General Theory of Markov Processes

MICHAEL SHARPE

General Theory of Markov Processes

This is Volume 133 in
PURE AND APPLIED MATHEMATICS
H. Bass, A. Borel, J. Moser, and S.-T. Yau, editors
Paul A. Smith and Samuel Eilenberg, founding editors
A list of titles in this series appears at the end of this volume.

General Theory of Markov Processes

Michael Sharpe
Department of Mathematics
University of California at San Diego
La Jolla, California

ACADEMIC PRESS, INC.
Harcourt Brace Jovanovich, Publishers
Boston San Diego New York
Berkeley London Sydney
Tokyo Toronto

ACADEMIC PRESS, INC.
1250 Sixth Avenue, San Diego, CA 92101

United Kingdom Edition published by
ACADEMIC PRESS INC. (LONDON) LTD.
24-28 Oval Road, London NW1 7DX

Library of Congress Cataloging-in-Publication Data

Sharpe, Michael, Date -
General theory of Markov processes / Michael Sharpe.
p. cm. — (Pure and applied mathematics : 133)
Bibliography: p.
Includes indexes.
ISBN 0-12-639060-6
1. Markov processes. I. Title. II. Series: Pure and applied
mathematics (Academic Press) : 133.
QA3.P8 vol. 133
[QA274.7]
519.2—dc19 88-18088
CIP

Printed in the United States of America
88 89 90 91 9 8 7 6 5 4 3 2 1

To my wife Sheila and my son Colin

This Page Intentionally Left Blank

Preface

This work is intended to serve as a reference to the theory of *right processes*, a very general class of right continuous strong Markov processes. The use of the term *general theory* is meant to suggest both the absence of hypotheses of special type other than those for right processes, and the coordination of the methods with those of the general theory of processes, as exposed in the first two volumes of *Probabilités et Potentiel* by Dellacherie and Meyer. We do provide in the appendix a fairly extensive discussion and summary of the general techniques needed in the text, with hopes that it may lead the reader to a fuller appreciation of the Dellacherie-Meyer volumes.

The original definition of right process (*processus droit*) was set down twenty years ago by Meyer as an abstraction of certain properties possessed by standard Markov processes, which had been up to that time the largest class of strong Markov processes that could be shown to have an intimate connection with abstract potential theory. The hypotheses of Meyer were weakened in the subsequent lecture notes of Getoor (1975). Right processes in the sense of Meyer or Getoor do form a class large enough to encompass most right continuous Markov processes of practical interest such as Brownian motion, diffusions, Lévy processes (processes with stationary independent increments), Feller processes and so on, constructed from reasonable transition semigroups. However, the form of hypotheses discussed by Meyer and Getoor contains a serious flaw, in that their hypotheses are not invariant under the classical transformations of Markov processes such as killing, time-change, mappings of the state space, and Doob's h-transforms. Motivated by the wish to have a setting which is preserved by essentially all these transformations, we propose hypotheses for right processes weaker than those of either Meyer or Getoor, but which remain strong enough to guarantee a rich theory of sample path behavior and close links with potential theory.

The point of view of the book is chiefly to study the probabilistic structure of a given right process as expressed through such objects as its homogeneous functionals, its additive and multiplicative functionals, its associated stochastic calculus, and to consider the transformations of right processes that yield other right processes. It has been a constant goal to avoid imposing secondary hypotheses which would limit the domain of applicability. There is only one section concerning the construction of a right process from a nice (Ray) semigroup, and while adequate for constructing some classical examples, it is not of great generality. There is no discussion of construction of Markov processes by solving Stroock-Varadhan type martingale problems. The recent book of Ethier-Kurtz (1986) has much on these matters.

Explicit examples of right processes are discussed principally in the exercises. The connections between right processes and abstract potential theory are discussed though not always in full detail. For example, though there is a discussion of the Hunt-Shih identification of hitting operators and réduite of an excessive function on a set, we do not present a complete proof. The reader interested in questions of more direct potential theoretic type is referred to volumes III and IV of Dellacherie-Meyer.

The sections on multiplicative functionals and homogeneous random measures, the latter a generalization of additive functionals, bring up to date the older books of Meyer and Blumenthal-Getoor. Especially in the sections on Lévy systems and exit systems, there is a penalty to be paid for the breadth of the hypotheses, requiring us to construct kernels on spaces larger than the state space so that the statements of the results will look a bit unusual to experts familiar with their forms under restrictive measurability conditions. However, the applications of these constructions do not appear to be affected in any essential way by this complication.

It is a pleasure to thank those individuals whose comments on earlier versions have eliminated many inaccuracies, inconsistencies and irrelevancies. Marti Bachman, Ron Getoor, Joe Glover, Bernard Maisonneuve, Joanna Mitro, Wenchuan Mo, Art Pittenger, Phil Protter, Tom Salisbury and Michel Weil provided me with valuable feedback for which I am very grateful. Thanks are also due to Neola Crimmins, whose expert entry of part of the first draft simplified the task of assembling the final document in TEX format.

La Jolla, 1988.

Contents

This Page Intentionally Left Blank

I Fundamental Hypotheses

1. Markov Property, Transition Functions and Entrance Laws

A stochastic process indexed by a subset of the real line has the **Markov property** if, roughly speaking, the past and future are conditionally independent given the present, for every possible value of the present. See (1.1) below for the precise specification. In this definition, the state space is required to have only measurable structure—no algebra or topology is involved. Nevertheless, because of applications to special examples and our focus on path regularizations which would otherwise take a different form, we shall work exclusively with topological state spaces. The minimal hypothesis on every state space E shall be that E is a Radon topological space. See §A1. This is not a burdensome restriction. Every Polish (:=complete, separable, metrizable) space and every locally compact Hausdorff space with countable base (LCCB) is Radonian.

The notation $\mathcal{B}(E)$ stands for the Borel σ-algebra on E, but we shall use the simpler notation $\mathcal{E}$ in its place unless clarity dictates otherwise. The notation $\mathcal{E}^u$ will, following the pattern described in the Appendix, denote the σ-algebra of universally measurable subsets of E. Other σ-algebras intermediate to $\mathcal{E}$ and $\mathcal{E}^u$ will be introduced later. We shall always denote a generic such σ-algebra by $\mathcal{E}^\bullet$ with the superscript $^\bullet$ usually being one of $0, r, e$, referring to the σ-algebras on E generated by the Borel, Ray and excessive functions respectively. Thus $\mathcal{E}^0$ is just another name for $\mathcal{E}$. See §10. In later sections, we shall make a distinction between $\mathcal{E}^0$ and $\mathcal{E}$, identifying $\mathcal{E}$ with $\mathcal{E}^r$ instead of $\mathcal{E}^0$. This will require minor reinterpretation of some of the constructs in this chapter, but to do otherwise would lead to serious notational complications later.

The reader is now assumed to be familiar with the terminology established in the Appendix, especially in A0–A3. In particular, given a σ-algebra $\mathcal{M}$ on a space M, b$\mathcal{M}$ (resp., p$\mathcal{M}$) stands for the class of *bounded* (resp., *positive*) $\mathcal{M}$-measurable functions on M. (Positive always refers to values in $[0,\infty]$, rather than the positive reals).

Let $(\Omega, \mathcal{G}, \mathbf{P})$ be a probability space, I be an index set contained in the real line $\mathbf{R}$, and let $X = (X_t)_{t\in I}$ be a **stochastic process** indexed by I, with values in E. That is, $(X_t)_{t\in I}$ is a collection of measurable maps of $(\Omega, \mathcal{G})$ into $(E, \mathcal{E})$. In order to emphasize the dependence here on $\mathcal{E}$, we call X an $\mathcal{E}$-stochastic process. Similar definitions will apply when $\mathcal{E}$ is replaced with a larger σ-algebra $\mathcal{E}^\bullet$. It is, of course, a more demanding condition for X to be an $\mathcal{E}^\bullet$-stochastic process, as it is required in this case that for every $t \in I$, $\{X_t \in F\} := \{\omega \in \Omega : X_t(\omega) \in F\}$ be in $\mathcal{G}$ for every set F in $\mathcal{E}^\bullet$ rather than for every F in $\mathcal{E}$.

Corresponding to a fixed σ-algebra $\mathcal{E}^\bullet$ on E and a fixed $\mathcal{E}^\bullet$-stochastic process X on Ω, the natural σ-algebra $\mathcal{F}^\bullet_{\leq t}$ (or, more simply, $\mathcal{F}^\bullet_t$) is defined as $\sigma\{f(X_r) : r \in I, r \leq t, f \in \mathcal{E}^\bullet\}$. A similar definition specifies the σ-algebra $\mathcal{F}^\bullet_{\geq t}$ of the future from t. Thus, for example, $\mathcal{F}^0_t$ (resp., $\mathcal{F}^u_t$) denotes the σ-algebra generated by the maps $f(X_r)$ with $r \leq t$ and f in $\mathcal{E}^0 (:= \mathcal{E})$ (resp., f in $\mathcal{E}^u$).

The process X has the **$\mathcal{E}^\bullet$-Markov property** if the σ-algebras $\mathcal{F}^\bullet_{\leq t}$, $\mathcal{F}^\bullet_{\geq t}$ are conditionally independent given X_t, for every $t \in I$. That is, for $t \in I$, $A \in \mathcal{F}^\bullet_{\leq t}$ and $B \in \mathcal{F}^\bullet_{\geq t}$,

$$\mathbf{P}\{A \cap B \mid X_t\} = \mathbf{P}\{A \mid X_t\} \cdot \mathbf{P}\{B \mid X_t\}. \tag{1.1}$$

The need for the prefix $\mathcal{E}^\bullet$ is only temporary, as we shall see after the discussion of augmentation procedures in §6. Under the condition (1.1), one may compute, using the well known properties of conditional expectations,

$$\begin{aligned}\mathbf{P}\{\, A \cap B \,\} &= \mathbf{P}\{\, \mathbf{P}\{A \cap B \mid X_t\} \,\} \\ &= \mathbf{P}\{\, \mathbf{P}\{A \mid X_t\}\, \mathbf{P}\{B \mid X_t\} \,\} \\ &= \mathbf{P}\{\, \mathbf{P}\{B \mid X_t\}\, ; A \,\}.\end{aligned}$$

As $A \in \mathcal{F}^\bullet_t$ is arbitrary, it follows that (1.1) implies

$$\mathbf{P}\{B \mid \mathcal{F}^\bullet_t\} = \mathbf{P}\{B \mid X_t\} \tag{1.2}$$

for every $B \in \mathcal{F}^\bullet_{\geq t}$, $t \in I$. That is, prediction of future behavior of X based on the entire past is only as valuable as the predictor based on the present value X_t alone. Conversely, the condition (1.2) implies (1.1) by similar manipulations, and consequently (1.2) is also referred to as the $\mathcal{E}^\bullet$-Markov property of X. In many respects, (1.2) is more convenient to manipulate and generalize. In the first place, it is reasonable and useful to replace the filtration $(\mathcal{F}_t^\bullet)$ with a more general filtration $(\mathcal{G}_t)$ to which (X_t) is $\mathcal{E}^\bullet$-adapted. This leads us to say that (X_t) is **$\mathcal{E}^\bullet$-Markovian** with respect to $(\mathcal{G}_t)$ if X is $\mathcal{E}^\bullet$-adapted to $(\mathcal{G}_t)$, and if, for all $t \in I$ and all $B \in \mathcal{F}_{\geq t}^\bullet$,

$$(1.3) \qquad \mathbf{P}\{B \mid \mathcal{G}_t\} = \mathbf{P}\{B \mid X_t\}.$$

In applications, (1.3) has a more convenient form

$$(1.4) \qquad \mathbf{P}\{H \mid \mathcal{G}_t\} = \mathbf{P}\{H \mid X_t\}, \qquad H \in \mathrm{p}\mathcal{F}_{\geq t}^\bullet.$$

Formula (1.4) is an immediate consequence of (1.3), starting with the case $H = 1_B$, $B \in \mathcal{F}_{\geq t}^\bullet$, and making use of the Monotone Class Theorem (A0.1).

The definition above is too crude to be useful except when I is a discrete subset of $\mathbf{R}$. We bring more precision to bear by introduction of the notion of a transition function $(P_{s,t})$ for X.

(1.5) DEFINITION. *A family $(P_{s,t})$ of Markov kernels on $(E, \mathcal{E}^\bullet)$ indexed by pairs $s, t \in I$ with $s \leq t$ is a* **transition function** *on $(E, \mathcal{E}^\bullet)$ if, for all $r \leq s \leq t$ in I and all $x \in E$, $B \in \mathcal{E}^\bullet$*

$$P_{r,t}(x, B) = \int_E P_{r,s}(x, dy)\, P_{s,t}(y, B).$$

In accordance with the discussion of kernels in A3, $P_{s,t}(x, dy)$ is a kernel on $(E, \mathcal{E}^\bullet)$ provided that, for all $x \in E$, $P_{s,t}(x, dy)$ is a positive measure on $(E, \mathcal{E}^\bullet)$, and for every $B \in \mathcal{E}^\bullet$, $x \to P_{s,t}(x, B)$ is $\mathcal{E}^\bullet$ measurable. In addition, $P_{s,t}(x, dy)$ is a **Markov** kernel if $P_{s,t}(x, E) = 1$ for all $x \in E$. The equation in (1.5) is called the **Chapman-Kolmogorov equation.**

Define the action of the Markov kernel $P_{s,t}$ on $\mathrm{b}\mathcal{E}^\bullet$ (resp., $\mathrm{p}\mathcal{E}^\bullet$) by

$$P_{s,t}f(x) := \int P_{s,t}(x, dy)\, f(y), \qquad f \in \mathrm{p}\mathcal{E}^\bullet \cup \mathrm{b}\mathcal{E}^\bullet,$$

so that $P_{s,t}f \in \mathrm{b}\mathcal{E}^\bullet$ (resp., $\mathrm{p}\mathcal{E}^\bullet$.) See §A3. We say that a transition function $(P_{s,t})$ on $(E, \mathcal{E}^\bullet)$ is the transition function for a process $(X_t)_{t \in I}$ with values in E, and satisfying the Markov property (1.4) relative to $(\mathcal{G}_t)$ in case

$$(1.6) \qquad \mathbf{P}\{f(X_t) \mid \mathcal{G}_s\} = P_{s,t}f(X_s), \quad s \leq t \in I, f \in \mathrm{b}\mathcal{E}^\bullet.$$

(1.7) THEOREM. *Let $(X_t)_{t \in I}$ be $\mathcal{E}^\bullet$-adapted to $(\mathcal{G}_t)$, and suppose that $(P_{s,t})$ is a transition function on $(E, \mathcal{E}^\bullet)$ such that (1.6) holds for every $s \leq t \in I$ and every $f \in \mathrm{b}\mathcal{E}^\bullet$. Then X has the Markov property (1.4).*

PROOF: The class $\mathcal{H}$ of random variables in $\mathrm{b}\mathcal{F}_{\geq t}^\bullet$ for which (1.4) holds is clearly an MVS (see A0) because of monotonicity properties of conditional expectations. By hypothesis, $\mathcal{H}$ contains every H of the form $f(X_t)$ with $f \in b\mathcal{E}^\bullet$. As $b\mathcal{F}^\bullet_{\geq t}$ is generated by the multiplicative class $\mathcal{V} = \cup_n \mathcal{V}_n$, where $\mathcal{V}_n$ is the collection of products $F_1 F_2 \cdots F_n$ with $F_j = f_j(X_{t_j})$, $t \leq t_1 \leq t_2 \leq \cdots \leq t_n$, $f_j \in b\mathcal{E}^\bullet$, it suffices by the Monotone Class Theorem to verify that $\mathcal{V} \subset \mathcal{H}$. Proceed by induction on n to get $\mathcal{V}_n \subset \mathcal{H}$ for all $n \geq 1$. By our first remarks above, $\mathcal{V}_1 \subset \mathcal{H}$. Suppose, inductively, that $\mathcal{V}_n \subset \mathcal{H}$ and let $G = F_1 \ldots F_{n+1} \in \mathcal{V}_{n+1}$. Compute $\mathbf{P}\{G \mid \mathcal{G}_t\}$ by first conditioning relative to $\mathcal{G}_{t_n}$, so that

$$\mathbf{P}\{G \mid \mathcal{G}_t\} = \mathbf{P}\{\mathbf{P}\{G \mid \mathcal{G}_{t_n}\} \mid \mathcal{G}_t\} = \mathbf{P}\{F_1 \ldots F_n P_{t_n, t_{n+1}} f_{n+1}(X_{t_n}) \mid \mathcal{G}_t\}.$$

However, the random variable being conditioned in the last term is clearly in $\mathcal{V}_n$, and thus, by inductive hypothesis, $\mathcal{G}_t$ may be replaced by X_t. The same calculation with $\mathcal{G}_t$ replaced throughout by X_t completes the inductive step by proving $\mathbf{P}\{G \mid \mathcal{G}_t\} = \mathbf{P}\{G \mid X_t\}$, which finishes the proof.

We shall be interested primarily in the case $I = \mathbf{R}^+ := [0, \infty[$ though the cases $]0, \infty[$, $]-\infty, \infty[$ and $]0, 1[$ also arise frequently in practice.

A family $(P_t)_{t \geq 0}$ of Markov kernels on $(E, \mathcal{E}^\bullet)$ is called a **Markov transition semigroup** or simply a transition semigroup in case

$$P_{t+s} f(x) = P_t(P_s f)(x), \quad t, s \geq 0, x \in E, f \in b\mathcal{E}^\bullet.$$

A transition function $(P_{s,t})$ indexed by $s \leq t \in \mathbf{R}^+$ is called **temporally homogeneous** if there is a transition semigroup (P_t) with $P_{s,t} = P_{t-s}$ for all $s \leq t$. Starting with a transition semigroup (P_t), $P_{s,t} := P_{t-s}$ defines a temporally homogeneous transition function.

A Markov process X satisfying (1.6) with a homogeneous transition function (P_t) has the characteristic property

$$\text{(1.8)} \qquad \mathbf{P}\{f(X_{t+s}) \mid \mathcal{G}_t\} = P_s f(X_t), \quad t, s \geq 0, \ f \in b\mathcal{E}^\bullet.$$

This is the **simple Markov property** of X relative to (P_t).

The Markov processes considered here will be temporally homogeneous for the most part. See however exercise (1.15), which deals with the so-called space-time process connected with a general Markov process.

Suppose now that $(X_t)_{t \geq 0}$ has the Markov property (1.8) relative to $(\Omega, \mathcal{G}, \mathcal{G}_t, \mathbf{P})$, with transition semigroup (P_t). The distribution μ_0 of X_0 is called the **initial law** of X, and the distribution μ_t of X_t then satisfies $\mu_t = \mu_0 P_t$ for all $t \geq 0$. That is, for $f \in b\mathcal{E}^\bullet$,

$$\mu_t(f) := \int f \, d\mu_t = \mathbf{P} f(X_t) = \mathbf{P}\, P_t f(X_0) = \mu_0(P_t f).$$

If the index set for X were instead $]0,\infty[$, there would be no initial law μ_0 definable as above. However, the μ_t would obviously satisfy the identities

$$\mu_{t+s} = \mu_t P_s, \qquad t, s > 0. \tag{1.9}$$

A family $(\mu_t)_{t>0}$ of positive measures on $(E, \mathcal{E}^\bullet)$ satisfying (1.9) is called an **entrance law** for the semigroup (P_t). It is called *finite* in case $\mu_t(E) < \infty$ $\forall t > 0$, *bounded* if $\sup_t \mu_t(E) < \infty$, *probability* if $\mu_t(E) = 1$ for all t. If there is a measure μ_0 such that $\mu_t = \mu_0 P_t$ for all $t > 0$, then μ_0 is said to **close** the entrance law $(\mu_t)_{t>0}$. A probability entrance law $(\mu_t)_{t>0}$ need not have a closing element μ_0. For example, let E be the open right half line $\mathbf{R}^{++}$ and let $P_t(x, dy) := \epsilon_{x+t}(dy)$ —unit mass a location $x+t$. Then, for $t > 0$, $\mu_t(dy) := \epsilon_t(dy)$ defines a probability entrance law for (P_t) without a closing element. See Chapter V for the compactification theory needed to permit the representation of a closing element for an arbitrary probability entrance law.

A (temporally homogeneous) Markov process $(X_t)_{t\geq 0}$ satisfying (1.8) and having initial law μ_0 necessarily satisfies the more general identities

$$\begin{aligned} \mathbf{P}\{f_1(X_{t_1}) f_2(X_{t_2}) \cdots f_n(X_{t_n})\} \\ = \mu_0 \left(P_{t_1} \left(f_1 \cdot P_{t_2 - t_1} \left(f_2 \cdots (f_{t_n} \cdot P_{t_n - t_{n-1}} f_n) \cdots \right)\right)\right), \end{aligned} \tag{1.10}$$

for $0 \leq t_1 \leq t_2 \leq \cdots \leq t_n$, $f_1, \cdots, f_n \in b\mathcal{E}^\bullet$. This is a simple consequence of (1.8) via an induction argument. The last formula is perhaps more intuitive in its differential version, which states that under the same conditions as above,

$$\begin{aligned} \mathbf{P}\{X_0 \in dx_0, X_{t_1} \in dx_1, \cdots, X_{t_n} \in dx_n\} \\ = \mu_0(dx_0) P_{t_1}(x_0, dx_1) \cdots P_{t_n - t_{n-1}}(x_{n-1}, dx_n). \end{aligned} \tag{1.11}$$

In this form, the Markov property corresponds to the Huygens principle in wave propagation—in order to compute $\mathbf{P}\{X_t \in dx\}$, one may imagine interposing a barrier at time $t_1 < t$ and, knowing the position X_{t_1}, perform calculations supposing that the process starts afresh at X_{t_1}. The integral version (1.10) asserts that the total probability that $X_t \in dx$ is obtained by adding the above probabilities over all possible positions X_{t_1}, weighted by the probabilities of reaching the points X_{t_1} in the first place.

(1.12) EXERCISE. *Formulate the appropriate versions of (1.10) and (1.11) in the case where X is homogeneous Markov with time parameter set* $]0,\infty[$.

The next pair of exercises is designed to give the reader a little practice with arguments involving completions. This kind of "sandwiching" will be used repeatedly in later sections. Exercise (1.14) will show that there is no need to maintain any distinction between different $\mathcal{E}^\bullet$-Markov properties, provided the filtration is sufficiently rich.

(1.13) EXERCISE. *Let* (P_t) *preserve each of the* σ*-algebras* $\mathcal{E}^\bullet$, $\mathcal{E}'$, *with* $\mathcal{E} \subset \mathcal{E}^\bullet \subset \mathcal{E}' \subset \mathcal{E}^u$. *Let* $(X_t)_{t\geq 0}$ *be defined on* $(\Omega, \mathcal{G}, \mathcal{G}_t, \mathbf{P})$ *with* X *satisfying (1.8) for all* $t, s \geq 0$, $f \in \mathrm{b}\mathcal{E}^\bullet$. *Assume that* X *is* $\mathcal{E}'$*-adapted to* $(\mathcal{G}_t)$. *Prove that (1.8) holds then for* $f \in \mathrm{b}\mathcal{E}'$. *(Hint: choose* $f_1 \leq f \leq f_2$ *with* $f_1, f_2 \in \mathcal{E}^\bullet$ *and* $f_2 - f_1$ *null for the measure* $g \to \mathbf{P}g(X_{t+s}) = \mathbf{P}\, P_s g(X_t)$ $(g \in \mathrm{b}\mathcal{E}^\bullet)$. *Remember that a conditional expectation is an equivalence class of random variables.)*

(1.14) EXERCISE. *Let* $\mathcal{E}^\bullet \subset \mathcal{E}'$ *be* σ*-algebras preserved by* (P_t), *and assume that* X *satisfies (1.8) for every* $f \in \mathrm{b}\mathcal{E}^\bullet$. *Prove that for all* $f \in \mathrm{b}\mathcal{E}'$,

$$\mathbf{P}\{f(X_{t+s}) \mid \mathcal{F}'_t\} = P_s f(X_t).$$

(Hint: by (1.13), one may reduce to the case $f \in \mathrm{b}\mathcal{E}^\bullet$. *Show using monotone classes that, for every* $H \in \mathrm{b}\mathcal{F}'_t$, *there exist* $H_1 \leq H \leq H_2$ *with* $H_1, H_2 \in \mathrm{b}\mathcal{F}^\bullet_t$ *and* $H_2 - H_1$ *null for the measure* $G \to \mathbf{P}G$ $(G \in \mathrm{b}\mathcal{F}^\bullet_t)$.*)*

(1.15) EXERCISE. *Let* $(X_t)_{t\geq 0}$ *be Markov with transition function* $(P_{s,t})$. *Suppose also that* $(P_{s,t})$ *satisfies the measurability condition*

$$(s,t,x) \to P_{s,t}(x,B)1_{\{s\leq t\}} \quad \text{is in } \mathcal{B}(\mathbf{R}) \otimes \mathcal{B}(\mathbf{R}) \otimes \mathcal{E}^\bullet \quad \forall\, B \in \mathcal{E}^\bullet.$$

Show that, with ϵ_u *denoting unit mass at* $u \in E$,

$$\tilde{P}_t\big((r,x),(ds,dy)\big) := \epsilon_{r+t}(ds)P_{r,r+t}(x,dy)$$

defines a Markov transition semigroup on $(\mathbf{R} \times E, \mathcal{B}(\mathbf{R}) \otimes \mathcal{E}^\bullet)$ *and the* **space-time** *process* $\tilde{X}_t := (t, X_t)$ *has the Markov property relative to* $(\Omega, \mathcal{G}, \mathcal{G}_t, \mathbf{P})$, *with transition semigroup* $(\tilde{P}_t)$.

(1.16) EXERCISE. *Verify, using the Kolmogorov existence theorem, that if* (P_t) *is a Markov transition semigroup on the Radon space* E, *and if* $(\mu_t)_{t>0}$ *is an arbitrary probability entrance law for* (P_t), *then there exists a unique probability measure* $\mathbf{P}$ *on the product space* $\Omega = E^{]0,\infty[}$ *with product* σ*-algebra* $\mathcal{G}$ *so that the coordinate maps* X_t *form a Markov process with transition function* (P_t) *and entrance law* (μ_t). *Formulate and check the temporally inhomogeneous version of this result.*

(1.17) EXERCISE. *Let* X_t *be a process with (not necessarily stationary) independent increments in* $\mathbf{R}^d$. *Show that* X *is Markovian and satisfies (1.6) for some transition function* $(P_{s,t})$.

2. The First Regularity Hypothesis

A stochastic process $(X_t)_{t \in I}$ defined on $(\Omega, \mathcal{G}, \mathbf{P})$ and having values in a topological space E is **right continuous** in case every sample path $t \to X_t(\omega)$ is a right continuous map of I into E. The following hypothesis is essentially the first of Meyer's **hypothèses droites**, which is to say in rough English translation, the regularity hypotheses for **right processes**. It is formulated as a condition on the transition semigroup rather than on the stochastic process.

(2.1) DEFINITION (HD1). *A Markov semigroup (P_t) on a Radon space E is said to satisfy HD1 if, given an arbitrary probability law μ on E, there exists a σ-algebra $\mathcal{E}^\bullet$ with $\mathcal{E} \subset \mathcal{E}^\bullet \subset \mathcal{E}^u$ and $P_t(\mathrm{b}\mathcal{E}^\bullet) \subset \mathrm{b}\mathcal{E}^\bullet$, and an E-valued right continuous $\mathcal{E}^\bullet$-process $(X_t)_{t \geq 0}$ on some filtered probability space $(\Omega, \mathcal{G}, \mathcal{G}_t, \mathbf{P})$ so that $X = (\Omega, \mathcal{G}, \mathcal{G}_t, \mathbf{P}, X_t)$ is (temporally homogeneous) Markov with transition semigroup (P_t) and initial law μ.*

It is implicit in (2.1) that X_t is $\mathcal{E}^\bullet$-adapted to $(\mathcal{G}_t)$ and that (1.8) is verified. Obviously, under the conditions of (2.1), one may replace $(\mathcal{G}_t)$ by $(\mathcal{F}_t^\bullet)$ without affecting anything. Notice that exercise (1.14) shows that if X satisfies all the conditions described in (2.1), and if (X_t) is $\mathcal{E}^u$-adapted to $(\mathcal{G}_t)$ then X also satisfies (2.1) with $\mathcal{E}^\bullet$ replaced everywhere with $\mathcal{E}^u$. That is, (2.1) does not really depend on the particular $\mathcal{E}^\bullet$.

The Markov property in the form (1.4) is rather awkward to manage, and in order to facilitate compututions, we shall make use of Dynkin's set-up for Markov processes, which brings in a family of measures governing a Markov process—one for each initial value $x \in E$—rather than one fixed measure $\mathbf{P}$. Loosely speaking, we shall work with a fixed collection of random variables X_t defined on some probability space, and a collection $\mathbf{P}^x$ of measures specified in such a way that $\mathbf{P}^x\{X_0 = x\}$, and, under every $\mathbf{P}^x$, X_t is Markov with semigroup (P_t). The $\mathbf{P}^x$ may be thought of as the conditional distributions for $\mathbf{P}$, given $X_0 = x$. (There is much of interest to be said in connection with Markov processes run under one distinguished measure $\mathbf{P}$, especially in case $\mathbf{P}$ is not necessarily a finite measure and the time-parameter set is the entire real line. We shall not go into such matters.)

(2.2) DEFINITION. *Let E be a Radon space, (P_t) a Markov semigroup on $(E, \mathcal{E}^\bullet)$ preserving $\mathcal{E}^\bullet$. The collection $X = (\Omega, \mathcal{G}, \mathcal{G}_t, X_t, \theta_t, \mathbf{P}^x)$ is a* **right continuous simple $\mathcal{E}^\bullet$-Markov process** *with state space E and transition semigroup (P_t) in case X satisfies conditions (2.3–5) below:*

(2.3) *$(\Omega, \mathcal{G}, \mathcal{G}_t)$ is a filtered measurable space, and X_t is an E-valued right continuous process $\mathcal{E}^\bullet$-adapted to $(\mathcal{G}_t)$;*

(2.4) *$(\theta_t)_{t\geq 0}$ is a collection of* **shift operators for** X, *viz, maps of* Ω *into itself satisfying, identically for* $t, s \geq 0$,

$$\theta_t \circ \theta_s = \theta_{t+s} \text{ and } X_t \circ \theta_s = X_{t+s};$$

(2.5) *For every* $x \in E$, $\mathbf{P}^x\{X_0 = x\} = 1$, *and the process* $(X_t)_{t\geq 0}$ *has the Markov property (1.8) with transition semigroup* (P_t) *relative to* $(\Omega, \mathcal{G}, \mathcal{G}_t, \mathbf{P}^x)$.

The condition $\mathbf{P}^x\{X_0 = x\} = 1$ in (2.5) is not always built into the definition of a simple Markov process. Markov processes enjoying this property are usually called **normal**. All Markov processes here will be assumed normal, unless explicit mention is made to the contrary. Note that (2.3) imposes the requirement that $\mathcal{G}_t \supset \mathcal{F}_t^\bullet$ for every $t \geq 0$.

(2.6) LEMMA. *Given a collection* $(\Omega, \mathcal{G}, \mathcal{G}_t, X_t, \theta_t, \mathbf{P}^x)$ *as above, then for every* $H \in \mathrm{b}\mathcal{F}^\bullet$, $x \to \mathbf{P}^x H$ *is* $\mathcal{E}^\bullet$*-measurable.*

PROOF: The formula (1.10) and the fact that the initial law for $\mathbf{P}^x$ is ϵ_x, the unit mass at x, shows that $\mathbf{P}^x\{f_1(X_{t_1}) \cdots f_n(X_{t_n})\}$ is in $\mathcal{E}^\bullet$ whenever $f_1, \ldots, f_n \in \mathrm{b}\mathcal{E}^\bullet$, $0 \leq t_1 \leq t_2 \leq \cdots \leq t_n$. An application of the MCT completes the proof.

Let $(\Omega, \mathcal{G}, \mathcal{G}_t, X_t, \theta_t, \mathbf{P}^x)$ satisfy (2.2). Given an arbitrary probability law μ on E, define $\mathbf{P}^\mu$ on $(\Omega, \mathcal{F}^\bullet)$ by $\mathbf{P}^\mu(H) := \int \mu(dx)\, \mathbf{P}^x(H)$, $H \in \mathrm{b}\mathcal{F}^\bullet$. It is a routine exercise to verify that (X_t) continues to have the Markov property relative to $(\Omega, \mathcal{G}, \mathcal{G}_t, \mathbf{P}^\mu)$, with transition function (P_t) and initial law μ.

One says that the collection $X = (\Omega, \mathcal{G}, \mathcal{G}_t, X_t, \theta_t, \mathbf{P}^x)$ satisfying (2.2) is a **realization** of the semigroup (P_t). The idea here is that (P_t) may be the prime object of study, and all information about (P_t) is embodied in X, which may then be studied by the methods of stochastic processes rather than those of functional analysis.

Under HD1, there is a realization of (P_t) which is in some respects canonical.

(2.7) THEOREM. *Let* (P_t) *be a Markov transition semigroup on* E *satisfying* HD1. *Then* (P_t) *has a right continuous realization* $(\Omega, \mathcal{G}, \mathcal{G}_t, X_t, \theta_t, \mathbf{P}^x)$.

PROOF: Let $x \in E$. According to (2.1), there is a process $(Y_t)_{t\geq 0}$ on a filtered probability space $(W, \mathcal{H}, \mathcal{H}_t, \mathbf{P})$ and a σ-algebra $\mathcal{E}^\bullet$ on E preserved by (P_t) such that Y is right continuous, $\mathcal{E}^\bullet$-adapted to $(\mathcal{H}_t)$, and satisfies (1.1) and (1.8) with initial law ϵ_x, so that $\mathbf{P}\{Y_0 = x\} = 1$. Let Ω denote the space of all right continuous maps of $\mathbf{R}^+$ into E. Let $X_t(\omega) := \omega(t)$ denote the coordinate variables on Ω and let $\mathcal{G}^\bullet := \sigma\{f(X_t) : t \geq 0, f \in \mathrm{b}\mathcal{E}^\bullet\}$,

$\mathcal{G}_t^\bullet := \sigma\{f(X_s) : 0 \le s \le t, f \in b\mathcal{E}^\bullet\}$. The map $\Phi : W \to \Omega$ defined by $\Phi(w) := \omega$, where $\omega(s) := Y_s(w)$ for all $s \ge 0$, is characterized by the formulas $Y_s \equiv X_s \circ \Phi$, $s \ge 0$. It follows trivially that $\Phi \in \mathcal{H}/\mathcal{G}^\bullet$. Let $\mathbf{P}^x$ be the image of $\mathbf{P}$ under the map Φ so that $\mathbf{P}^x F = \mathbf{P}\{F \circ \Phi\}$ for every $F \in b\mathcal{G}^\bullet$. This means that for $0 \le t_1 \le t_2 \le \cdots \le t_n$ and $f_1, \cdots, f_n \in b\mathcal{E}^\bullet$,

$$\mathbf{P}^x\{f_1(X_{t_1}) \cdots f_n(X_{t_n})\} = \mathbf{P}\{f_1(Y_{t_1}) \cdots f_n(Y_{t_n})\}.$$

The result now follows from (1.14), taking $\mathcal{G} := \mathcal{G}^u$.

(2.8) PROPOSITION. *Let $(\Omega, \mathcal{G}, \mathcal{G}_t, X_t, \theta_t, \mathbf{P}^x)$ be a right continuous simple $\mathcal{E}^\bullet$-Markov process as defined in (2.2). For every $F \in b\mathcal{F}^\bullet$ and all $t \ge 0$, $F \circ \theta_t \in b\mathcal{F}^\bullet$, and*

$$\mathbf{P}^\mu\{F \circ \theta_t \mid \mathcal{G}_t\} = \mathbf{P}^{X_t}\{F\}.$$

(More precisely, $\mathbf{P}^{X_t}\{F\}$ means $f(X_t)$, where $f(x) := \mathbf{P}^x F$. This notation, which looks rather confusing at first, will be used consistently.) Since $\mathcal{G}_t \supset \mathcal{F}_t^\bullet$, $\mathcal{G}_t$ may be replaced by $\mathcal{F}_t^\bullet$ in the above conditional expectation.

PROOF: The map $x \to \mathbf{P}^x F$ is in $\mathcal{E}^\bullet$ by (2.6). Thus, $\mathbf{P}^{X_t}\{F\} \in b\mathcal{F}_t^\bullet \subset \mathcal{G}_t$. For F of the form $F = f_1(X_{t_1}) \cdots f_n(X_{t_n})$ with $f_1, \ldots, f_n \in b\mathcal{E}^\bullet$ and $0 \le t_1 \le t_2 \le \cdots \le t_n$, the sought identity follows at once from the proof of (1.7), and the general case is completed by an appeal to the MCT.

The inexperienced reader is cautioned that HD1 (2.1) is a very substantial hypothesis whose verification seems possible only in very special situations, such as under strong analytic conditions on (P_t). See §9 for one such result. Another approach to HD1 is to start with a process known to satisfy HD1 and deform it in some probabilistic manner, verifying that the new process thus obtained continues to satisfy HD1. Examples of this type will be given in the course of the next few chapters, particularly in Chapter II. The examples will also illustrate why it is desirable to avoid hypotheses that would mandate that (P_t) preserve Borel functions. A third avenue to HD1 is to take a Markov process not necessarily satisfying HD1 and regularize it in some manner so that the new semigroup satisfies HD1. Part of the theory of Ray-Knight completions deals with this matter.

It should be mentioned that the definitions given in this section differ slightly from those in the earlier work of Blumenthal-Getoor [**BG68**]. The requirements imposed on $(\mathcal{G}_t)$ in this section involve only that the conditional expectations of the $H \in b\mathcal{F}^u$ relative to $\mathcal{G}_t$ be the same as the conditional expectations relative to $\mathcal{F}_t^u$. If one wishes to condition $\mathcal{G}$-measurable random variables, further measurability conditions must be imposed on $(\mathcal{G}_t)$. We take up this issue in §6.