\}}(\Delta X_s) \equiv \sum_{i=1}^{\infty} I_{\left[\tau_i^{\Lambda}\leq t\right]}(t),$$

then $(N_t^{\Lambda})_{t\in\mathbb{R}_+}$ is a nonexplosive counting process, and, more specifically, we have the following theorem.

Theorem 2.246. *Let $\Lambda \in \mathcal{B}_{\mathbb{R}}$, with $0 \notin \bar{\Lambda}$. Then $(N_t^{\Lambda})_{t\in\mathbb{R}_+}$ is a time-homogeneous Poisson process with intensity*

$$\nu(\Lambda) = E\left[N_1^{\Lambda}\right].$$

Remark 2.247. If the Lévy process $(X_t)_{t\in\mathbb{R}_+}$ has bounded jumps, then $\nu(\Lambda) < +\infty$.

Theorem 2.248. *For any $t \in \mathbb{R}_+$ the mapping*

$$\Lambda \to N_t(\Lambda) \equiv N_t^{\Lambda}, \qquad \Lambda \in \mathcal{B}_{\mathbb{R}}; 0 \notin \bar{\Lambda},$$

is a random (counting) measure. Furthermore, the mapping

$$\Lambda \to \nu(\Lambda), \qquad \Lambda \in \mathcal{B}_{\mathbb{R}}; 0 \notin \bar{\Lambda},$$

is a σ-finite measure.

Proof. See, e.g., Protter (2004, p. 26). □

Definition 2.249. The measure ν given by

$$\nu(\Lambda) = E\left[\sum_{0<s\leq 1} I_{\{\Lambda\}}(\Delta X_s)\right], \qquad \Lambda \in \mathcal{B}_{\mathbb{R}\setminus\{0\}},$$

is called the *Lévy measure* of the Lévy process $(X_t)_{t\in\mathbb{R}_+}$.

It can be shown that the following proposition holds.

Proposition 2.250. *The Lévy measure ν is a measure on $\mathbb{R}\setminus\{0\}$ such that*

$$\int_{\mathbb{R}\setminus\{0\}} \min\left\{1, x^2\right\} \nu(dx) < +\infty.$$

Proof. See, e.g., Medvegyev (2007, p. 481). □

Hence $(N_t)_{t\in\mathbb{R}_+}$ is a Poisson random measure on $\mathcal{B}_{\mathbb{R}\setminus\{0\}}$, with intensity measure ν (see also Sect. 2.10).

Theorem 2.251. *Under the assumptions of Theorem 2.248, let f be a measurable function, finite on Λ. Then*

$$\int_\Lambda f(x)N_t(dx) = \sum_{0<s\leq t} f(\Delta X_s)I_{\{\Lambda\}}(\Delta X_s).$$

Because by Theorem 2.246 $(N_t^\Lambda)_{t\in\mathbb{R}_+}$ is a time-homogeneous Poisson process, we also have the following proposition.

Proposition 2.252. *Under the assumptions of Theorem 2.251, the process $(\int_\Lambda f(x)N_t(dx))_{t\in\mathbb{R}_+}$ is a Lévy process. In particular, if $f(x) = x$, then the process is nonexplosive almost surely for any $t \in \mathbb{R}_+$.*

Proof. See, e.g., Protter (2004, p. 27). □

Theorem 2.253. *Let $\Lambda \in \mathcal{B}_\mathbb{R}$, $0 \notin \bar{\Lambda}$. Then the process*

$$\left(X_t - \int_\Lambda f(x)N_t(dx)\right)_{t\in\mathbb{R}_+}$$

is a Lévy process.

Now, if we define

$$J_t = \int_{\{|x|\geq 1\}} xN_t(dx) = \sum_{0<s\leq t} \Delta X_s I_{\{|\Delta X_s|\geq 1\}}(|\Delta X_s|),$$

then because $(X_t)_{t\in\mathbb{R}_+}$ has RCLL paths for each $\omega \in \Omega$, its trajectory has only finitely many jumps bigger than 1 during the interval $[0, t]$. Therefore $(J_t)_{t\in\mathbb{R}_+}$ has paths of finite variation on compacts.

Both $(J_t)_{t\in\mathbb{R}_+}$ (by Proposition 2.252) and $V_t = X_t - J_t$ (by Theorem 2.253) are Lévy processes, where in particular the latter has jumps bounded by 1. Hence all moments of $(V_t)_{t\in\mathbb{R}_+}$ exist and are finite. Because $E[V_1] = \mu$ (and $E[V_0] = 0$), we have $E[V_t] = \mu t$, by the stationarity of the increments. If we define $Y_t = V_t - E[V_t]$, for all $t \in \mathbb{R}_+$, then $(Y_t)_{t\in\mathbb{R}_+}$ has independent increments and mean zero. Hence it is a martingale. If we further define $Z_t = J_t + \mu t$, then the following decomposition theorem holds.

Theorem 2.254. *Let $(X_t)_{t\in\mathbb{R}_+}$ be a Lévy process. Then it can be decomposed as*

$$X_t = Y_t + Z_t,$$

where Y_t and Z_t are both Lévy processes and, furthermore, Y_t is a martingale with bounded jumps and $Y_t \in L^p$, for all $p \geq 1$, whereas Z_t has trajectories of finite variation on compacts.

Proposition 2.255. *Let $(X_t)_{t\in\mathbb{R}_+}$ be a Lévy process and ν its Lévy measure. Then for any $a \in \mathbb{R}_+^*$*

$$Z_t = \int_{\{|x|<a\}} x[N_t(dx) - t\nu(dx)] \tag{2.38}$$

is a zero-mean martingale.

By Theorem 2.199, the process $(N_t - \lambda t)_{t\in\mathbb{R}_+}$ is also a zero-mean martingale and $(Z_t)_{t\in\mathbb{R}_+}$ can be interpreted as a mixture of compensated Poisson processes.

Theorem 2.256. *Let $(X_t)_{t\in\mathbb{R}_+}$ be a Lévy process with jumps bounded by $a \in \mathbb{R}_+^*$, and let*

$$V_t = X_t - E[X_t] \qquad \forall t \in \mathbb{R}_+.$$

Then $(V_t)_{t\in\mathbb{R}_+}$ is a zero-mean martingale that can be decomposed as

$$V_t = Z_t^c + Z_t \qquad \forall t \in \mathbb{R}_+,$$

where Z_t^c is a martingale with continuous paths and Z_t as defined in (2.38). In fact, $Z_t^c = W_t$ is Brownian motion.

Theorem 2.256 can be interpreted by saying that a Lévy process with bounded jumps can be decomposed as the sum of a continuous martingale (Brownian motion) and another martingale that is a mixture of compensated Poisson processes. More generally, a third component would be due to the presence of unbounded jumps.

An updated detailed account of the preceding equations can be found in Medvegyev (2007, Chap. 7); an additional and important general reference is Bertoin (1996).

Theorem 2.257 (Lévy-Itô decomposition). *Let $(X_t)_{t\in\mathbb{R}_+}$ be a Lévy process and $\mu \in \mathbb{R}$. Then*

$$\begin{aligned} X_t &= \sigma W_t + \mu t + \int_{\{|x|<1\}} x[N_t(dx) - t\nu(dx)] + \sum_{0<s\le t} \Delta X_s I_{\{|\Delta X_s|\ge 1\}} \\ &= W_t + \mu t + \int_{\{|x|<1\}} x[N_t(dx) - t\nu(dx)] + \int_{\{|x|\ge 1\}} x N_t(dx), \end{aligned}$$

where

1. *W_t is a standard Brownian motion.*
2. *For any set $\Lambda \in \mathcal{B}_{\mathbb{R}\setminus\{0\}}$, $0 \notin \bar{\Lambda}$:*
   - *$N_t^\Lambda \equiv \int_\Lambda N_t(dx)$ is a Poisson process independent of W_t.*
   - *N_t^Λ is independent of $N_t^{\Lambda'}$ if $\Lambda \cap \Lambda' = \emptyset$.*
   - *N_t^Λ has intensity $t\,\nu(\Lambda)$.*
   - *ν is a measure on $\mathbb{R}\setminus\{0\}$ such that $\int_{\mathbb{R}\setminus\{0\}} \min\{1, x^2\}\, \nu(dx) < +\infty$.*

Proof. See, e.g., Sato (1999, p. 119), and Medvegyev (2007, p. 480). □

Remark 2.258. In the preceding formula we may notice that the process

$$\int_{\{|x|\geq 1\}} xN_t(dx), \quad t \in \mathbb{R}_+,$$

describing the "large jumps" of a Levy process, is a compound Poisson process. We may further notice that while the process

$$X_t - \int_{\{|x|\geq 1\}} xN_t(dx), \quad t \in \mathbb{R}_+,$$

has finite moments of any order, the process $\int_{\{|x|\geq 1\}} xN_t(dx),\ t \in \mathbb{R}_+$, may have no finite moments (e.g., Applebaum 2004, p. 110).

Proposition 2.259.

(i) Let $X = (X_t)_{t\in\mathbb{R}_+}$ be a Lévy process. For any $t \in \mathbb{R}_+$, the distribution of X_t is infinitely divisible.
(ii) For any infinitely divisible law P one can construct a Lévy process $X = (X_t)_{t\in\mathbb{R}_+}$ such that X_1 has law P.

Proof. See Exercise 2.30. □

Proposition 2.260. *The characteristic function of a Lévy process $X = (X_t)_{t\in\mathbb{R}_+}$ at time $t \in \mathbb{R}_+$ admits the following representation:*

$$\phi_{X_t}(u) = (\phi_{X_1}(u))^t, \quad u \in \mathbb{R}.$$

Proof. See, e.g., Mikosch (2009, p. 343). □

The following result is a trivial consequence of the foregoing proposition (see also Exercise 1.13).

Theorem 2.261 (Lévy–Khintchine formula). *Under the assumptions of Theorem 2.257, the characteristic function of a Lévy process $X = (X_t)_{t\in\mathbb{R}_+}$ at time $t \in \mathbb{R}_+$ is given by*

$$\phi_{X_t}(u) = E\left[e^{iuX_t}\right] = \exp\{-t\psi(u)\}, \quad u \in \mathbb{R}$$

where ψ is the characteristic exponent of X_1,

$$\psi(u) = \frac{1}{2}\sigma^2u^2 - i\mu u + \int_{\{|x|<1\}} (1 - \exp\{iux\} + iux)\nu(dx) + \int_{\{|x|\geq 1\}} (1 - \exp\{iux\})\nu(dx),$$

having the same ingredients as in Theorem 2.257. Further, the triplet (μ, σ^2, ν) *characterizes the probability law of the Lévy process* $(X_t)_{t\in\mathbb{R}_+}$ *in a unique way.*

Proof. See, e.g., Mikosch (2009, p. 344) and references therein. □

This is why the triplet (μ, σ^2, ν) is called the *characteristic triplet* of the Lévy process $X = (X_t)_{t\in\mathbb{R}_+}$.

The following result is a natural consequence of Theorems 2.257 and 2.261.

Theorem 2.262. *Any Lévy process* $X = (X_t)_{t\in\mathbb{R}_+}$ *can be decomposed as follows:*

$$X_t = \mu t + \sigma W(t) + S(t), \quad t \in \mathbb{R}_+, \tag{2.39}$$

where $\mu \in \mathbb{R}$ *and* $\sigma \in \mathbb{R}_+$ *are constants and* $W(t)$ *is a standard Brownian motion independent of the process* S.

Notice that the component $(\mu t+\sigma W(t))_{t\in\mathbb{R}_+}$ has continuous sample paths almost surely, while $(S(t))_{t\in\mathbb{R}_+}$ has almost surely discontinuous sample paths (for any $t \in \mathbb{R}_+$, $S(t)$ can be obtained as the weak limit of a sequence of compound Poisson random variables); this is why S is called the *pure jump process* of X, and the Lévy measure ν is known as the *jump measure* of X.

Corollary 2.263. *Let* $X = (X_t)_{t\in\mathbb{R}_+}$ *be a Lévy process. If* $X_t \in \mathcal{L}^2$ *for* $t \in \mathbb{R}_+$, *then*

$$\int_{|x|\geq 1} |x|^2 \nu(dx) < +\infty,$$

and the representation (2.39) can be written more explicitly as

$$X(t) = \mu_1 t + \sigma W(t) + \int_{\mathbb{R}-\{0\}} x[N_t(dx) - t\nu(dx)], \quad t \in \mathbb{R}_+, \tag{2.40}$$

with $\mu_1 = \mu + \int_{|x|\geq 1} x\nu(dx)$.

Proof. See, e.g., Di Nunno et al. (2009, p. 162). □

Using the notations of marked point processes, the Poisson random measure $N = (N_t)_{t\in\mathbb{R}_+}$ on $\mathbb{R} \setminus \{0\}$ can be seen as

$$N_t(B) = \int_0^t \int_B N(dt, dx) \qquad B \in \mathcal{B}_{\mathbb{R}},\ 0 \notin \bar{B},\ t \in \mathbb{R}_+.$$

Consequently we may introduce the compensated measure

$$\widetilde{N}(dt, dx) = N(dt, dx) - dt\nu(dx).$$

Accordingly, (2.40) may be rewritten as

$$X(t) = \mu_1 t + \sigma W(t) + \int_0^t \int_{\mathbb{R}-\{0\}} x\widetilde{N}(dt, dx), \quad t \in \mathbb{R}_+. \tag{2.41}$$

Translation-Invariant Semigroup

Let, for $a \in \mathbb{R}$, τ_a denote the translation operator

$$f \in BC(\mathbb{R}) \mapsto \tau_a f \in BC(\mathbb{R})$$

such that

$$(\tau_a f)(x) = f(x-a), \qquad \text{for any} \quad x \in \mathbb{R}.$$

Definition 2.264. Given a one-parameter semigroup $(T_t)_{t\in\mathbb{R}_+}$ on $BC(\mathbb{R})$, we say that it is translation invariant if, for any $a \in \mathbb{R}$ and any $t \in \mathbb{R}_+$,

$$T_t\tau_a = \tau_a T_t.$$

According to Definition 2.56 and the following remark, since a Lévy process $X = (X_t)_{t\in\mathbb{R}_+}$ is a time-homogenous process with independent increments (with $X_0 = 0$), it is completely characterized by the convolution semigroup of the probability laws $\mu_t = P_{X_t}$.

Thanks to Theorem 2.108 we may state the following proposition.

Proposition 2.265. *Any Lévy process is a Markov process; further, it is a Feller process.*

Proof. See, e.g., Applebaum (2004, p. 126). □

Theorem 2.266. *Let $(X_t)_{t\in\mathbb{R}_+}$ be a Lévy process and T a stopping time. Then the process $(Y_t)_{t\in\mathbb{R}_+}$, given by*

$$Y_t = X_{T+t} - X_T,$$

is a Lévy process on the set $]T, \infty[$, adapted to $\mathcal{F}_{T+t}$. Furthermore, Y_t is independent of $\mathcal{F}_T$ and has the same distribution as X_t.

Proof. See, e.g., Protter (2004, p. 23). □

The following result is a consequence of the stationarity and independence of increments of a Lévy process.

Theorem 2.267. *The one-parameter semigroup $(T_t)_{t\in\mathbb{R}_+}$ associated with a Feller process X such that $X(0) = 0$ is translation invariant if and only if X is a Lévy process.*

Proof. See, e.g., Bauer (1981, p. 410) and Applebaum (2004, p. 137). □

In particular, if $(\mu_t)_{t\in\mathbb{R}_+}$ denotes the convolution semigroup of probability measures associated with the Lévy process $(X_t)_{t\in\mathbb{R}_+}$,

$$\mu_t(B) = P(X_t \in B), \quad B \in \mathcal{B}_{\mathbb{R}}, \quad t \in \mathbb{R}_+,$$

then the following relation holds:

$$\mu_t(B-x) = P(X_{t+s} \in B | X_s = x), \quad B \in \mathcal{B}_{\mathbb{R}}, \quad x \in \mathbb{R}, \quad s,t \in \mathbb{R}_+.$$

As a consequence, the transition semigroup of a Lévy process $X = (X_t)_{t\in\mathbb{R}_+}$ is a one-parameter contraction semigroup $(T_t)_{t\in\mathbb{R}_+}$ given by

$$(T(t)f)(x) = \int_{\mathbb{R}} f(x+y)\mu_t(dx), \qquad x \in \mathbb{R},$$

for any $BC(\mathbb{R})$ (e.g., Bauer 1981, p. 405).

If the characteristic triplet of the Lévy process is (μ, σ^2, ν), then one can show (e.g., Sato 1999, p. 208 and references therein) that the infinitesimal generator $\mathcal{A}$ of the semigroup $(T_t)_{t\in\mathbb{R}_+}$ is well defined on $BC(\mathbb{R}) \cap C^2(\mathbb{R})$, and it is given by

$$(\mathcal{A}f)(x) = -\mu f'(x) + \frac{1}{2}\sigma^2 f''(x) + \int_{\mathbb{R}} (f(x+y) - f(x) - I_{[|y|<1]}\, y\, f'(x))\, \nu(dy), \; x \in \mathbb{R}. \quad (2.42)$$

The following examples are trivial consequences of (2.42).

Example 2.268. For the standard Brownian motion the triplet is $(0,1,0)$, so that the infinitesimal generator is

$$(\mathcal{A}f)(x) = \frac{1}{2}f''(x), \qquad x \in \mathbb{R},$$

for $f \in BC(\mathbb{R}) \cap C^2(\mathbb{R})$.

Example 2.269. For a Brownian motion with drift the triplet is of the form $(\mu, \sigma^2, 0)$, so that the infinitesimal generator is

$$(\mathcal{A}f)(x) = -\mu f'(x) + \frac{1}{2}\sigma^2 f''(x), \qquad x \in \mathbb{R},$$

for $f \in BC(\mathbb{R}) \cap C^2(\mathbb{R})$.

Example 2.270. For a Poisson process with triplet $(0, 0, \lambda\epsilon_1)$ the infinitesimal generator is

$$(\mathcal{A}f)(x) = \lambda(f(x+y) - f(x)), \qquad x \in \mathbb{R},$$

for $f \in C_0(\mathbb{R})$.

Example 2.271. For a compound Poisson process with triplet $(0, 0, \nu)$ the infinitesimal generator is

$$(\mathcal{A}f)(x) = \int_{\mathbb{R}} (f(x+y) - f(x))\nu(dy), \qquad x \in \mathbb{R},$$

for $f \in C_0(\mathbb{R})$.

Stable Lévy Processes

As a particular important subclass of Lévy processes we will briefly mention the case of stable Lévy processes.

Definition 2.272. A Lévy process $X = (X_t)_{t\in\mathbb{R}_+}$ is stable if, for any $t \in \mathbb{R}_+$, X_t is a stable random variable.

According to Definition 1.212 and Proposition 1.215, symmetric stable distributions have a characteristic function of the form

$$\phi(u) = \exp\{-\sigma^\alpha |u|^\alpha\}, \quad u \in \mathbb{R},$$

for $\sigma \in \mathbb{R}^*$ and $\alpha \in (0, 2]$. The case $\alpha = 2$ corresponds to the Normal distribution $N(0, 2\sigma^2)$.

Corollary 2.273. *A Lévy symmetric stable process* $(X_t)_{t\in\mathbb{R}_+}$ *has the* scaling property, *i.e., the rescaled process* $(t^{\frac{1}{\alpha}} X_t)_{t\in\mathbb{R}_+}$ *has the same probability law as* $(X_t)_{t\in\mathbb{R}_+}$. *This is a generalization of the specific case for the Wiener process (for which* $\alpha = 2$*) of Proposition 2.176.*

This is one reason why Lévy symmetric stable processes are so important in applications. Another reason why general Lévy stable processes are important in applications is that they exhibit *heavy tails*, i.e., for any $t \in \mathbb{R}_+$, $P(|X_t| > y) \propto y^{-\alpha}$ for $y \to +\infty$, as opposed to the exponential decay of the Gaussian case (e.g., Applebaum 2004 and references therein).

2.12 Exercises and Additions

2.1. Let $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ be a filtration on the measurable space $(\Omega, \mathcal{F})$. Show that $\mathcal{F}_{t^+} = \bigcap_{u>t} \mathcal{F}_u$ is a σ-algebra (Theorem 2.118 and Remark 2.119).

2.2. Prove that two processes that are modifications of each other are equivalent.

2.3. A real-valued stochastic process, indexed in $\mathbb{R}$, is *strictly stationary* if and only if all its joint finite-dimensional distributions are invariant under a parallel time shift, i.e.,

$$F_{X_{t_1},\dots,X_{t_n}}(x_1, \dots, x_n) = F_{X_{t_1+h},\dots,X_{t_n+h}}(x_1, \dots, x_n)$$

for any $n \in \mathbb{N}$, any choice of $t_1, \dots, t_n \in \mathbb{R}$ and $h \in \mathbb{R}$, and any $x_1, \dots, x_n \in \mathbb{R}$.

1. Prove that a process of i.i.d. random variables is strictly stationary.
2. Prove that a time-homogeneous process with independent increments is strictly stationary.
3. Prove that a Gaussian process $(X_t)_{t\in\mathbb{R}}$ is strictly stationary if and only if the following two conditions hold:

(a) $E[X_t] = constant$ for any $t \in \mathbb{R}$.
(b) $Cov[s,t] = K(t-s)$ for any $s, t \in \mathbb{R},\ s < t$.

2.4. An L^2 real-valued stochastic process indexed in $\mathbb{R}$ is *weakly stationary* if and only if the following two conditions hold:

(a) $E[X_t] = constant$ for any $t \in \mathbb{R}$.
(b) $Cov[s,t] = K(t-s)$ for any $s, t \in \mathbb{R},\ s < t$.

1. Prove that an L^2 strictly stationary process is also weakly stationary.
2. Prove that a weakly stationary Gaussian process is also strictly stationary.

2.5. Show that Brownian motion is not stationary.

2.6 (Prediction). Let $(X_{r-j}, \ldots, X_r)$ be a family of random variables representing a sample of a (weakly) stationary stochastic process in L^2. We know that the best approximation in L^2 of an additional random variable X_{r+s}, for any $s \in \mathbb{N}^*$, in terms of $(X_{r-j}, \ldots, X_r)$ is given by $E[Y|X_{r-j}, \ldots, X_r]$. The evaluation of this quantity is generally a hard task. On the other hand, the problem of the best linear approximation can be handled in terms of the covariances of the random variables $X_{r-j}, \ldots, X_r, X_{r+s}$, as follows.

Prove that the best approximation of X_{r+s} in terms of a linear function of $(X_{r-j}, \ldots, X_r)$, is given by

$$\widehat{X}_{r+s} = \sum_{k=0}^{j} a_k X_{r-k},$$

where the a_k satisfy the linear system

$$\sum_{k=0}^{j} a_k c(|k-i|) = c(s+i) \text{ for } 0 \leq i \leq j.$$

Here we have denoted $c(m) = Cov[X_i, X_{i+m}]$.

2.7. Refer to Proposition 2.47. Prove that $\mathcal{F}_T$ is a σ-algebra of the subsets of Ω.

2.8. Prove all the statements of Theorem 2.48.

2.9. Prove Lemma 2.131 by considering the sequence

$$T_n = \sum_{k=1}^{\infty} k 2^{-n} I_{(k-1)2^{-n} \leq T \leq k2^{-n}}.$$

2.10. Let $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$ be a filtration and prove that T is a stopping time if and only if the process $X_t = I_{\{T \leq t\}}$ is adapted to $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$. Show that if T and S are stopping times, then so is $T + S$.

2.11. Show that any (sub- or super-) martingale remains a (sub- or super-) martingale with respect to the induced filtration.

2.12. Let $(X_t)_{t\in\mathbb{R}_+}$ be a martingale in L^2. Show that its increments on nonoverlapping intervals are orthogonal.

2.13. Prove Proposition 2.105. (*Hint:* To prove that 1⇒2, it suffices to use the indicator function on B; to prove that 2⇒1, it should first be shown for simple measurable functions, and then the theorem of approximation of measurable functions through elementary functions is invoked.)

2.14. Prove Remark 2.122.

2.15. Verify Example 2.152.

2.16. Determine the infinitesimal generator of a time-homogeneous Poisson process.

2.17. We say that $(Z_t)_{t\in\mathbb{R}_+}$ is a *compound Poisson process* if it can be expressed as

$$Z_0 = 0$$

and

$$Z_t = \sum_{k=1}^{N_t} Y_k \text{ for } t > 0,$$

where N_t is a Poisson process with intensity parameter $\lambda \in \mathbb{R}_+^*$ and $(Y_k)_{k\in\mathbb{N}^*}$ is a family of i.i.d. random variables, independent of N_t. Show that the compound Poisson process $(Z_t)_{t\in\mathbb{R}_+}$ is a stochastic process with time-homogeneous (stationary) independent increments.

2.18. Show that

1. The Brownian motion and the compound Poisson process are both almost surely continuous at any $t \geq 0$.
2. The Brownian motion is sample continuous, but the compound Poisson process is not sample continuous.

Hence almost sure continuity does not imply sample continuity.

2.19. In the compound Poisson process, assume that the random variables Y_n are i.i.d. with common distribution

$$P(Y_n = a) = P(Y_n = -a) = \frac{1}{2},$$

where $a \in \mathbb{R}_+^*$.

1. Find the characteristic function ϕ of the process $(Z_t)_{t\in\mathbb{R}_+}$.

2. Discuss the limiting behavior of the characteristic function ϕ when $\lambda \to +\infty$ and $a \to +\infty$ in such a way that the product λa^2 is constant.

2.20. An integer-valued stochastic process $(N_t)_{t\in\mathbb{R}_+}$ with stationary (time-homogeneous) independent increments is called a *generalized Poisson process*.

1. Show that the characteristic function of a generalized Poisson process necessarily has the form
$$\phi_{N_t}(u) = e^{\lambda t[\phi(u)-1]}$$
for some $\lambda \in \mathbb{R}_+^*$ and some characteristic function ϕ of a nonnegative integer-valued random variable. The Poisson process corresponds to the degenerate case $\phi(u) = e^{iu}$.
2. Let $(N_t^{(k)})_{t\in\mathbb{R}_+}$ be a sequence of independent Poisson processes with respective parameters λ_k. Assume that $\lambda = \sum_{k=1}^{+\infty} \lambda_k < +\infty$. Show that the process
$$N_t^{(k)} = \sum_{k=1}^{+\infty} kN_t^{(k)},\ t \in \mathbb{R}_+,$$
is a generalized Poisson process, with characteristic function
$$\phi(u) = \sum_{k=1}^{+\infty} \frac{\lambda_k}{\lambda} e^{iku}.$$
3. Show that any generalized Poisson process can be represented as a compound Poisson process. Conversely, if the random variables Y_k in the compound Poisson process are integer-valued, then the process is a generalized Poisson process.

2.21. Let $(X_n)_{n\in\mathbb{N}} \subset E$ be a *Markov chain*, i.e., a discrete-time Markov jump process, where E is a countable set. Let $i, j \in E$ be *states* of the process; j is said to be *accessible* from state i if for some integer $n \geq 0$, $p_{ij}(n) > 0$, i.e., state j is accessible from state i if there is positive probability that in a finite number of transition states j can be reached starting from state i. Two states i and j, each accessible to the other, are said to *communicate*, and we write $i \leftrightarrow j$. If two states i and j do not communicate, then
$$p_{ij}(n) = 0 \quad \forall n \geq 0,$$

$$p_{ji}(n) = 0 \quad \forall n \geq 0,$$
or both relations are true.

We define the period of state i, written $d(i)$, as the greatest common divisor of all integers $n \geq 1$ for which $p_{ii}(n) > 0$ (if $p_{ii}(n) = 0$ for all $n \geq 1$, define $d(i) = 0$).

1. Show that the concept of communication is an equivalence relationship.
2. Show that, if $i \leftrightarrow j$, then $d(i) = d(j)$.
3. Show that if state i has period $d(i)$, then there exists an integer N depending on i such that for all integers $n \geq N$

$$p_{ii}(nd(i)) > 0.$$

2.22.

1. Consider two urns A and B containing a total of N balls. A ball is selected at random (all selections are equally likely) at time $t = 1, 2, \ldots$ from among the N balls. The drawn ball is placed with probability p in urn A and with probability $q = 1 - p$ in urn B. The state of the system at each trial is represented by the number of balls in A. Determine the transition matrix for this Markov chain.
2. Assume that at each time t there are exactly k balls in A. At time $t+1$ an urn is selected at random proportionally to its content (i.e., A is chosen with probability k/N and B with probability $(N-k)/N$). Then a ball is selected either from A with probability p or from B with probability $1-p$ and placed in the previously chosen urn. Determine the transition matrix for this Markov chain.
3. Now assume that at time $t+1$ a ball and an urn are chosen with probability depending on the contents of the urn (i.e., a ball is chosen from A with probability $p = k/N$ or from B with probability q; urn A is chosen with probability p and B with probability q). Determine the transition matrix of the Markov chain.
4. Determine the equivalence classes in parts 1, 2, and 3.

2.23. Let $(X_n)_{n\in\mathbb{N}}$ be a Markov chain whose transition probabilities are $p_{ij} = 1/[e(j-i)!]$ for $i = 0, 1, \ldots$ and $j = i, i+1, \ldots$. Verify the martingale property for

- $Y_n = X_n - n$
- $U_n = Y_n^2 - n$
- $V_n = \exp\{X_n - n(e-1)\}$

2.24. Let $(X_t)_{t\in\mathbb{R}_+}$ be a process with the following properties:

- $X_0 = 0$.
- For any $0 \leq t_0 < t_1 < \cdots < t_n$ the random variables $X_{t_k} - X_{t_{k-1}}$ $(1 \leq k \leq n)$ are independent.
- If $0 \leq s < t$, $X_t - X_s$ is normally distributed with

$$E(X_t - X_s) = (t-s)\mu, \qquad E\left[(X_t - X_s)^2\right] = (t-s)\sigma^2$$

  where μ, σ are real constants $(\sigma \neq 0)$.

The process $(X_t)_{t\in\mathbb{R}_+}$ is called a Brownian motion with *drift* μ and *variance* σ^2. (Note that if $\mu = 0$ and $\sigma = 1$, then X_t is the so-called standard Brownian motion.) Show that $Cov(X_t, X_s) = \sigma^2 \min\{s, t\}$ and $(X_t - \mu t)/\sigma$ is a standard Brownian motion.

2.25. Show that if $(X_t)_{t\in\mathbb{R}_+}$ is a Brownian motion, then the processes

$$Y_t = cX_{t/c^2} \quad \text{for fixed} \quad c > 0,$$

$$U_t = \begin{cases} tX_{1/t} & \text{for} \ \ t > 0, \\ 0 & \text{for} \ \ t = 0, \end{cases}$$

and

$$V_t = X_{t+h} - X_h \quad \text{for fixed} \quad h > 0$$

are all Brownian motions.

2.26. Let $(X_t)_{t\in\mathbb{R}_+}$ be a Brownian motion; given $a \in \mathbb{R}$, let τ_a denote the first passage time of the process started at $X(0) = 0$.

1. Show that for any $t \in \mathbb{R}_+$

$$P(\tau_a \leq t) = \frac{2}{\sqrt{2\pi t}} \int_a^{+\infty} \exp\left\{-\frac{y^2}{2t}\right\} dy.$$

   Conclude that the first passage time through any given point is a.s. finite, but its mean value is infinite.
2. Use the preceding result to show that one-dimensional Brownian motion is *recurrent*, in the sense that, for any $x \in \mathbb{R}$ and any $T \in \mathbb{R}_+$,

$$P(W_t = x \quad \text{for some} \quad t > T) = 1.$$

   This means that a Brownian motion returns to every point infinitely many times, for arbitrary large times.

2.27. Let $(X_t)_{t\in\mathbb{R}_+}$ be a Brownian motion, and let $M_t = \max_{0\leq s\leq t} X_s$. Prove that $Y_t = M_t - X_t$ is a continuous-time Markov process. (*Hint:* Note that for $t' < t$

$$Y(t) = \max\left\{\max_{t'\leq s\leq t}\{(X_s - X_{t'})\}, Y_{t'}\right\} - (X_t - X_{t'}).)$$

2.28. Let T be a stopping time for a Brownian motion $(X_t)_{t\in\mathbb{R}_+}$. Then the process

$$Y_t = X_{t+T} - X_T, \qquad t \geq 0,$$

is a Brownian motion, and $\sigma(Y_t, t \geq 0)$ is independent of $\sigma(X_t, 0 \leq t \leq T)$.

(*Hint:* At first consider T constant. Then suppose that the range of T is a countable set and finally approximate T by a sequence of stopping times such as in Lemma 2.131.)

2.29. Let $(X_t)_{t\in\mathbb{R}_+}$ be an n-dimensional Brownian motion starting at 0, and let $U \in \mathbb{R}^{n\times n}$ be a (constant) orthogonal matrix, i.e., $UU^T = I$. Prove that

$$\widetilde{X}_t \equiv UX_t$$

is also a Brownian motion.

2.30. Let $(X_t)_{t\in\mathbb{R}_+}$ be a Lévy process:

1. Show that the characteristic function of X_t is infinitely divisible.
2. Suppose that the law of X_1 is $P_{X_1} = \mu$. Then, for any $t > 0$ the law of X_t is $P_{X_1} = \mu^t$.
3. Given two Lévy processes $(X_t)_{t\in\mathbb{R}_+}$ and $(X'_t)_{t\in\mathbb{R}_+}$, if $P_{X_1} = P_{X'_1}$, then the two processes are identical in law.

We call $\mu = P_{X_1}$ the infinitely divisible distribution of the Łévy process $(X_t)_{t\in\mathbb{R}_+}$.

2.31. Consider a Lévy process $X = (X_t)_{t\in\mathbb{R}_+}$. Show that the finite-dimensional distributions of X are determined by the one-dimensional marginal
distributions.

2.32. A Lévy process $(X_t)_{t\in\mathbb{R}_+}$ is a *subordinator* if it is also a real and non-negative process.

1. Show that sample paths of a subordinator are increasing.
2. Show that a Lévy process $(X_t)_{t\in\mathbb{R}_+}$ is a subordinator if and only if $X_1 \geq 0$ almost surely.

2.33. Show that Brownian motions with drift, i.e.,

$$X_t = \sigma W_t + \alpha t \text{ for } \alpha, \sigma \in \mathbb{R},$$

are the only Lévy processes with continuous paths.

2.34. Consider two sequences of real numbers $(\alpha_k)_{k\in\mathbb{N}}$ and $(\beta_k)_{k\in\mathbb{N}}$ such that $\sum_{k\in\mathbb{N}} \beta_k^2 \alpha_k < +\infty$. Let N_t^k be a sequence of Poisson processes with intensities α_k and $k \in \mathbb{N}$, respectively.

Then the process

$$X_t = \sum_{k\in\mathbb{N}} \beta_k (N_t^k - \alpha_k t), \quad t \in \mathbb{R}_+,$$

is a Lévy process having ν as its Lévy measure.

2.35. Show that

1. Any Lévy process is a Markov process.
2. Conversely, any stochastically continuous and temporarily homogeneous Markov process on $\mathbb{R}$ is a Lévy process.

2.36. According to, e.g., Grigoriu (2002), we define as a classical *semimartingale* any adapted RCLL process X_t that admits the following decomposition:

$$X_t = X_0 + M_t + A_t,$$

where M_t is a local martingale and A_t is a finite variation (on compacts) RCLL process such that $M_0 = A_0 = 0$.

1. Show that any Lévy process is a semimartingale.
2. Show that the Poisson process is a semimartingale.
3. Show that the square of a Wiener process is a semimartingale.

2.37 (Poisson process and order statistics). Let $X_1, \ldots, X_n$ denote a *sample*, i.e., a family of nondegenerate i.i.d. random variables with common cumulative distribution function F. We define an *ordered sample* as the family

$$X_{n,n} \leq \cdots \leq X_{1,n},$$

so that $X_{n,n} = \min\{X_1, \ldots, X_n\}$ and $X_{1,n} = \max\{X_1, \ldots, X_n\}$. The random variable $X_{k,n}$ is called the *k-order statistic*.

Let $N = (N_t)_{t\in\mathbb{R}_+}$ be a homogeneous Poisson process with intensity $\lambda > 0$. Prove that the arrival times T_i of N in $]0,t]$, conditionally upon $\{N_t = n\}$, have the same distribution as the order statistics of a uniform sample on $]0,t[$ of size n, i.e., for all Borel sets A in $\mathbb{R}_+$ and any $n \in \mathbb{N}$ we have

$$P((T_1, T_2, \ldots, T_{N_t}) \in A | N_t = n) = P((U_{n,n}, \ldots, U_{1,n}) \in A).$$

2.38. (Self-similarity). A real-valued stochastic process $(X_t)_{t\in\mathbb{R}_+}$ is said to be *self-similar* with *index* $H > 0$ (H-ss) if its finite-dimensional distributions satisfy the relation

$$(X_{at_1}, \ldots, X_{at_n}) \stackrel{d}{=} a^H (X_{t_1}, \ldots, X_{t_n})$$

for any choice of $a > 0$ and $t_1, \ldots, t_n \in \mathbb{R}_+$. Show that a Gaussian process with mean function $m_t = E[X_t]$ and covariance function $K(s,t) = Cov(X_s, X_t)$ is H-ss for some $H > 0$ if and only if

$$m_t = ct^H, \ \text{ and } \ K(s,t) = s^{2H} C(t/s, 1)$$

for some constant $c \in \mathbb{R}$ and some nonnegative definite function C. As a consequence, show that the standard Brownian motion is 1/2-ss. Also, show that any α-stable process is $1/\alpha$-ss.

2.39 (Affine processes). Let $\Phi = (\Phi_t)_{t\in\mathbb{R}_+}$ be a process on a given probability space $(\Omega, \mathcal{F}, P)$ such that $E[\|\Phi_t\|] < +\infty$ for each $t \in \mathbb{R}_+$. The past-future filtration associated with Φ is defined as the family

$$\mathcal{F}_{s,T} = \sigma\{\Phi_u | u \in [0,s] \cup [T, +\infty[\}.$$

We shall call Φ an *affine process* if it satisfies

$$E[\Phi_t|\mathcal{F}_{s,T}] = \frac{T-t}{T-s}\Phi_s + \frac{t-s}{T-s}\Phi_T, \ s < t < T.$$

Show that the preceding condition is equivalent to the property that for $s \leq t < t' \leq u$ the quantity

$$E\left[\frac{\Phi_t - \Phi_{t'}}{t-t'}\middle| \mathcal{F}_{s,u}\right] = \frac{\Phi_u - \Phi_s}{u-s}$$

and, hence, does not depend on the pair (t, t').

2.40. Prove that Brownian motion is an affine process.

2.41. Let $X = (X_t)_{t\in\mathbb{R}_+}$ be a Lévy process such that $E[\|X_t\|] < +\infty$ for each $t \in \mathbb{R}_+$. Show that X is an affine process.

2.42. Consider a process $M = (M_t)_{t\in\mathbb{R}_+}$ that is adapted to the filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ on a probability space $(\Omega, \mathcal{F}, P)$ and satisfies

$$E[\|M_t\|] < +\infty \text{ and } E\left[\int_0^t du|M_u|\right] < +\infty \text{ for any } t > 0.$$

Prove that the following two conditions are equivalent:

1. M is an $\mathcal{F}_t$-martingale.
2. For every $t > s$,

$$E\left[\frac{1}{t-s}\int_s^t duM_u\middle| \mathcal{F}_s\right] = M_s.$$

2.43 (Empirical process and Brownian bridge). Let $U_1, \ldots, U_n, \ldots,$ be a sequence of i.i.d. random variables uniformly distributed on $[0, 1]$. Define the stochastic process $b^{(n)}$ on the interval $[0, 1]$ as follows:

$$b^{(n)}(t) = \sqrt{n}\left(\frac{1}{n}\sum_{k=1}^{n} I_{[0,t]}(U_k) - t\right), \ t \in [0, 1].$$

1. For any s and t in $[0, 1]$, compute $E[b^{(n)}(t)]$ and $Cov[b^{(n)}(s), b^{(n)}(t)]$.
2. Prove that, as $n \to \infty$, the finite-dimensional distributions of the process $(b^{(n)}(t))_{t\in[0,1]}$ converge weakly toward those of a Gaussian process on $[0, 1]$ whose mean and covariance functions are the same as those of $b^{(n)}$.

An Introduction to Continuous-Time Stochastic Processes
Theory, Models, and Applications to Finance, Biology, and Medicine
Capasso, V.; Bakstein, D.
2012, XIII, 434 p. 14 illus.,
ISBN: 978-0-8176-8346-7
A product of Birkhäuser Basel