I. I. Gihman A.V. Skorohod

The Theory of Stochastic Processes II

Translated from the Russian
by S. Kotz

I. I. Gihman A.V. Skorohod

The Theory of Stochastic Processes II

Translated from the Russian
by S. Kotz

Reprint of the First Edition 1975

Springer-Verlag
Berlin Heidelberg New York Tokyo 1983

Iosif Il'ich Gihman
Academy of Sciences of the Ukrainian SSR, Institute of Applied Mathematics and Mechanics, Donetsk

Anatolii Vladimirovich Skorohod
Academy of Sciences of the Ukrainian SSR, Institute of Mathematics, Kiev

Translator:

Samuel Kotz
Department of Mathematics, Temple University, Philadelphia

Title of the Russian Original Edition: Teoriya sluchainyh protsessov, Vol. II. Publisher: Nauka, Moscow 1973

AMS Subject Classification (1970): 60-02: 60 Jxx, 60 Kxx

ISBN 3-540-07247-0 Springer-Verlag Berlin Heidelberg New York Tokyo
ISBN 0-387-07247-0 Springer-Verlag New York Heidelberg Berlin Tokyo

Library of Congress Cataloging in Publication Data. Gihman, Iosif Il'ich. The theory of stochastic processes. (Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen mit besonderer Berücksichtigung der Anwendungsgebiete; Bd. 210, 218) DM 129.00 (v.1). Translation of Teoriia sluchaïnykh protsessov. Bibliography: v. 1, p. ; v. 2, p. Includes indexes. 1. Stochastic processes. I. Skorokhod, Anatolii Vladimirovich, joint author. II. Title. III. Series: Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen; Bd. 210 [etc.]. QA274.G5513. 519.2. 74-2552. ISBN 3-540-06573-3 (v. 1).

Printed in Germany
Printing: Weihert-Druck, Darmstadt. Bookbinding: Konrad Triltsch, Würzburg.
2141/3140-543210

Preface

The second volume of "The Theory of Stochastic Processes" is devoted mainly to Markov processes. The first and second chapters contain the general theory of Markov and homogeneous Markov processes. In the succeeding Chapters the following important classes of Markov processes are discussed: jump processes, semi-Markov processes, branching processes and processes with independent increments. A substantial number of the results in this volume are appearing in (non-periodical) literature for the first time.

The portion of material dealing with the theory of Markov processes not included in this volume such as diffusion, conditional Markov processes and some other topics are scheduled for incorporation in the third volume. The theory of stochastic differential equations will also appear in this volume.

I. I. Gihman and A. V. Skorohod

Table of Contents

Introduction

Markov processes play a special role in the theory of random processes. This is due to the fact that the definition of a Markov process utilizes to a great extent the notions which distinguish probability theory—within the encompassing framework of general measure theory—as a separate independent science. The probabilistic intuition based on the notion of independence manifests itself most prominently in the theory of Markov processes.

Another important feature of the theory of Markov processes is the fact that this theory allows us to describe all the finite-dimensional distributions of the process in terms of a small number of constructively defined characteristics and thus allows us to evaluate the distribution of various functionals of the process.

Note that for other general classes of processes (with the exception of the class of Gaussian processes) we are able in general to define only those events whose probability is either 0 or 1.

Finally, the most important feature of a Markov process is the evolutionary character of its development: the state of the process at present completely determines its probabilistic behavior in the future. This allows us in many cases, by appropriately extending the phase space of a process, to transform it into a Markov process. On the other hand, the evolutionary character of development of the process permits us to derive recurrence relationships (in the discrete case) or evolutionary equations (in the continuous case) for the determination of the probabilistic characteristics of the process.

Current investigations in the theory of random processes are to a great extent devoted to the study of various classes of Markov processes.

The notion of a Markov process originated as a generalization of a sequence of trials connected into a chain which was studied by A. A. Markov. Unlike Bernoulli's scheme, Markov studied the case in which the probability of the occurrence of an event in a subsequent trial depends on the outcome of the previous trial. The general concept of a Markov process was given by A. N. Kolmogorov in his paper "Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung" (1931), Math. Ann. **104**, 415—458. In this paper, Kolmogorov studied stochastically determined systems, i. e. systems whose probabilistic behavior in the future is completely determined by the state of the system at the present instant of time. These are described by the function $P(s,x,t,B)$ namely by the probability that the system will be at time t ($t>s$) at one of the states of the set $B\in\mathfrak{B}$, where $\mathfrak{B}$ is the σ-algebra of the subsets of the phase space $\mathscr{X}$, given that at time s the system is in the state x. It follows from the formula of total probability and stochastic determinancy that the function $P(s,x,t,B)$—called the transition probability—satisfies the following relation

$$P(s,x,t,B)=\int_{\mathscr{X}} P(s,x,u,dy)\,P(u,y,t,B) \qquad (s<u<t), \tag{1}$$

where $\mathscr{X}$ is the phase space of the system. This relation is called the Chapman-Kolmogorov equation.

One of the initial problems which arises in this connection is to describe the various classes of solutions of equation (1).

In the case in which the phase space $\mathscr{X}$ consists of at most a countable set of points $x_1, x_2,\ldots$ the transition probability is determined by the collection of functions $p_{ij}(s,t)=P(s,x_i,t,\{x_j\})$, where $\{x_j\}$ is the singleton containing x_j. A. N. Kolmogorov has shown that under certain assumptions the functions $p_{ij}(s,t)$ satisfy the following systems of differential equations

$$\frac{dp_{ij}(s,t)}{ds}=\sum_k a_{ik}(s)p_{kj}(s,t), \qquad \frac{dp_{ij}(s,t)}{dt}=\sum_k p_{ik}(s,t)a_{kj}(t).$$

Another important class of processes investigated by A. N. Kolmogorov is the class of processes with finite-dimensional Euclidean phase space whose transition probabilities possess density functions $p(s,x,t,y)$. By imposing on the function $p(s,x,t,y)$ certain restrictions which are in agreement with the intuitive notion of continuous movement of a system, A. N. Kolmogorov obtained the following equations in partial derivatives for the function $p(s,x,t,y)$:

$$\frac{\partial p(s,x,t,y)}{\partial s}+\sum_k a_k(s,x)\frac{\partial p(s,x,t,y)}{\partial x_k}+\frac{1}{2}\sum_{i,k} b_{ik}(s,x)\frac{\partial^2 p(s,x,t,y)}{\partial x_i\partial x_k}=0,$$

$$\frac{\partial p(s,x,t,y)}{\partial t}+\sum_k \frac{\partial}{\partial y_k}(a_k(t,y)p(s,x,t,y))-\frac{1}{2}\sum_{i,k}\frac{\partial^2}{\partial y_i\partial y_k}(b_{ik}(t,y)p(s,x,t,y))=0.$$

A. N. Kolmogorov also obtained equations for more general processes in Euclidean space, whose states may vary not only continuously but also stepwise. In all these cases, A. N. Kolmogorov succeeded in reducing the non-linear functional equation (1) to the more common linear differential equations of the evolutionary type (in the case of discontinuous processes integro-differential equations appear). The processes are, moreover, characterized by the coefficients of the corresponding equations which have a simple probabilistic meaning and which are infinitesimal characteristics of the process.

I. G. Petrovskiĭ and A. Ya. Khinchin utilized continuous Markov processes for construction of the probabilistic model of diffusion. Such processes were subsequently referred to as diffusion processes. Moreover, it was discovered that the infinitesimal characteristics of the process introduced by A. N. Kolmogorov allow us to determine not only the transition probabilities but also to evaluate the distributions of various functionals of the processes (such as the time required to reach a certain region and the distribution of the values of the process at the moment the boundary of the region has been reached).

A. N. Kolmogorov's ideas constituted the basis of the mathematical theory of Markov processes and provided the direction for further investigations such as the description of the infinitesimal characteristics of the process and construction of transition probabilities corresponding to the given infinitesimal characteristics.

However, these infinitesimal characteristics do not always exist, and in the case when they do exist they do not necessarily determine the process uniquely. In this connection, the idea proposed by W. Feller of utilizing the theory of semigroups of operators associated with the transition probabilities proved to be fruitful. The theory of semigroups of operators is particularly applicable to processes which are homogeneous in time, i. e. processes in which the transition probability $P(s,x,t,B)$ depends only on the difference of the arguments $t-s$, i. e. $P(s,x,t,B)=P(t-s,x,B)$. This restriction is not essential, since an arbitrary Markov process can easily be converted into a homogeneous one by an appropriate modification of the phase space.

Let $\mathscr{F}_{\mathfrak{X}}$ be the space of all $\mathfrak{B}$-measurable real-valued bounded functions. The family of operators $\mathbf{T}_t$, determined by the relation

$$\mathbf{T}_t f(x) = \int f(y) P(t,x,dy), \qquad f \in \mathscr{F}_{\mathfrak{X}}, t>0,$$

is called the semigroup of operators associated with the transition probability $P(t,x,B)$. This semi-group completely determines the transition probabilities. On the other hand, this group is in many cases uniquely determined by its *infinitesimal operator* $\mathbf{A}$, where

$$\mathbf{A} f(x) = \lim_{t \downarrow 0} \frac{\mathbf{T}_t f(x) - f(x)}{t},$$

provided this limit exists for all $x \in \mathfrak{X}$. W. Feller suggested that the infinitesimal operator $\mathbf{A}$ be considered as an infinitesimal characteristic of the process. He gave a description of all the diffusion processes defined on a bounded interval by means of semi-group methods. In this case, the infinitesimal operator of the process is of the form

$$\mathbf{A} f = a \frac{df}{dx} + \frac{1}{2} b \frac{d^2 f}{dx^2},$$

where a and b are the coefficients appearing in A. N. Kolmogorov's equation and the domain of definition of the operator $\mathbf{A}$ depends on the behavior of the process on the boundary points and is selected from the set of all twice differentiable functions by means of certain additional (boundary) conditions. Various additional conditions were described by W. Feller and A. D. Ventzel.

E. B. Dynkin improved on Feller's purely analytical method by introducing into consideration the trajectories of the process. He introduced the general definition of a Markov process which is commonly used nowadays and studied in detail the strong Markovian property of the process (i. e. the preservation of the Markovian property of a process with respect to a random moment of time which is independent of the future). E. B. Dynkin determined the characteristic operator of the process $\mathfrak{U}$ for a strong Markov process. If $\mathscr{D}_{\mathfrak{U}}$ is the domain of definition of operator $\mathfrak{U}$, then under some very natural assumptions $\mathscr{D}_{\mathrm{A}} \subset \mathscr{D}_{\mathfrak{U}}$ and $\mathrm{A}f = \mathfrak{U}f$ on $\mathscr{D}_{\mathrm{A}}$.

The advantage of the characteristic operator over the infinitesimal one is due to the fact that for computing the characteristic operator it is sufficient to know the behavior of the trajectory up to the instant of time at which the trajectory leaves an arbitrary small neighborhood of the initial point (including this instant). Therefore the characteristic operator is easily computable in many cases (for example in the case of jump processes and continuous processes on the line). Therefore, if the characteristic operator is known, then in order to determine the infinitesimal operator it is sufficient to find the domain of its definition $\mathscr{D}_{\mathrm{A}}$ which is the contraction of $\mathscr{D}_{\mathfrak{U}}$ and which is consequently selected from $\mathscr{D}_{\mathfrak{U}}$ by means of certain additional (boundary) conditions. As we can see, the situation in the general case is similar to that which exists in the case of a diffusion process on the interval.

The general problem of describing the domain of definition of an infinitesimal operator corresponding to a given characteristic operator remains as yet unsolved. It was discovered that the solution of this problem is connected with the study of harmonic and excessive functions of the process which was investigated by G. A. Hunt. On the other hand, the notions of multiplicative and additive functionals—introduced by E. B. Dynkin—which are also connected with the notion of an excessive function are of great importance for the construction of various transformations of the processes (such transformations may significantly simplify the study of the processes). The study of excessive functions and additive and multiplicative functionals of a process constitutes at present a substantial part of the general theory of Markov processes. This theory was presented for the first time in rather complete form in E. B. Dynkin's monographs "Basic notions of the theory of Markov processes" and "Markov processes"*.

Simultaneously with the development of the general theory, a detailed investigation was undertaken of various particular classes of Markov processes with more specific properties. We shall mention a few of the most important of these classes.

Processes with independent increments, i. e. processes $\xi(t)$ for which the variables

$$\xi(0),\ \xi(t_1)-\xi(0), \ldots, \xi(t_n)-\xi(t_{n-1}), \qquad 0<t_1<t_2<\cdots<t_n,$$

are independent for any n and $t_1, t_2, \ldots, t_n$ constitute an important class of Markov processes. These processes can be viewed as continuous-time random walks. They were originally utilized for the description of Brownian motions.

* English translation 1965

The general processes of this class are utilized as evolution models of arbitrary systems in a (spatial) homogeneous random medium. Stochastically continuous processes with independent increments were comprehensively described in the the works of B. de Finetti, A. N. Kolmogorov and P. Levy. From the point of view of the general theory, processes with independent increments are spatially homogeneous Markov processes.

In addition to an analytic description of distributions of the processes with independent increments, the properties of the sample functions (paths) of these processes were also studied. P. Lévy has shown that stochastically continuous processes with independent increments have no discontinuities of the second kind. A. Ya. Khinchin studied the local growth of processes with independent increments and in particular, he derived the well known law of iterated logarithms.

To describe the size of biological populations F. Galton and H. W. Watson suggested a certain random process. A. N. Kolmogorov and N. A. Dmitriev introduced a special class of Markov processes with a countable number of states which they called branching processes. Subsequently, these processes had numerous applications in biology and physics for a description of systems with birth, extinction and transformation of species (or particles). The state of a branching process in a given instant of time is determined by the number of particles of each type which are present in the system (for example the number of species of each sex in a biological population). Each particle is subject to a transformation as a result of which the particle either disappears, or is replaced by other particles of arbitrary types and in arbitrary numbers. If the further evolution of a particle present in the system is independent of its age and of the evolution of other particles, the process becomes a Markov branching process. The infinitesimal characteristics of the process are the extinction probabilities or the probability of transformation of a particle of each type—during an infinitesimal period of time—into a collection of (other) particles. By means of these characteristics one can write differential equations for the generating functions of the number of particles present in the system.

It is of interest to study the asymptotic behavior of the number of particles in the system as $t \to \infty$ and in particular the calculation of the probabilities of disappearance of all particles from the system (its degeneration) as well as the probabilities that their number increases indefinitely (explosion).

For a more accurate description of real-world systems, it is natural to consider branching processes in which the transformation probabilities depend on the age of the particles. This can be achieved by introducing a phase space in which the position of the particle is subject to change and moreover, the probabilities of transformation of a particle may depend on its position in the phase space. In this manner one arrives at the general definition of a Markov branching process whose state is determined by the number of particles of each type as well as by their position in a certain phase space, and moreover, the movement of each particle in the phase space is described by a Markov process with transition probabilities depending only on the type of the particle.

Another interesting generalization of a branching process is obtained in the case when the type of particle is characterized not by number but mass assuming that the latter varies continuously. The development of the theory of general branching processes is a recent one and only initial results have been obtained related to the determination of infinitesimal characteristics of the process and construction of a process on the basis of these characteristics.

A large number of papers including those of an applied nature, are devoted to queueing theory problems. A queueing system is characterized by an incoming stream of customers (demands), the number of servers and the service time of a customer by each one of the servers. If the stream of customers is Poisson and the service times are exponentially distributed, then such a queueing system is described by a Markov process with a countable number of states. To describe a more general queueing system a special class of Markov processes is utilized. A semi-Markov process has a countable number of states and the transition probabilities depend on the duration of stay in a given state. If the state of the system is described by the pair, the state of the semi-Markov process and the duration of stay in this state, the system becomes Markovian. The basic problems of the theory of semi-Markov processes which follow from the nature of their applications are the calculation of transition probabilities, stationary distributions in the phase space and determination of conditions for applicability of the ergodic theorem.

A natural generalization of semi-Markov processes is the general process with a discrete chance interference. This is a process which is Markovian between two consecutive interferences of the chance. The effect of the interference is that the process suddenly changes its state in the phase space (in a non-Markovian manner), while the pair—the state of the process and the duration of time from the instant of the last chance interference—is a Markov process.

The present volume is completely devoted to the theory of Markov processes. The general theory as well as the most important classes of Markov processes are studied in this volume. Diffusion processes, however, will be investigated in detail in the third volume.

Chapter I deals with the general definitions of a Markov process, of a Markov random function, and the strict Markov property of the process. In this chapter the criterion for the strict Markov property is established and multiplicative functionals and subprocesses of a Markov process are studied as well as the properties of sample functions. Markov processes in the wide sense are studied before the general theory is presented. (The theory of wide-sense Markov processes does not involve the notion of sample function of the process.) Kolmogorov's equations for various classes of processes are derived as well.

Chapter II is devoted to homogeneous Markov processes. Here the semigroup corresponding to a Markov process is introduced, resolvent and generating operators of a process are discussed and the Hille-Yosida theorem on the existence of a semigroup with a given generating operator is proved. A substantial part of the second chapter is devoted to the study of Feller processes in compact and locally compact spaces. The conditions are derived under which a given operator is a characteristic operator of a Feller process in a locally compact space, and all the processes with a given characteristic operator are described. Additive functionals of a Markov process are studied. A description of all continuous additive functionals of Feller processes is obtained and the random substitution of time is considered.

Chapter III is devoted to jump processes. A general definition of these processes is given and their structure is investigated. Also studied are homogeneous processes with a countable number of states, semi-Markov processes, processes with a semi-Markov chance interference and general processes with a discrete chance interference. Characteristic and generating operators for these processes are obtained and the ergodic theory is proved. Moreover, Markov processes with a discrete component are investigated and their characteristic operators are obtained.

Processes with independent increments are studied in Chapter IV. Here the properties of sample functions of these processes are investigated as well as their local growth and the growth at infinity. For a one-dimensional process the distributions of the basic functionals of the process are obtained, such as the first passage time of a certain level, the size of the jump through this level as well as the joint distributions of the supremum, infimum and the value of the process. A certain class of non-negative continuous additive functionals of a process is described.

Chapter V deals wit Markov branching processes. Branching processes with a finite number of types of particles, processes with a continuous mass and general Markov branching processes are studied.

In many cases the main text does not contain references to the original papers. To a certain extent this is done in the Remarks at the end of the book In the bibliography the authors tried to list all the basic papers on the theory of Markov processes dealing with problems discussed in this book.

Chapter I

Basic Definitions and Properties of Markov Processes

§ 1. Wide-Sense Markov Processes

Definition. The idea of a process "without an aftereffect" is the underlying characteristic of a Markov process. Consider a system (or a particle) which may be found in various states. The possible states of the system form a set $\mathscr{X}$ called the phase space of the system. Assume that the system changes in time. The state of the system at time t is denoted by x_t. If $x_t \in B$, where $B \subset \mathscr{X}$, we say that the system at time t is situated in the set B. Assume that the evolution of the system is of a stochastic nature, i.e. the state of the system at time t is, in general, not uniquely determined by the states of the system at the times preceding time s, where $s<t$, but is random and is described by certain probabilistic laws. Denote by $P(s,x,t,B)$ the probability of the event $x_t \in B$ $(s<t)$ given that $x_s = x$.

The function $P(s,x,t,B)$ is called the *transition probability* of the given system. A system is termed *without an aftereffect* if the probability of its being situated at time t in the set B, under the condition that the movement of the system up to time s $(s<t)$ is completely known, equals $P(s,x,t,B)$ and thus depends only on the state of the system at time s. A complete formal definition will be given in succeeding sections. Here we present a simple definition of this concept which is sufficient for a number of applications. Denote by $P(s,x,u,y,t,B)$ the conditional probability of the event $x_t \in B$ under the assumptions $x_s = x$, $x_u = y$ $(s \leqslant u \leqslant t)$. From the general properties of conditional expectations we have

$$P(s,x,t,B) = \int_{\mathscr{X}} P(s,x,u,y,t,B)\,P(s,x,u,dy). \tag{1}$$

For a system without an aftereffect $P(s,x,u,y,t,B) = P(u,y,t,B)$. In this case equality (1) becomes

$$P(s,x,t,B) = \int_{\mathscr{X}} P(u,y,t,B)\,P(s,x,u,dy) \quad (s<u<t). \tag{2}$$

Equation (2) is called the *Chapman-Kolmogorov equation*. It may serve as the basis for a definition of a process without aftereffect; such a process will be referred to in what follows as a Markov process.

Let $\{\mathscr{X}, \mathfrak{B}\}$ be a measurable space. The function $P(x,B)$, $x \in \mathscr{X}$, $B \in \mathfrak{B}$, satisfying the conditions

a) For a fixed x $P(x,B)$ is a measure on $\mathfrak{B}$ and $P(x,\mathscr{X})\leqslant 1$,

b) for a fixed B $P(x,B)$ is a $\mathfrak{B}$-measurable function of x,

is called a *semi-stochastic kernel*. If $P(x,\mathscr{X})=1$ for all $x\in\mathscr{X}$, then $P(x,B)$ is called a *stochastic kernel*.

This terminology will be used also in a somewhat more general case, when the argument x of the function $P(x,B)$ takes on values in a measurable space $\{\mathscr{X}_0,\mathfrak{B}_0\}$ different from $\{\mathscr{X},\mathfrak{B}\}$.

Let $\mathscr{I}$ be a finite or infinite half-interval. A family of semi-stochastic (stochastic) kernels

$$\{P_{st}(x,B)=P(s,x,t,B),\ s<t,\ (s,t)\in\mathscr{I}\times\mathscr{I}\}$$

satisfying Chapman-Kolmogorov equations is called a *Markov family of semi-stochastic (stochastic) kernels*.

Definition 1. *A wide-sense Markov process* is a collection of the following objects:

a) a measurable space $\{\mathscr{X},\mathfrak{B}\}$,

b) an interval $\mathscr{I}$ (half-interval, segment) on the real line,

c) a Markov family of stochastic kernels

$$\{P_{st}(x,B),\ s<t,\ (s,t)\in\mathscr{I}\times\mathscr{I}\}.$$

The family of kernels $P_{st}(x,B)=P(s,x,t,B)$ is called *the transition probability of a Markov process*, the space $\{\mathscr{X},\mathfrak{B}\}$ is called *the phase space of the system*, the points of $\mathscr{I}$ are interpreted as the instants of time, and the quantity $P_{st}(x,B)=P(s,x,t,B)$ as the conditional probability that the system at time t be situated in B given that at time s it was situated at point x of the phase space $(s<t)$.

Henceforth, we shall assume that the kernel $P_{st}(x,B)$ is defined also for $s=t$. Namely, it is natural to define

$$P_{tt}(x,B)=\chi(B,x),$$

where $\chi(B,x)$ denotes the indicator of the set B: $\chi(B,x)=1$ if $x\in B$ and $\chi(B,x)=0$ if $x\notin B$.

Clearly, equality (2) is satisfied with $u=s$ or $u=t$ for such a definition of the kernel $P_{tt}(x,B)$.

The Chapman-Kolmogorov equation shows that the kernel $P_{st}(x,B)$ is the convolution of the kernels $P_{su}(x,B)$ and $P_{ut}(x,B)$ $(s\leqslant u\leqslant t)$. The definition of a convolution of kernels was given in Volume I.

Cut-off Markov Processes. In what follows, we shall consider Markov processes defined by stochastic as well as semi-stochastic kernels. In the latter case the relation $P(s,x,t,\mathscr{X})<1$ can be naturally interpreted as the possibility of the disappearance of the system from the phase space. Moreover, if $x_s=x$, then the probability of disappearance $\tilde{p}(s,x,t)$ of the system during the time interval $(s,t]$ is set to be equal to $1-P(s,x,t,\mathscr{X})$. Note that it follows from the

Chapman-Kolmogorov equation that $\tilde{p}(s,x,t)$ is a non-decreasing function of t. Indeed for $h>0$

$$P(s,x,t+h,\mathscr{X})=\int P(s,x,t,dy)P(t,y,t+h,\mathscr{X})\leqslant\int P(s,x,t,dy)=P(s,x,t,\mathscr{X}).$$

The possibility of such an interpretation of the relation $P(s,x,t,\mathscr{X})<1$ is based on the following argument. The disappearance of the system from the phase space is interpreted as its reaching a certain special state $\mathfrak{v}$, $\mathfrak{v}\notin\mathscr{X}$. We extend the space $\mathscr{X}$ by adjoining to it the new point $\mathfrak{v}$ and denote the extended space by $\mathscr{X}_\mathfrak{v}$. We introduce in $\mathscr{X}_\mathfrak{v}$ the σ-algebra $\mathfrak{B}_\mathfrak{v}$ consisting of all possible sets $B\in\mathfrak{B}$ and the sets of the form $B\cup\{\mathfrak{v}\}$, $B\in\mathfrak{B}$. We complete the definition of the function $P(s,x,t,B)$ for $x\in\mathscr{X}_\mathfrak{v}$ and $B\in\mathfrak{B}_\mathfrak{v}$, by setting

$$\tilde{P}(s,x,t,B)=P(s,x,t,B\setminus\{\mathfrak{v}\})+\chi(B,\mathfrak{v})\tilde{p}(s,x,t)\quad\text{for}\quad x\neq\mathfrak{v}$$

and

$$\tilde{P}(s,\mathfrak{v},t,B)=\chi(B,\mathfrak{v}).$$

Lemma 1. *The family of stochastic kernels* $\tilde{P}(s,x,t,B)$ $(s\in\mathscr{I},t\in\mathscr{I},s<t)$, $x\in\mathscr{X}_\mathfrak{v}$, $B\in\mathfrak{B}_\mathfrak{v}$, *is Markovian.*

To prove the lemma it is sufficient to verify that $\tilde{P}$ satisfies the Chapman-Kolmogorov equation. We have

$$\begin{aligned}\tilde{P}(s,\mathfrak{v},t,B)&=\tilde{P}(s,\mathfrak{v},u,\{\mathfrak{v}\})\tilde{P}(u,\mathfrak{v},t,B)\\&=\int_{\mathscr{X}_\mathfrak{v}}\tilde{P}(s,\mathfrak{v},u,dx)\tilde{P}(u,x,t,B),\qquad B\in\mathfrak{B}_\mathfrak{v},s<u<t.\end{aligned}$$

If $B\in\mathfrak{B}$, $x\in\mathscr{X}$, then

$$\tilde{P}(s,x,t,B)=P(s,x,t,B)=\int_{\mathscr{X}}P(s,x,u,dy)P(u,y,t,B)=\int_{\mathscr{X}_\mathfrak{v}}\tilde{P}(s,x,u,dy)\tilde{P}(u,y,t,B).$$

Now let $x\in\mathscr{X}$, $B_0\in\mathfrak{B}$, $B=B_0\cup\{\mathfrak{v}\}$. Then

$$\begin{aligned}\tilde{P}(s,x,t,B)&=P(s,x,t,B_0)+P(s,x,t,\{\mathfrak{v}\})\\&=\int_{\mathscr{X}_\mathfrak{v}}\tilde{P}(s,x,u,dy)\tilde{P}(u,y,t,B_0)+\tilde{P}(s,x,t,\{\mathfrak{v}\}).\end{aligned}$$

But

$$\begin{aligned}\tilde{P}(s,x,t,\{\mathfrak{v}\})&=1-P(s,x,t,\mathscr{X})=1-\int_{\mathscr{X}}P(s,x,u,dy)P(u,y,t,\mathscr{X})\\&=1-\int_{\mathscr{X}_\mathfrak{v}}P(s,x,u,dy)P(u,y,t,\mathscr{X})=\int_{\mathscr{X}_\mathfrak{v}}\tilde{P}(s,x,u,dy)\tilde{P}(u,y,t,\{\mathfrak{v}\}).\end{aligned}$$

Therefore

$$\begin{aligned}P(s,x,t,B_0\cup\{\mathfrak{v}\})&=\int_{\mathscr{X}_\mathfrak{v}}\tilde{P}(s,x,u,dy)\tilde{P}(u,y,t,B_0)+\int_{\mathscr{X}_\mathfrak{v}}\tilde{P}(s,x,u,dy)\tilde{P}(u,y,t,\{\mathfrak{v}\})\\&=\int_{\mathscr{X}_\mathfrak{v}}\tilde{P}(s,x,u,dy)\tilde{P}(u,y,t,B_0\cup\{\mathfrak{v}\}).\end{aligned}$$

The lemma is thus proved. □

Although, as it was just shown, the case of a Markov family of semi-stochastic kernels can be rather easily reduced to the case of a Markov family of stochastic kernels, nevertheless the addition of a new point to the phase space changes its topological structure. It is therefore meaningful to distinguish between these two cases.

Definition 2. A system of objects consisting of the phase space $\{\mathscr{X},\mathfrak{B}\}$, a time half-interval $\mathscr{I}$ and a Markov family of semi-stochastic kernels $\{P_{st}(x,B), s<t, (s,t)\in\mathscr{I}\times\mathscr{I}\}$ in $\{\mathscr{X},\mathfrak{B}\}$ is called a *wide-sense cut-off Markov process*. (The word "cut-off" indicates the possibility of the disappearance of the system from the phase space (the cut-off of the process).) If, however, $P_{st}(x,\mathscr{X})\equiv 1$, such a process is called *non-cut-off*.

Input distribution of a Markov process. In this section only wide-sense Markov processes are studied. For this reason the words "wide-sense" will often be omitted. We return to the definition of a Markov process given above (cut-off or non-cut-off), and note that these definitions in general do not stipulate that the probability of the event $\{x_t\in B\}$ be defined.

However, if we define on $\mathfrak{B}$ (or on $\mathfrak{B}_0$) a probability measure q_s and assume that $\mathsf{P}\{x_s\in B\}=q_s(B)$, then for $t>s$ it follows from the general formulas of probability theory that the probability $q_t(B)$ of the event $\{x_t\in B\}$ should be defined as

$$q_t(B)=\mathsf{P}\{x_t\in B\}=\int P(s,x,t,B)q_s(dx). \tag{3}$$

This definition is meaningful in the following sense:

Compute the quantities q_u and q_t $(s<u<t)$ using formula (3), and then set $s=u$, $q_s=q_u$ in (3) and compute q_t again. The measures q_t, computed using different methods, will coincide. More precisely, if the operation which constructs q_t by means of formula (3) for given t, s and $q_s=q$ is denoted by $q_t=F_t(s,q)$ then for any $u\in(s,t)$

$$q_t=F_t(u,q_u)=F_t(u,F_u(s,q)). \tag{4}$$

To prove this assertion the following simple lemma will be needed.

Let $\{\mathscr{X}_i,\mathfrak{B}_i\}$ $(i=1,2)$ be two measurable spaces, let m be a measure on $\mathfrak{B}_1$, and let $q(x,B)$ $(x\in\mathscr{X}_1, B\in\mathfrak{B}_2)$ be a semi-stochastic kernel.

Set

$$q(B)=\int_{\mathscr{X}_1} q(x,B)m(dx).$$

Clearly $q(B)$ is a measure on $\mathfrak{B}_2$ and moreover $q(B)\leqslant m(\mathscr{X}_1)$.

Lemma 2. *If the measure m is finite, we have for an arbitrary bounded and $\mathfrak{B}_2$-measurable function $f(y)$*

$$\int_{\mathscr{X}_1} m(dx)\int_{\mathscr{X}_2} f(y)q(x,dy)=\int_{\mathscr{X}_2} f(y)q(dy). \tag{5}$$

Proof. Denote by K the class of functions for which (5) holds. The class K contains the indicators of sets in $\mathfrak{B}_2$ and being linear, it thus contains all the