Springer Optimization and Its Applications 80

Michał Kisielewicz

Stochastic Differential Inclusions and Applications

Springer

Springer Optimization and Its Applications

VOLUME 80

Aims and Scope
Optimization has been expanding in all directions at an astonishing rate during the last few decades. New algorithmic and theoretical techniques have been developed, the diffusion into other disciplines has proceeded at a rapid pace, and our knowledge of all aspects of the field has grown even more profound. At the same time, one of the most striking trends in optimization is the constantly increasing emphasis on the interdisciplinary nature of the field. Optimization has been a basic tool in all areas of applied mathematics, engineering, medicine, economics, and other sciences.

The series *Springer Optimization and Its Applications* publishes undergraduate and graduate textbooks, monographs and state-of-the-art expository work that focus on algorithms for solving optimization problems and also study applications involving such problems. Some of the topics covered include nonlinear optimization (convex and nonconvex), network flow problems, stochastic optimization, optimal control, discrete optimization, multi-objective programming, description of software packages, approximation techniques and heuristic approaches.

For further volumes:
http://www.springer.com/series/7393

Michał Kisielewicz

Stochastic Differential Inclusions and Applications

Springer

Michał Kisielewicz
Faculty of Mathematics
University of Zielona Góra
Zielona Góra
Poland

ISSN 1931-6828
ISBN 978-1-4614-6755-7 ISBN 978-1-4614-6756-4
DOI 10.1007/978-1-4614-6756-4
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2013934966

Mathematics Subject Classification (2010): 34K09, 39A50, 49J21, 65M75

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

To my wife with love
and gratitude for support

Contents

Preface

There has been a great deal of interest in optimal control systems described by stochastic and partial differential equations. These optimal control problems lead to stochastic and partial differential inclusions. The aim of this book is to present a unified theory of stochastic differential inclusions written in integral form with both types of stochastic set-valued integrals defined as subsets of the space $\mathbb{L}^2(\Omega, \mathbb{R}^n)$ and as multifunctions with closed values in the space $\mathbb{R}^n$. Such defined inclusions are therefore divided into two types: stochastic functional inclusions ($SFI(F, G)$) and stochastic differential inclusions ($SDI(F, G)$), respectively. The main results of the book deal with properties of solution sets of stochastic functional inclusions and some of their applications in stochastic optimal control theory and in the theory of partial differential inclusions. In particular, apart from the existence of weak solutions for initial value problems of stochastic functional inclusions, the existence of their strong and weak viable solutions is also investigated. An important role in applications is played by theorems on weak compactness of solution sets of weak and viable weak solutions for the above initial value problems. As a result of these properties, some optimal control problems for dynamical systems described by stochastic and partial differential inclusions are obtained. Let us remark that for a given pair (F, G) of multifunctions, the sets $\mathcal{X}(F, G)$ and $\mathcal{S}(F, G)$ of all weak solutions of $SFI(F, G)$ and $SDI(F, G)$, respectively, are defined as families of systems $(\mathcal{P}_{\mathbb{F}}, x, B)$ consisting of a filtered probability space $\mathcal{P}_{\mathbb{F}}$, a continuous process $x = (x_t)_{t\geq 0}$, and an $\mathbb{F}$-Brownian motion $B = (B_t)_{t\geq}$ satisfying these inclusions. Immediately from the definitions of $SFI(F, G)$ and $SDI(F, G)$, it follows that $\mathcal{X}(F, G) \subset \mathcal{S}(F, G)$. It is natural to extend the results of this book to the set $\mathcal{S}(F, G)$ and consider weak solutions with x a càdlàg process instead of a continuous one. These problems are quite complicated and need new methods. Therefore, in this book, they are left as open problems.

The first papers dealing with stochastic functional inclusions written in integral form are due to Hiai [38] and Kisielewicz [50–56,58,60–62]. Independently, Ahmed [2], Da Prato and Frankowska [23], Aubin and Da Prato [9], and Aubin et al. [10] have considered stochastic differential inclusions symbolically written in the differential form $\mathrm{d}x_t \in F(t, x_t)\mathrm{d}t + G(t, x_t)\mathrm{d}B_t$ and understood as a problem consisting in finding a system $(\mathcal{P}_{\mathbb{F}}, x, B)$ consisting of a filtered probability space $\mathcal{P}_{\mathbb{F}}$, a continuous process $x = (x_t)_{t\geq 0}$, and an $\mathbb{F}$-Brownian motion such that $x_t = x_0 + \int_0^t f_\tau \mathrm{d}\tau + \int_0^t g_\tau \mathrm{d}B_\tau$ with $f_t \in (F \circ x)_t =: F(t, x_t)$ and $g_t \in (G \circ x)_t =: F(t, x_t)$ a.s. for $t \geq 0$. Stochastic functional inclusions defined by Hiai [38] and Kisielewicz [51] are in the general case understood as a problem consisting in finding a system $(\mathcal{P}_{\mathbb{F}}, x, B)$ such that $x_t - x_s \in \mathrm{cl}_{\mathbb{L}}\{J_{st}(F \circ x) + \mathcal{J}_{st}(G \circ x)\}$ for every $0 \leq s \leq t < \infty$, where $J_{st}(F \circ x)$ and $\mathcal{J}_{st}(G \circ x)$ denote set-valued functional integrals on the interval $[s, t]$ of $F \circ x$ and $G \circ x$, respectively. It is evident that some properties of stochastic functional inclusions written in integral form follow from properties of set-valued stochastic integrals. Such properties are difficult to obtain for stochastic differential inclusions written in differential form.

The first results dealing with set-valued stochastic integrals with respect to the Wiener process with application to some set-valued stochastic differential equations are due to Bocşan [22]. More general definitions and properties of set-valued stochastic integrals were given in the above-cited papers of Hiai and Kisielewicz, where set-valued stochastic integrals are defined as certain subsets of the spaces $\mathbb{L}^2(\Omega, \mathbb{R}^n)$ and $\mathbb{L}^2(\Omega, \mathcal{X})$ of all square integrable random variables with values at $\mathbb{R}^n$ and $\mathcal{X}$, respectively, where $\mathcal{X}$ is a Hilbert space. In this book, such integrals are called stochastic functional set-valued integrals. Unfortunately, such integrals do not admit a representation by set-valued random variables with values in $\mathbb{R}^n$ and $\mathcal{X}$, because they are not decomposable subsets of $\mathbb{L}^2(\Omega, \mathbb{R}^n)$ and $\mathbb{L}^2(\Omega, \mathcal{X})$, respectively. Later, Jung and Kim [46] (see also [98]) defined a set-valued stochastic integral as a set-valued random variable determined by a closed decomposable hull of the above-mentioned set-valued stochastic functional integral. Unfortunately, the authors did not obtain any properties of such integrals. In Chap. 3, we apply the above approach to the theory of set-valued stochastic integrals of $\mathbb{F}$-nonanticipative multiprocesses and obtain some properties of such integrals.

The first results dealing with partial differential inclusions were in fact simple generalizations of ordinary differential inclusions. They dealt with hyperbolic partial differential inclusions of the form $z''_{x,y} \in F(x, y, z)$. Later on, partial differential inclusions $z''_{x,y} \in F(x, y, z, z'_x, z'_y)$ were also investigated. Such partial differential inclusions have been considered by Kubiaczyk [65], Dawidowski and Kubiaczyk [24], Dawidowski et al. [25], and Sosulski [92, 93], among others. Some hyperbolic partial differential inclusions were considered in Aubin and Frankowska [11]. A new idea dealing with partial differential inclusions was given by Bartuzel and Fryszkowski in their papers [15–17], where partial differential inclusions of the form $Du \in F(u)$ with a lower semicontinuous multifunction F and a partial differential operator D are considered. The existence and properties of solutions of initial and boundary value problems of such inclusions follow from classical results dealing with abstract differential inclusions. As usual, certain types of continuous selection theorems for set-valued mappings play an important role in investigations of such inclusions.

The partial differential inclusions considered in this book have the forms $u'_t(t, x) \in (\mathbb{L}_{FG}u)(t, x) + c(t, x)u(t, x)$ and $\psi(t, x) \in (\mathbb{L}_{FG}u)(t, x) + c(t, x)u(t, x)$, where c and ψ are given functions and $\mathbb{L}_{FG}$ denotes the set-valued diffusion generator defined by given multifunctions F and G. The first results dealing with such partial differential inclusions are due to Kisielewicz [60, 61]. The initial and boundary value problems of such inclusions are investigated by stochastic methods. Their solutions are characterized by weak solutions of stochastic functional inclusions $SFI(F, G)$. Such an approach leads to natural methods of solving some optimal control problems for systems described by the above type of partial differential inclusions. It is a consequence of weak compactness with respect to the convergence in distribution of sets of all weak solutions of considered stochastic functional inclusions.

The content of the book is divided into seven parts. Chapter 1 covers basic notions and theorems of the theory of stochastic processes. Chapter 2 contains the fundamental notions of the theory of set-valued mappings and the theory of set-valued stochastic processes. Chapter 3 is devoted to the theory of set-valued stochastic integrals. Apart from their properties, it contains some important selection theorems. The main results of Chap. 4 deal with properties of stochastic functional and differential inclusions. In particular, it contains theorems dealing with weak compactness with respect to convergence in distribution of solution sets of weak solutions of initial value problems for stochastic functional inclusions. Chapter 5 contains some results dealing with viability theory for forward and backward stochastic functional and differential inclusions, whereas Chaps. 6 and 7 are devoted to some applications of the above-mentioned results to partial differential inclusions and to some optimal control problems for systems described by stochastic functional and partial differential inclusions.

The present book is intended for students, professionals in mathematics, and those interested in applications of the theory. Selected probabilistic methods and the theory of set-valued mappings are needed for understanding the text. Formulas, theorems, lemmas, remarks, and corollaries are numbered separately in each chapter and denoted by pairs of numbers. The first stands for the section number, the second for the number of the formula, theorem, etc. If we need to quote some formula or theorem given in the same chapter, we always write only this pair. In other cases, we will use this pair with information indicated the chapter number. The ends of proofs, theorems, remarks, and corollaries are denoted by $\square$.

The manuscript of this book was read by my colleagues M. Michta and J. Motyl, who made many valuable comments. The last version of the manuscript was read by Professor Diethard Pallaschke. His remarks and propositions were very useful in my last correction of the manuscript. It is my pleasure to thank all of them for their efforts.

Zielona Góra, Poland Michał Kisielewicz

List of Symbols

$\mathbb{F}$	– filtration of a probability space $(\Omega, \mathcal{F}, P)$, 1
$\mathcal{P}_{\mathbb{F}}$	– filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$, 1
$\mathbb{R}^+$	– set of all non-negative real numbers, 2
$\in$	– is an element of, 1
$\mathbb{R}^n$	– n-dimensional Euclidean spaces, 2
$C(\mathbb{R}^+, \mathbb{R}^n)$	– metric space of continuous functions, 2
$\mathcal{D}(\mathbb{R}^+, \mathbb{R}^n)$	– metric space of càdlàg functions, 2
Q	– set of all rational numbers, 2
$\subset$	– subset of (set inclusion relation), 3
$\cap$	– intersection of sets, 2
$\cup$	– union of sets, 2
$A \setminus B$	– complement of B with respect to A, 3
$\notin$	– is not an element of, 3
τ_D^X	– first exit time of a stochastic process X from a set D, 3
$S \wedge T$	– minimum of stopping times S and T, 3
$S \vee T$	– maximum of stopping times S and T, 3
$\mathcal{F}_T$	– σ-algebra induced by a stopping time T, 3
cad($\mathbb{F}$)	– family of $\mathbb{F}$-adapted càdlàg processes, 3
$\sigma(\mathcal{M})$	– σ-algebra generated by a family $\mathcal{M}$ of random variables, 4
$\beta(\mathcal{X})$	– Borel σ-algebra of subsets of a metric space $(\mathcal{X}, \rho)$, 4
$\mathcal{M}(\mathcal{X})$	– space of probability measures on $\beta(\mathcal{X})$, 4
$P_n \Rightarrow P$	– weak convergence of a sequence of probability measures, 4
PX^{-1}	– distribution of a random variable X, 6
$X_n \xrightarrow{P} X$	– convergence in probability of a sequence of random variables, 6
$X_n \to X$ $a.s.$	– convergence a.s. of a sequence of random variables, 6
$X_n \Rightarrow X$	– convergence in distribution of a sequence of random variables, 6

l.s.c.	– lower semicontinuity, 71
u.s.c.	– upper semicontinuity, 71
H – l.s.c.	– lower semicontinuity with respect to the Hausdorff metric, 71
H – u.s.c.	– upper semicontinuity with respect to the Hausdorff metric, 70
$\mathrm{Comp}(Y)$	– space of all nonempty compact subsets of a topological space Y, 71
$\sigma(\cdot, A)$	– support function of a set $A \subset \mathbb{R}^d$, 77
$\mathrm{Conv}(\mathbb{R}^d)$	– space of all nonempty compact convex subsets of $\mathbb{R}^d$, 77
s(A)	– Steiner point of a set $A \in \mathrm{Conv}(\mathbb{R}^d)$, 78
$\langle \cdot, \cdot \rangle$	– inner product in the space $\mathbb{R}^d$, 78
$\mathrm{Graph}(F)$	– graph of a multifunction F, 82
$\mathrm{cl}(A)$	– closure of a subset A of a topologigal space, 83
$S(F)$	– set of all selectors $f \in \mathbb{L}^p(T, \mathbb{R}^d)$ of a multifunction F, 84
$\mathcal{M}(T, \mathbb{R}^d)$	– space of all measurable multifunctions $F : T \to \mathrm{Cl}(\mathbb{R}^d)$, 84
$\mathcal{A}(T, \mathbb{R}^d)$	– subset of $\mathcal{M}(T, \mathbb{R}^d)$ such that $S(F) \neq \emptyset$, 84
$\overline{\mathrm{co}}\, S(F)$	– closed convex hull of $S(F)$, 85
$\mathrm{dec}\{C\}$	– decomposable hull of a set $C \subset \mathbb{L}^p(T, \mathbb{R}^d)$, 89
$\overline{\mathrm{dec}}\{C\}$	– closed decomposable hull of a set $C \subset \mathbb{L}^p(T, \mathbb{R}^d)$, 89
$\Sigma_{\mathbb{F}}$	– σ-algebra of $\mathbb{F}$-nonanticipative subsets of $T \times \Omega$, 96
$S_{\mathbb{F}}(\Phi)$	– set of all $\mathbb{F}$-nonaticipatine selectors of multifunction Φ, 96
$\mathcal{M}(\mathbb{R}^+ \times \Omega, \mathbb{R}^d)$	– space of all measurable set-valued processes, 97
$\mathcal{M}_{\mathbb{F}}(\mathbb{R}^+ \times \Omega, \mathbb{R}^d)$	– space of $\Sigma_{\mathbb{F}}$-measurable multifunctions, 97
$\mathcal{L}^2(\mathbb{R}^+ \times \Omega, \mathbb{R}^d)$	– space of measurable square integrable multifunctions, 97
$\mathcal{L}^2_{\mathbb{F}}(\mathbb{R}^+ \times \Omega, \mathbb{R}^d)$	– space of square integrable $\Sigma_{\mathbb{F}}$-measurable multifunctions, 97
$E[\Phi \vert \mathcal{G}]$	– $\mathcal{G}$-conditional expectation of a set-valued mapping Φ, 99
J	– linear mapping defined by $J(\phi) = (\int_0^T \phi_t \mathrm{d}t)(\cdot)$, 103
$\mathcal{J}$	– linear mapping defined by $\mathcal{J}(\psi) = (\int_0^T \psi_t \, \mathrm{d}B_t)(\cdot)$, 103
$(\mathcal{A}) \int_0^T \Phi_t \mathrm{d}t$	– set-valued stochastic Aumann's integral, 115
$\int_0^T \Phi_t dt$	– set-valued stochastic Aumann's integral, 115
$\int_0^T \Psi_t dB_t$	– set-valued stochastic Itô integral, 115
$D(\Psi)$	– set-valued mapping $D(\Psi)_t(\omega) = \{v \cdot v^* : v \in \Psi_t(\omega)\}$, 133
$\mathcal{C}_r$	– metric space of continuous functions $\varphi : \mathbb{R}^r \to \mathbb{R}^r$, 133
$\mathcal{C}_{r \times r}$	– metric space of continuous functions $\psi : \mathbb{R}^r \to \mathbb{R}^{r \times r}$, 133
$\varphi(h)$	– gradient of a function $h \in C_0^2(\mathbb{R}^r, \mathbb{R})$, 133
$\psi(h)$	– matrix of second partial derivatives of $h \in C_0^2(\mathbb{R}^r, \mathbb{R})$, 133

$\mathbb{L}^x_{fg}(\varphi,\psi)$	–	semi-elliptic differential operator, 136
$\mathbb{L}^x_{fg}h$	–	semi-elliptic differential operator, 136
$SFI(F,G)$	–	stochastic functional inclusion, 147
$\overline{SFI}(F,G)$	–	stochastic functional inclusion, 147
$\mathcal{X}_\mu(F,G)$	–	set of all weak solutions of $SFI(F,G)$, 148
$\overline{\mathcal{X}}_\mu(F,G)$	–	set of all weak solutions of $\overline{SFI}(F,G)$, 148
$\mathbb{L}^x_{fg}$	–	semi-elliptic partial differential operator, 151
$\mathbb{L}^x_{AB}$	–	set of all $\mathbb{L}^x_{fg}$ for $(f,g)\in A\times B$, 151
$\mathcal{M}^x_{AB}$	–	family of all $\mathbb{L}^x_{fg}\in\mathbb{L}^x_{AB}$ generating local martingales, 152
$\mathbb{L}^x_{FG}$	–	set $\mathbb{L}^x_{S_{\mathbb{F}}(F\circ x)S_{\mathbb{F}}(G\circ x)}$, 152
$SDI(F,G)$	–	stochastic differential inclusion, 163
$\overline{SDI}(F,G)$	–	stochastic differential inclusion, 163
$BSDI(F,H)$	–	backward stochastic differential inclusion, 165
$\mathcal{B}(F,H)$	–	set of all weak solutions of $BSDI(F,H)$, 166
$\mathcal{CB}(F,H)$	–	set of all continuous weak solutions of $BSDI(F,H)$, 166
$\mathcal{S}(\mathbb{F},\mathbb{R}^d)$	–	space of d-dimensional continuous $\mathbb{F}$-semimartingales, 167
$\mathbb{D}(\mathbb{F},\mathbb{R}^d)$	–	space of d-dimensional $\mathbb{F}$-adapted càdlàg processes, 167
$\mathbf{C}(\mathbb{F},\mathbb{R}^{\mathbf{d}})$	–	space of d-dimensional $\mathbb{F}$-adapted continuous processes, 167
$\mathcal{T}_K(t,x)$	–	stochastic tangent set to $K\subset\mathbb{R}^d$, 198
$\mathcal{S}_K(t,x)$	–	stochastic tangent set to $K\subset\mathbb{R}^d$, 199
$\tau_K(t,x)$	–	stochastic contingent set to $K\subset\mathbb{R}^d$, 200
$\mathcal{A}_{fg}$	–	infinitesimal diffusion generator, 217
$\mathbb{L}_{FG}$	–	set-valued semi-elliptic partial differential operator, 218
$\mathcal{C}(F),\mathcal{C}(G)$	–	sets of continuous selectors of multifunctions F and G, 218
$\mathcal{A}_{FG}$	–	set of the form $\{\mathcal{A}_{fg}:(f,g)\in\mathcal{C}(F)\times\mathcal{C}(G)\}$, 218
$\mathcal{L}_{FG}$	–	set of the form $\{\mathcal{L}_{fg}:(f,g)\in\mathcal{C}(F)\times\mathcal{C}(G)\}$, 218

Chapter 1
Stochastic Processes

In this chapter we give a survey of concepts of the theory of stochastic processes. It is assumed that the basic notions of measure and probability theories are known to the reader.

1 Filtered Probability Spaces and Stopping Times

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ a family of sub-σ-algebras $\mathcal{F}_t$ of σ-algebra $\mathcal{F}$ such that $\mathcal{F}_s \subset \mathcal{F}_t$ for $0 \leq s \leq t < \infty$. A system $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ is said to be a filtered probability space. It is called complete if P is a complete measure, i.e., $2^B \subset \mathcal{F}$ for every $B \in \mathcal{F}$ such that $P(B) = 0$. We say that a filtration $\mathbb{F}$ satisfies the usual conditions if $\mathcal{F}_0$ contains all P-null sets of $\mathcal{F}$ and $\mathcal{F}_t = \bigcap_{\varepsilon > 0} \mathcal{F}_{t+\varepsilon}$ for every $t \geq 0$. If the last condition is satisfied, we say that a filtration $\mathbb{F}$ is right continuous. We call a filtration $\mathbb{F}$ left continuous if $\mathcal{F}_t$ is generated by a family $\{\mathcal{F}_s : 0 \leq s < t\}$ for every $t \geq 0$, i.e., $\mathcal{F}_t = \sigma(\{\mathcal{F}_s : 0 \leq s < t\})$ for every $t \geq 0$. A filtration $\mathbb{F}$ is said to be continuous if it is right and left continuous.

Remark 1.1. From a practical point of view, a filtered probability space $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ is usually regarded as a probability model of a given experiment with results belonging to Ω. The family $\mathcal{F}$ is treated as a set of informations on elements of Ω, whereas the filtration contains all informations contained in $\mathcal{F}$ given up to $t \geq 0$. □

Given a filtered probability space $\mathcal{P}_{\mathbb{F}}$ and a metric space $(\mathcal{X}, \rho)$, by an $\mathcal{X}$-random variable on $\mathcal{P}_{\mathbb{F}}$ we mean an $(\mathcal{F}, \beta_{\mathcal{X}})$-measurable mapping $X : \Omega \to \mathcal{X}$, i.e., such that $X^{-1}(A) \in \mathcal{F}$ for every $A \in \beta(\mathcal{X})$, where as usual, $\beta(\mathcal{X})$ denotes the Borel σ-algebra on $\mathcal{X}$ and $X^{-1}(A) = \{\omega \in \Omega : X(\omega) \in A\}$. We shall also say that X is a random variable on $\mathcal{P}_{\mathbb{F}}$ with values at $\mathcal{X}$. In particular, if $\mathcal{X} = \mathbb{R}^n$ then, an $\mathcal{X}$-random variable is also called an n-dimensional random variable. Given a

M. Kisielewicz, *Stochastic Differential Inclusions and Applications*,
Springer Optimization and Its Applications 80, DOI 10.1007/978-1-4614-6756-4_1,

random variable $X : \Omega \to \mathcal{X}$, we denote by $\mathcal{F}_X$ the σ-algebra generated by X, i.e., the smallest σ-algebra on Ω containing all sets $X^{-1}(U)$ for all open sets $U \subset \mathcal{X}$. It is easy to see that $\mathcal{F}_X = \{X^{-1}(A) : A \in \beta(\mathcal{X})\}$.

Remark 1.2. It can be verified that if $X, Y : \Omega \to \mathbb{R}^n$ are given functions, then Y is $\mathcal{F}_X$-measurable if and only if there exists a Borel-measurable function $g : \mathbb{R}^n \to \mathbb{R}^n$ such that $Y = g(X)$. □

From a practical point of view, random variables can be applied to mathematical modeling of static random processes. In the case of dynamic ones, instead of random variables, we have to apply families $X = (X_t)_{t \geq 0}$ of random variables parameterized by a parameter $t \geq 0$ usually treated as the time at which the modeled dynamical process is taking place. Families $X = (X_t)_{t \geq 0}$ of n-dimensional random variables $X_t : \Omega \to \mathbb{R}^n$ are called n-dimensional stochastic processes on $\mathcal{P}_{\mathbb{F}}$. Such processes are called continuous if for a.e. $\omega \in \Omega$ mappings $\mathbb{R}^+ \ni t \to X_t(\omega) \in \mathbb{R}^n$, called trajectories of X, are continuous. In a similar way, we define càdlàg and càglàd stochastic processes on $\mathcal{P}_{\mathbb{F}}$. An n-dimensional process X is said to be a càdlàg process if for a.e. $\omega \in \Omega$, its trajectory $\mathbb{R}^+ \ni t \to X_t(\omega) \in \mathbb{R}^n$ is right continuous and possesses the left-hand limit $X_{t-}(\omega)$ for every $t > 0$. Similarly, a process X is called a càglàd process if for a.e. $\omega \in \Omega$, its trajectory $\mathbb{R}^+ \ni t \to X_t(\omega) \in \mathbb{R}^n$ is left continuous and possesses the right-hand limit $X_{t+}(\omega)$ for every $t > 0$. If for every $t \geq 0$, a random variable X_t is $\mathcal{F}_t$-measurable, then a process X is called $\mathbb{F}$-adapted. Many more notions and properties dealing with stochastic processes are given in Sect. 3.

Remark 1.3. It can be proved that all random variables $X : \Omega \to \mathcal{C}$ and $X : \Omega \to \mathcal{D}$ with $\mathcal{C} = C(\mathbb{R}^+, \mathbb{R}^n)$ and $\mathcal{D} = \mathcal{D}(\mathbb{R}^+, \mathbb{R}^n)$, where $C(\mathbb{R}^+, \mathbb{R}^n)$ and $\mathcal{D}(\mathbb{R}^+, \mathbb{R}^n)$ denote the metric spaces of all continuous and càdlàg functions $x : \mathbb{R}^+ \to \mathbb{R}^n$ with appropriate metrics, can be described respectively as n-dimensional continuous and càdlàg processes. □

A random variable $T : \Omega \to [0, \infty]$ on $\mathcal{P}_{\mathbb{F}}$ such that $\{T \leq t\} \in \mathcal{F}_t$ for every $t \geq 0$ is said to be an $\mathbb{F}$-stopping time. If a filtration $\mathbb{F}$ is right continuous, then the condition $\{T \leq t\} \in \mathcal{F}_t$ in the above definition can be replaced by $\{T < t\} \in \mathcal{F}_t$ for every $t \geq 0$. This follows from the following theorem.

Theorem 1.1. *If a filtered probability space $\mathcal{P}_{\mathbb{F}}$ is such that $\mathbb{F}$ is right continuous, then a random variable $T : \Omega \to [0, \infty]$ is an $\mathbb{F}$-stopping time on $\mathcal{P}_{\mathbb{F}}$ if and only if $\{T < t\} \in \mathcal{F}_t$ for every $t \geq 0$.*

Proof. Let $\{T < t\} \in \mathcal{F}_u$ for $u > t$ and $t \geq 0$. Since $\{T \leq t\} = \bigcap_{t+\varepsilon > u > t}\{T < u\}$ for every $\varepsilon > 0$ and $\mathbb{F}$ is right continuous, we have $\{T \leq t\} \in \bigcap_{u>t} \mathcal{F}_u = \mathcal{F}_t$ for $t \geq 0$. Therefore, the condition $\{T < t\} \in \mathcal{F}_t$ for $t \geq 0$ implies that $\{T \leq t\} \in \mathcal{F}_t$ for $t \geq 0$. Conversely, if $\{T \leq t\} \in \mathcal{F}_t$ for $t \geq 0$, then we also have $\{T < t\} = \bigcup_{\varepsilon \in Q} \bigcup_{s \in Q \cap [0, t-\varepsilon]} \{T \leq s\} \in \mathcal{F}_t$, where Q is the set of all rational numbers of the real line $\mathbb{R}$. □

Example 1.1. Let $X = (X_t)_{t \geq 0}$ be a càdlàg process and $\Lambda \subset \mathbb{R}$ a Borel set. We define a hitting time of Λ for X by taking $T(\omega) = \inf\{t > 0 : X_t(\omega) \in \Lambda\}$ for $\omega \in \Omega$. If Λ is an open set, then by right continuity of X, we have $\{T < t\} \subset \bigcup_{s \in Q \cap [0,t)} \{X_s \in \Lambda\}$. If furthermore, X is $\mathbb{F}$-adapted, then $\{X_s \in \Lambda\} = X_s^{-1}(\Lambda) \in \mathcal{F}_s$ for $s \in Q \cap [0,t)$. Therefore, for such a process X, one has $\{T < t\} \in \bigcup_{s \in Q \cap [0,t)} \mathcal{F}_s = \mathcal{F}_t$ for every $t \geq 0$. From the above theorem, it follows that if a filtration $\mathbb{F}$ is right continuous, then for the above process X and an open set $\Lambda \subset \mathbb{R}$, a hitting time of Λ for X is an $\mathbb{F}$-stopping time.

Theorem 1.2. *Let $X = (X_t)_{t \geq 0}$ be a càdlàg and $\mathbb{F}$-adapted process on $\mathcal{P}_{\mathbb{F}}$. Then for every closed set $\Lambda \subset \mathbb{R}$, the random variable $T : \Omega \to \mathbb{R}$ defined by $T(\omega) = \inf\{t > 0 : X_t(\omega) \in \Lambda$ or $X_{t-}(\omega) \in \Lambda\}$ for $\omega \in \Omega$ is an $\mathbb{F}$-stopping time.*

Proof. Let $A_n = \{x \in \mathbb{R} : \text{dist}(x, \Lambda) < 1/n\}$. It is easy to see that A_n is an open set. But $X_{t-}(\omega) = \lim_{s \to t, s < t} X_s(\omega)$ for $\omega \in \Omega$. Therefore, $\{X_{t-} \in \Lambda\} = \bigcap_{n \geq 1} \bigcup_{s \in Q \cap [0,t)} \{X_s \in A_n\}$ for $t \geq 0$. Then $\{T \leq t\} = \{X_t \in \Lambda\} \cup \{X_{t-} \in \Lambda\} = \{X_t \in \Lambda\} \cup \bigcap_{n \geq 1} \bigcup_{s \in Q \cap [0,t)} \{X_s \in A_n\}$ for $t \geq 0$. By the properties of a family X it follows that $\{X_t \in \Lambda\} \in \mathcal{F}_t$ and $\bigcap_{n \geq 1} \bigcup_{s \in Q \cap [0,t)} \{X_s \in A_n\} \in \mathcal{F}_t$ for $t \geq 0$. Therefore, for every $t \geq 0$ one has $\{T \leq t\} \in \mathcal{F}_t$. □

The above result can be easily extended for n-dimensional càdlàg and $\mathbb{F}$-adapted processes.

Theorem 1.3. *Let $X = (X_t)_{t \geq 0}$ be an n-dimensional càdlàg and $\mathbb{F}$-adapted process. Then for every domain D in $\mathbb{R}^n$, the random variable $T : \Omega \to \mathbb{R}$ defined by $T(\omega) = \inf\{t > 0 : X_t(\omega) \notin D\}$ for $\omega \in \Omega$ is an $\mathbb{F}$-stopping time.*

Proof. Let $\Lambda = \mathbb{R}^n \setminus D$. The set Λ is closed and $T(\omega) = \inf\{t > 0 : X_t(\omega) \in \Lambda\}$ for $\omega \in \Omega$. Hence, similarly as in the proof of Theorem 1.2, it follows that T is an $\mathbb{F}$-stopping time. □

The $\mathbb{F}$-stopping time defined in Theorem 1.3 is said to be the first exit time of the process X from D. Usually it is denoted by τ_D^X, or simply by τ_D if X is fixed.

Remark 1.4. Immediately from the definition of stopping times it follows that for all $\mathbb{F}$-stopping times S and T on $\mathcal{P}_{\mathbb{F}}$, also $S \wedge T$, $S \vee T$, $S + T$, and αS with $\alpha > 1$ are $\mathbb{F}$-stopping times on $\mathcal{P}_{\mathbb{F}}$. □

Given a filtered probability space $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ with $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$, the σ-algebra $\mathcal{F}_t$ can be thought as representing all (theoretically) observable events up to and including time t. We would like to have an analogous notion of events that are observable before a random time T. To get that, we have to define an $\mathbb{F}$-stopping time σ-algebra $\mathcal{F}_T$ induced by an $\mathbb{F}$-stopping time T. It is defined by setting $\mathcal{F}_T = \{A \in \mathcal{F} : A \cap \{T \leq t\} \in \mathcal{F}_t, \text{ for } t \geq 0\}$. The present definition represents "knowledge" up to time T. This follows from the following theorem.

Theorem 1.4. *Let $\text{cad}(\mathbb{F})$ denote the family of all $\mathbb{F}$-adapted càdlàg processes $X = (X_t)_{t \geq 0}$ on $\mathcal{P}_{\mathbb{F}}$. Then for every finite $\mathbb{F}$-stopping time T, one has $\mathcal{F}_T = \sigma(\{X_T : X \in \text{cad}(\mathbb{F})\})$.*

Proof. Let $\mathcal{G}_T = \sigma(\{X_T : X \in \mathrm{cad}(\mathbb{F})\})$ and let $A \in \mathcal{F}_T$. Define a process $X = (X_t)_{t\geq 0}$ on $\mathcal{P}_{\mathbb{F}}$ by setting $X_t = \mathbb{1}_A \cdot \mathbb{1}_{\{t\geq T\}}$ for $t \geq 0$. We have $\mathbb{1}_{\{T\geq T\}} = 1$. Therefore, $X_T = \mathbb{1}_A$. By the above definition of a process X, we have $X \in \mathrm{cad}(\mathbb{F})$, which implies that $A \in \mathcal{G}_T$. Then $\mathcal{F}_T \subset \mathcal{G}_T$.

Let $X \in \mathrm{cad}(\mathbb{F})$. We need to show that X_T is $\mathcal{F}_T$-measurable. We can consider X as a function $X : [0,\infty) \times \Omega \to \mathbb{R}$. Construct a function $\varphi : \{T \leq t\} \to [0,\infty) \times \Omega$ by setting $\varphi(\omega) = (T(\omega), \omega)$ for $\omega \in \{T \leq t\}$. Since $X \in \mathrm{cad}(\mathbb{F})$, then $X_T = X \circ \varphi$ is a measurable mapping from $(\{T \leq t\}, \mathcal{F}_t \cap \{T \leq t\})$ into $(\mathbb{R}, \beta(\mathbb{R}))$, where $\beta(\mathbb{R})$ denotes the Borel σ-algebra on $\mathbb{R}$. Therefore, $\{\omega \in \Omega : X(T(\omega), \omega) \in B\} \cap \{T \leq t\} \in \mathcal{F}_t$ for every $t \geq 0$ and $B \in \beta(\mathbb{R})$. Then X_T is $\mathcal{F}_T$-measurable. Thus $\mathcal{G}_T \subset \mathcal{F}_T$. □

The following result follows immediately from the above definitions of an $\mathbb{F}$-stopping time and an σ-algebra $\mathcal{F}_T$.

Theorem 1.5. *Let S and T be $\mathbb{F}$-stopping times on $\mathcal{P}_{\mathbb{F}}$ such that $S \leq T$ a.s. Then $\mathcal{F}_S \subset \mathcal{F}_T$ and $\mathcal{F}_S \cap \mathcal{F}_T = \mathcal{F}_{S\wedge T}$.* □

2 Weak Compactness of Sets of Random Variables

Let $(\mathcal{X}, \rho)$ be a separable metric space and $\beta(\mathcal{X})$ a Borel σ-algebra on $\mathcal{X}$. Denote by $\mathcal{M}(\mathcal{X})$ the space of all probability measures on $\beta(\mathcal{X})$ and let $C_b(\mathcal{X})$ be the space of all continuous bounded functions $f : \mathcal{X} \to \mathbb{R}$. We say that a sequence $(P_n)_{n=1}^\infty$ of $\mathcal{M}(\mathcal{X})$ weakly converges to $P \in \mathcal{M}(\mathcal{X})$ if $\lim_{n\to\infty} \int_{\mathcal{X}} f \mathrm{d}P_n = \int_{\mathcal{X}} f \mathrm{d}P$ for every $f \in C_b(\mathcal{X})$. We shall denote this convergence by $P_n \Rightarrow P$. We have the following theorem.

Theorem 2.1. *The following conditions are equivalent to weak convergence of a sequence $(P_n)_{n=1}^\infty$ of $\mathcal{M}(\mathcal{X})$ to $P \in \mathcal{M}(\mathcal{X})$:*

(i) $\limsup_{n\to\infty} P_n(F) \leq P(F)$ *for every closed set* $F \subset \mathcal{X}$.
(ii) $\liminf_{n\to\infty} P_n(G) \geq P(G)$ *for every open set* $G \subset \mathcal{X}$.

Proof. Let $P_n \Rightarrow P$. Hence it follows that $\limsup_{n\to\infty} P_n(F) \leq \lim_{n\to\infty} \int_{\mathcal{X}} f_k \mathrm{d}P_n = \int_{\mathcal{X}} f_k \mathrm{d}P$ for every closed set $F \subset \mathcal{X}$, where $f_k(x) = \psi(k \cdot \mathrm{dist}(x, F))$ with $\psi(t) = 1$ for $t \leq 0$, $\psi(t) = 0$ for $t \geq 1$, and $\psi(t) = 1 - t$ for $0 \leq t \leq 1$. Passing in the above inequality to the limit with $k \to \infty$, we see that (i) is satisfied. It is easy to see that (i) is equivalent to (ii). Indeed, by virtue of (i), for every open set $G \subset \mathcal{X}$ we obtain $\limsup_{n\to\infty} P_n(\mathcal{X} \setminus G) \leq P(\mathcal{X} \setminus G)$, which implies that $\liminf_{n\to\infty} P_n(G) \geq P(G)$. In a similar way, we can see that from (ii), it follows that $\limsup_{n\to\infty} P_n(F) \leq P(F)$ for every closed set $F \subset \mathcal{X}$.

Assume that (i) is satisfied and let $f \in C_b(\mathcal{X})$. We can assume that $0 < f(x) < 1$ for $x \in \mathcal{X}$. Then

$$\sum_{i=1}^{k} \frac{i-1}{k} \cdot P\left\{x \in \mathcal{X} : \frac{i-1}{k} \le f(x) < \frac{i}{k}\right\} \le \int_{\mathcal{X}} f(x) \mathrm{d}P$$
$$\le \sum_{i=1}^{k} \frac{i}{k} \cdot P\left\{x \in \mathcal{X} : \frac{i-1}{k} \le f(x) < \frac{i}{k}\right\}.$$

For every $F_i = \{x \in \mathcal{X} : i/k \le f(x)\}$, the right-hand side of the above inequality is equal to $\sum_{i=0}^{k-1} P_n(F_i)/k$, and the left-hand side to $\sum_{i=0}^{k-1} P_n(F_i)/k - 1/k$. This and (i) imply

$$\limsup_{n\to\infty} \int_{\mathcal{X}} f(x) \mathrm{d}P_n \le \limsup_{n\to\infty} \sum_{i=0}^{k-1} P_n(F_i)/k \le \sum_{i=0}^{k-1} P(F_i)/k \le 1/k + \int_{\mathcal{X}} f(x) \mathrm{d}P.$$

Then $\limsup_{n\to\infty} \int_{\mathcal{X}} f(x)\mathrm{d}P_n \le \int_{\mathcal{X}} f(x)\mathrm{d}P$. Repeating the above procedure with a function $g = 1 - f$, we obtain $\liminf_{n\to\infty} \int_{\mathcal{X}} f(x)\mathrm{d}P_n \ge \int_{\mathcal{X}} f(x)\mathrm{d}P$. Therefore,

$$\int_{\mathcal{X}} f(x)\mathrm{d}P \le \liminf_{n\to\infty} \int_{\mathcal{X}} f(x)\mathrm{d}P_n \le \limsup_{n\to\infty} \int_{\mathcal{X}} f(x)\mathrm{d}P_n \le \int_{\mathcal{X}} f(x)\mathrm{d}P.$$

Thus $\lim_{n\to\infty} \int_{\mathcal{X}} f(x)\mathrm{d}P_n = \int_{\mathcal{X}} f(x)\mathrm{d}P$ for every $f \in C_b(\mathcal{X})$. □

We can consider weakly compact subsets of the space $\mathcal{M}(\mathcal{X})$. Let us observe that we can define on $\mathcal{M}(\mathcal{X})$ a metric d such that weak convergence in $\mathcal{M}(\mathcal{X})$ of a sequence $(P_n)_{n=1}^{\infty}$ to P is equivalent to $d(P_n, P) \to 0$ as $n \to \infty$. Therefore, we say that a set $\Lambda \subset \mathcal{M}(\mathcal{X})$ is relatively weakly compact if every sequence $(P_n)_{n=1}^{\infty}$ of Λ possesses a subsequence $(P_{n_k})_{k=1}^{\infty}$ weakly convergent to $P \in \mathcal{M}(\mathcal{X})$. If $P \in \Lambda$ then Λ, is called weakly compact. We shall prove that for relative weak compactness of a set $\Lambda \subset \mathcal{M}(\mathcal{X})$, it suffices that Λ be tight, i.e., that for every $\varepsilon > 0$ there exist a compact set $K \subset \mathcal{X}$ such that $P(K) \ge 1 - \varepsilon$ for every $P \in \Lambda$.

Theorem 2.2. *Every tight set* $\Lambda \subset \mathcal{M}(\mathcal{X})$ *is relatively weakly compact.*

Proof. Assume first that $(\mathcal{X}, \rho)$ is a compact metric space. By the Riesz theorem, we have $\mathcal{M}(\mathcal{X}) = \{\mu \in C^*(\mathcal{X}) : \mu(f) \ge 0 \text{ for } f \ge 0 \text{ and } \mu(\mathbf{1}) = 1\}$, where $\mathbf{1}(x) = 1$ for $x \in \mathcal{X}$ and $C^*(\mathcal{X})$ is the dual space of $C(\mathcal{X})$. Since $C(\mathcal{X}) = C_b(\mathcal{X})$, weak convergence of probability measures is in this case equivalent to weak *-topology convergence on $C^*(\mathcal{X})$. Then $\mathcal{M}(\mathcal{X})$ is weakly compact, because every weakly *-closed subset of the unit ball of $C^*(\mathcal{X})$ is weakly *-compact.

In the general case, let us note that $\mathcal{X}$ is homeomorphic to a subset of a compact metric space. Therefore, we can assume that $\mathcal{X}$ is a subset of a compact metric space $\tilde{\mathcal{X}}$. For every probability measure μ on $(\mathcal{X}, \beta(\mathcal{X}))$ let us define on $(\tilde{\mathcal{X}}, \beta(\tilde{\mathcal{X}}))$ a probability measure $\tilde{\mu}$ by setting $\tilde{\mu}(\tilde{A}) = \mu(\tilde{A} \cap \mathcal{X})$ for $\tilde{A} \in \beta(\tilde{\mathcal{X}})$. Let us observe that $A \subset \mathcal{X}$ belongs to $\beta(\mathcal{X})$ if and only if $A = \tilde{A} \cap \mathcal{X}$ for every $\tilde{A} \in \beta(\tilde{\mathcal{X}})$.

We shall show now that if $\Lambda \subset \mathcal{M}(\mathcal{X})$ is tight, then every sequence $(\mu_n)_{n=1}^{\infty}$ of Λ possesses a subsequence weakly convergent to $\mu \in \mathcal{M}(\mathcal{X})$. Assume that a sequence $(\mu_n)_{n=1}^{\infty}$ is given and let $(\tilde{\mu}_n)_{n=1}^{\infty}$ be a sequence of probability measures defined on $\beta(\tilde{\mathcal{X}})$ by the sequence $(\mu_n)_{n=1}^{\infty}$ such as above, i.e., by taking $\tilde{\mu}_n(\tilde{A}) = \mu(\tilde{A} \cap \mathcal{X})$ for $\tilde{A} \in \beta(\tilde{\mathcal{X}})$ and $n \geq 1$. It is clear that a sequence $(\tilde{\mu}_n)_{n=1}^{\infty}$ possesses a subsequence $(\tilde{\mu}_{n_k})_{k=1}^{\infty}$ weakly convergent to a probability measure ν on $(\tilde{\mathcal{X}}, \beta(\tilde{\mathcal{X}}))$. We shall show that there exists a probability measure μ on $(\mathcal{X}, \beta(\mathcal{X}))$ such that $\tilde{\mu} = \nu$ and that a subsequence $(\mu_{n_k})_{k=1}^{\infty}$ converges weakly to μ. Indeed, by tightness of Λ, for every $r = 1, 2, \ldots$, there exists a compact set $K_r \subset \mathcal{X}$ such that $\mu_n(K_r) \geq 1 - 1/r$ for every $n \geq 1$. It is clear that K_r is also a compact subset of $\tilde{\mathcal{X}}$, and therefore, $K_r \in \beta(\mathcal{X}) \cap \beta(\tilde{\mathcal{X}})$ and $\tilde{\mu}_{n_k}(K_r) = \mu_{n_k}(K_r)$. But $\tilde{\mu}_{n_k} \Rightarrow \nu$. Therefore, $\nu(K_r) \geq \limsup_{k\to\infty} \mu_{n_k}(K_r) \geq 1 - 1/r$. Thus $E =: \bigcup_{r\geq 1} K_r \subset \mathcal{X}$ and $E \in \beta(\mathcal{X}) \cap \beta(\tilde{\mathcal{X}})$. For every $A \in \beta(\mathcal{X})$, we have $A \cap E \in \beta(\tilde{\mathcal{X}})$ because $A \cap E = \tilde{A} \cap \mathcal{X} \cap E = \tilde{A} \cap E$ for every $\tilde{A} \in \beta(\tilde{\mathcal{X}})$. Put $\mu(A) = \nu(A \cap E)$ for every $A \in \beta(\mathcal{X})$. It is clear that μ is a probability measure on $(\mathcal{X}, \beta(\mathcal{X}))$ and $\tilde{\mu} = \nu$. Finally, we verify that $\mu_{n_k} \Rightarrow \mu$. Indeed, let A be a closed subset of $\mathcal{X}$. Then $A = \tilde{A} \cap \mathcal{X}$ for every closed set $\tilde{A} \subset \tilde{\mathcal{X}}$ and $\tilde{\mu}_n(\tilde{A}) = \mu_n(A)$. Therefore, $\limsup_{k\to\infty} \mu_{n_k}(A) = \limsup_{k\to\infty} \tilde{\mu}_{n_k}(\tilde{A}) \leq \tilde{\mu}(\tilde{A}) = \mu(A)$, which by virtue of Theorem 2.1, implies that $\mu_{n_k} \Rightarrow \mu$ as $k \to \infty$. □

Let $(X_n)_{n=1}^{\infty}$ be a sequence of $\mathcal{X}$-random variables $X_n : \Omega_n \to \mathcal{X}$ on a probability space $(\Omega_n, \mathcal{F}_n, P_n)$ for $n \geq 1$. We say that $(X_n)_{n=1}^{\infty}$ converges in distribution to a random variable $X : \Omega \to \mathcal{X}$ defined on a probability space $(\Omega, \mathcal{F}, P)$ if the sequence $(PX_n^{-1})_{n=1}^{\infty}$ of distributions of random variables $X_n : \Omega_n \to \mathcal{X}$ is weakly convergent to the distribution PX^{-1} of X. It is denoted by $X_n \Rightarrow X$. If X_n and X are defined on the same probability space $(\Omega, \mathcal{F}, P)$, then we can define convergence of the above sequence $(X_n)_{n=1}^{\infty}$ in probability and a.s. to a random variable X. We denote the above types of convergence by $X_n \overset{\mathrm{P}}{\to} X$ and $X_n \to X$ a.s., respectively. We have the following important result.

Corollary 2.1. *If $(X_n)_{n=1}^{\infty}$ and X are as above, then $X_n \Rightarrow X$ if and only if $E_n\{f(X_n)\} \to E\{f(X)\}$ as $n \to \infty$ for every $f \in C_b(\mathcal{X})$, where E_n and E are mean value operators taken with respect to probability measures P_n and P, respectively.*

Proof. By the definitions of convergence of sequences of random variables and probability measures, it follows that $X_n \Rightarrow X$ if and only if $\int_{\mathcal{X}} f(x)\,\mathrm{d}[P(X_n)^{-1}] \to \int_{\mathcal{X}} f(x)\,\mathrm{d}[P(X)^{-1}]$ as $n \to \infty$ for every $f \in C_b(\mathcal{X})$. The result follows now immediately from the equalities $\int_{\mathcal{X}} f(x)\,\mathrm{d}[P(X_n)^{-1}] = \int_{\Omega_n} f(X_n)\,\mathrm{d}P_n = E_n\{f(X_n)\}$ and $\int_{\mathcal{X}} f(x)\,\mathrm{d}[P(X)^{-1}] = \int_{\Omega} f(X)\,\mathrm{d}P = E\{f(X)\}$. □

Theorem 2.3. *Let $(\mathcal{X}, \rho)$ be a Polish space, i.e., a complete separable metric space, and $(P_n)_{n=1}^{\infty}$ a sequence of $\mathcal{M}(\mathcal{X})$ weakly convergent to $\mathcal{P} \in \mathcal{M}(\mathcal{X})$ as $n \to \infty$. Then there exist a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ and $\mathcal{X}$-random variables*

X_n and X on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ for $n = 1, 2, \ldots$ such that (i) $P_n = PX_n^{-1}$ for $n = 1, 2, \ldots,$ $\mathcal{P} = PX^{-1}$, and (ii) $\rho(X_n, X) \to 0$ a.s. as $n \to \infty$.

Proof. Let $\tilde{\Omega} = [0, 1)$, $\tilde{\mathcal{F}} = \beta([0, 1))$, and $\tilde{P} = \mu$, where μ is Lebesgue measure on $\beta([0, 1))$. To every finite sequence $(i_1, \ldots, i_k)$ for $k = 1, 2, \ldots$ of positive integers we associate a set $S_{i_1,\ldots,i_k} \in \beta(\mathcal{X}))$ with the boundary $\partial S_{i_1,\ldots,i_k}$ such that

1° If $(i_1, \ldots, i_k) \neq (j_1, \ldots, j_k)$, then $S_{i_1,\ldots,i_k} \cap S_{j_1,\ldots,j_k} = \emptyset$ for $k = 1, 2, \ldots,$
2° $\bigcup_{k=1}^{\infty} S_{i_1,\ldots,i_k} = \mathcal{X}$ and $\bigcup_{j=1}^{\infty} S_{i_1,\ldots,i_k,j} = S_{i_1,\ldots,i_k}$ for $k = 1, 2, \ldots,$
3° $\operatorname{diam}(S_{i_1,\ldots,i_k}) \le 2^{-k}$ for $k = 1, 2, \ldots,$
4° $P_n(\partial S_{i_1,\ldots,i_k}) = 0$ and $\mathcal{P}(\partial S_{i_1,\ldots,i_k}) = 0$ for $k, n = 1, 2, \ldots.$

By virtue of 1° and 2°, a family $\{S_{i_1,\ldots,i_k}\}$ is for every fixed $k \in \mathbb{N}$ a disjoint covering of $\mathcal{X}$ that is a subdivision of a covering for $k' < k$. Such a system of subsets can be defined in the following way. For every k and $m = 1, 2, \ldots,$ we take balls $\sigma_m^{(k)}$ with radii not greater then $2^{-(k+1)}$ that cover $\mathcal{X}$ and are such that $P_n(\partial\sigma_m^{(k)}) = 0$ and $\mathcal{P}(\partial\sigma_m^{(k)}) = 0$ for every $n, k, m \in \mathbb{N}$. For fixed $k \in \mathbb{N}$ let $D_1^{(k)} = \sigma_1^{(k)}$, $D_2^{(k)} = \sigma_2^{(k)} \setminus \sigma_1^{(k)}, \ldots$ $D_n^{(k)} = \sigma_n^{(k)} \setminus \bigcup_{i=1}^{n-1} \sigma_i^{(k)}$, and $S_{i_1,\ldots,i_k} = D_{i_1}^{(1)} \cap D_{i_2}^{(2)} \cap \cdots \cap D_{i_k}^{(k)}$. It can be verified that such sets $S_{i_1,\ldots,i_k}$ satisfy the conditions presented above.

Now for fixed k, let us introduce in the set of all sequences $(i_1, \ldots, i_k)$ the lexicographic order and define in $[0, 1)$ intervals $\Delta_{i_1,\ldots,i_k}$ and $\Delta_{i_1,\ldots,i_k}^{(n)}$ such that

(I) $|\Delta_{i_1,\ldots,i_k}| = \mathcal{P}(S_{i_1,\ldots,i_k})$ and $|\Delta_{i_1,\ldots,i_k}^{(n)}| = P_n(S_{i_1,\ldots,i_k})$,
(II) If $(i_1, \ldots, i_k) < (j_1, \ldots, j_k)$, then $\Delta_{i_1,\ldots,i_k}$ and $\Delta_{i_1,\ldots,i_k}^{(n)}$ are on the left-hand side of $\Delta_{j_1,\ldots,j_k}$ and $\Delta_{j_1,\ldots,j_k}^{(n)}$, respectively,
(III) $\bigcup_{(i_1,\ldots,i_k)} \Delta_{i_1,\ldots,i_k} = [0, 1)$ and $\bigcup_{(i_1,\ldots,i_k)} \Delta_{i_1,\ldots,i_k}^{(n)} = [0, 1)$ for $n \ge 1$.

Such intervals are defined in a unique way. For every $(i_1, \ldots, i_k)$ such that $\overset{\circ}{S}_{i_1,\ldots,i_k} \neq \emptyset$, select a point $x_{i_1,\ldots,i_k} \in \overset{\circ}{S}_{i_1,\ldots,i_k}$, where $\overset{\circ}{S}_{i_1,\ldots,i_k}$ denotes the interior of $S_{i_1,\ldots,i_k}$. For every $\omega \in [0, 1)$, $k = 1, 2, \ldots,$ and $n = 1, 2, \ldots$ we define $X_n^k(\omega)$ and $X^k(\omega)$ by setting $X_n^k(\omega) = x_{i_1,\ldots,i_k}$ if $\omega \in \Delta_{i_1,\ldots,i_k}^{(n)}$ and $X^k(\omega) = x_{i_1,\ldots,i_k}$ if $\omega \in \Delta_{i_1,\ldots,i_k}$. For every $k, n, p \ge 1$ we have $\rho(X_n^k(\omega), X_n^{k+p}(\omega)) \le 1/2^k$ and $\rho(X^k(\omega), X^{k+p}(\omega)) \le 1/2^k$. Therefore, $X_n(\omega) = \lim_{k\to\infty} X_n^k(\omega)$ and $X(\omega) = \lim_{k\to\infty} X^k(\omega)$ exist. Furthermore, $P_n(S_{i_1,\ldots,i_k}) = |\Delta_{i_1,\ldots,i_k}^{(n)}| \to |\Delta_{i_1,\ldots,i_k}| = \mathcal{P}(S_{i_1,\ldots,i_k})$ as $n \to \infty$. Therefore, for every $\omega \in \overset{\circ}{\Delta}_{i_1,\ldots,i_k}$ there is n_k such that $\omega \in \overset{\circ}{\Delta}{}^{(n)}{}_{i_1,\ldots,i_k}$ for $n \ge n_k$. Then $X_n^k(\omega) = X^k(\omega)$ and therefore,

$$\rho(X_n(\omega), X(\omega)) \le \rho(X_n(\omega), X_n^k(\omega)) + \rho(X_n^k(\omega), X^k(\omega)) + \rho(X^k(\omega), X(\omega)) \le 2/2^k$$

for $n \ge n_k$. Thus for every $\omega \in \Omega_0 =: \bigcap_{k=1}^{\infty} \bigcup_{i_1,\ldots,i_k} \overset{\circ}{\Delta}_{i_1,\ldots,i_k}$ we get $X_n(\omega) \to X(\omega)$ as $n \to \infty$. It is easy to see that $\mathcal{P}(\Omega_0) = 1$.

Finally, we shall show that $PX_n^{-1} = P_n$ for $n = 1, 2, \dots$ and $PX^{-1} = \mathcal{P}$. Let us first observe that $\tilde{P}(\{X_n^{k+p} \in \overline{S}_{i_1,\dots,i_k}\}) = \tilde{P}(\{X_n^{k+p} \in \overset{\circ}{S}_{i_1,\dots,i_k}\}) = P_n(S_{i_1,\dots,i_k})$. Furthermore, every open set $O \subset \mathcal{X}$ can be defined as the union of a countable disjoint family of sets $S_{i_1,\dots,i_k}$. Then by Fatou's lemma, it follows that $\liminf_{p\to\infty} P(X_n^p)^{-1}(O) \geq P_n(O)$ for every open set $O \subset \mathcal{X}$. Therefore, by virtue of Theorem 2.1, we have $P(X_n^p)^{-1} \Rightarrow P_n$ as $p \to \infty$, which implies that $PX_n^{-1} = P_n$. Similarly, we also get $PX^{-1} = \mathcal{P}$. □

Consider now the case $\mathcal{X} = C$, where C is the space of all continuous functions $x : [0,\infty) \to \mathbb{R}^d$ with a metric ρ defined by $\rho(x_1, x_2) = \sum_{n=1}^{\infty} 2^{-n}[1 \wedge \max_{0\leq t\leq n} |x_1(t) - x_2(t)|]$ for $x_1, x_2 \in C$. It can be verified that (C, ρ) is a complete separable metric space. We prove the following theorem.

Theorem 2.4. *Let $(X_n)_{n=1}^{\infty}$ be a sequence of C-random variables X_n on a probability space $(\Omega_n, \mathcal{F}_n, P_n)$ for $n = 1, 2, \dots$ such that*

(i) $\lim_{N\to\infty} \sup_{n\geq 1} P_n(\{|X_n(0)| > N\}) = 0$ *and*
(ii) $\lim_{h\downarrow 0} \sup_{n\geq 1} P_n(\{\max_{t,s\in[0,T],|t-s|\leq h} |X_n(t) - X_n(s)| > \varepsilon\}) = 0$

for every $T > 0$ and $\varepsilon > 0$. Then there exist an increasing subsequence $(n_k)_{k=1}^{\infty}$ of $(n)_{n=1}^{\infty}$, a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$, and C-random variables $\tilde{X}_{n_k}$ and $\tilde{X}$ for $k = 1, 2, \dots$ on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ such that $PX_{n_k}^{-1} = P\tilde{X}_{n_k}^{-1}$ for $k = 1, 2, \dots$ and $\rho(\tilde{X}_{n_k}, \tilde{X}) \to 0$ a.s. as $k \to \infty$.

Proof. We shall show that conditions (i) and (ii) imply that the set $\Lambda = \{PX_n^{-1} : n \geq 1\}$ is a tight subset of $\mathcal{M}(C)$. Let us recall that by the Arzelà–Ascoli theorem, a set $A \subset C$ is relatively compact in (C, ρ) if and only if the following conditions are satisfied:

(I) A is uniformly bounded, i.e., $\sup_{x\in A} \max_{t\in[0,T]} |x(t)| < \infty$ for every $T > 0$,
(II) A is uniformly equicontinuous, i.e., $\lim_{h\downarrow 0} \sup_{x\in A} V_h^T(x) = 0$ for every $T > 0$,

where $V_h^T(x) = \max_{t,s\in[0,T],|t-s|\leq h} |X_n(t) - X_n(s)|$. By virtue of (i), for every $\varepsilon > 0$, there exists a number $a > 0$ such that $PX_n^{-1}(\{x : |x(0)| \leq a\}) > 1 - \varepsilon/2$ for $n \geq 1$. By (ii), for every $\varepsilon > 0$ and $k = 1, 2, \dots$ there exists $h_k > 0$ such that $h_k \downarrow 0$ and $PX_n^{-1}(\{x : V_{h_k}^T(x) > 1/k\}) \leq \varepsilon/2^{k+1}$ for every $n \geq 1$. Therefore, we have $PX_n^{-1}(\bigcap_{k=1}^{\infty}\{x : V_{h_k}^T(x) \leq 1/k\}) > 1 - \varepsilon/2$. Taking $K_\varepsilon = \{x \in C : |x(0)| \leq a\} \cap \left(\bigcap_{k=1}^{\infty}\{x : V_{h_k}^T(x) \leq 1/k\}\right)$, we can easily see that K_ε satisfies conditions (I) and (II). Therefore, K_ε is a compact subset of C such that $PX_n^{-1}(K_\varepsilon) > 1 - \varepsilon$ for $n \geq 1$. Then the set Λ is a tight subset of $\mathcal{M}(C)$. Hence, by virtue of Theorems 2.2 and 2.3, there exist an increasing subsequence $(n_k)_{k=1}^{\infty}$ of $(n)_{n=1}^{\infty}$, a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$, and C-random variables $\tilde{X}_{n_k}$ and $\tilde{X}$ for $k = 1, 2, \dots$ on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ such that conditions (i) and (ii) of Theorem 2.3 are satisfied, i.e., such that $PX_{n_k}^{-1} = P\tilde{X}_{n_k}^{-1}$ and $\rho(\tilde{X}_{n_k}, \tilde{X}) \to 0$ a.s. as $k \to \infty$. □

Remark 2.1. Theorem 2.4 holds if instead of a condition (ii) of this theorem, the following condition is satisfied for every $\varepsilon > 0$:

$$\lim_{\delta\to 0}\sup_{n\geq 1}\sum_{j<\delta^{-1}} P\left(\left\{\max_{j\delta\leq s\leq (j+1)\delta}|X_n(s)-X(j\delta)|>\varepsilon\right\}\right)=0. \tag{2.1}$$

Proof. Let us note that Theorem 2.4 holds if we consider its condition (ii) with 3ε instead of ε. For every $\delta\in(0,1)$ and $s,t\in[0,T]$ such that $|t-s|<\delta$ there is an integer $0\leq j<\delta^{-1}$ such that $s,t\in[j\delta,(j+1)\delta]$ or $s\wedge t\in[j\delta,(j+1)\delta]$ or $s\vee t\in[j\delta,(j+1)\delta]$. Therefore, for every $\delta\in(0,1)$ and $s,t\in[0,T]$ such that $|t-s|<\delta$ one has $s,t\in\bigcup_{0\leq j<\delta^{-1}}[j\delta,(j+1)\delta]$. Thus for every $\delta\in(0,1)$ and $n\geq 1$ we have

$$\begin{aligned}
&\max_{s,t\in[0,T],|t-s|<\delta}|X_n(s)-X_n(t)|\\
&\leq \sup\left\{|X_n(s)-X_n(t)| : s,t\in\bigcup_{j<\delta^{-1}}[j\delta,(j+1)\delta]\right\}\\
&\leq \sup\left\{|X_n(s)-X_n(j\delta)| : s\in\bigcup_{j<\delta^{-1}}[j\delta,(j+1)\delta]\right\}\\
&\quad+\sup\left\{|X_n(t)-X_n(j\delta)| : t\in\bigcup_{j<\delta^{-1}}[j\delta,(j+1)\delta]\right\}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
&P\left(\left\{\max_{s,t\in[0,T],|t-s|<\delta}|X_n(s)-X_n(t)|>3\varepsilon\right\}\right)\\
&\leq\sum_{j<\delta^{-1}}P\left(\left\{\max_{j\delta\leq s\leq(j+1)\delta}|X_n(s)-X_n(j\delta)|>\varepsilon\right\}\right).
\end{aligned}$$

Then condition (ii) of Theorem 2.4 is satisfied for every $\varepsilon>0$ if condition (2.1) is satisfied. □

Remark 2.2. If X and Y are given random variables defined on a probability space $(\Omega,\mathcal{F},P)$ with values in a metric space $(\mathcal{X},\rho)$, then X and Y are said to have equivalent distributions if $PX^{-1}(A)=0$ if and only if $PY^{-1}(A)=0$ for $A\in\beta(\mathcal{X})$. □

In what follows, we shall need the following results.

Lemma 2.1. *Let $(\mathcal{X},\rho)$ and $(Y,\mathcal{G})$ be a metric and a measurable space, respectively, and $\Phi : X\to Y$ a $(\beta(\mathcal{X}),\mathcal{G})$-measurable mapping, where $\beta(\mathcal{X})$ is a Borel σ-algebra on $\mathcal{X}$. If X and $\tilde{X}$ are $\mathcal{X}$-random variables defined on a probability* *space* $(\Omega, \mathcal{F} P)$ *and* $\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$, *respectively, such that* $PX^{-1} = P\tilde{X}^{-1}$, *then* $P(\Phi \circ X)^{-1} = P(\Phi \circ \tilde{X})^{-1}$.

Proof. Let $Z = \Phi \circ X$ and $\tilde{Z} = \Phi \circ \tilde{X}$. For every $A \in \mathcal{G}$ one has $P(\{Z \in A\}) = P(\{\Phi \circ X \in A\}) = P(X^{-1}(\Phi^{-1}(A))) = \tilde{P}(\tilde{X}^{-1}(\Phi^{-1}(A))) = \tilde{P}(\{\Phi \circ \tilde{X} \in A\}) = \tilde{P}(\{\tilde{Z} \in A\})$. Then $P(\Phi \circ X)^{-1} = P(\Phi \circ \tilde{X})^{-1}$. □

Lemma 2.2. *Let* $(\mathcal{X}, \rho)$ *and* (Y, d) *be metric spaces, and let* X^n *and* X *be* $\mathcal{X}$*-random variables defined on probability spaces* $(\Omega_n, \mathcal{F}_n P_n)$ *and* $(\Omega, \mathcal{F} P)$, *respectively for* $n = 1, 2, \ldots$ *such that* $X^n \Rightarrow X$ *as* $n \to \infty$. *For every continuous mapping* $\Phi : \mathcal{X} \to Y$ *one has* $\Phi \circ X^n \Rightarrow \Phi \circ X$ *as* $n \to \infty$.

Proof. By virtue of Theorem 2.1, for every open set $G \subset \mathcal{X}$ one has $\liminf_{n\to\infty} P(X^n)^{-1}(G) \geq PX^{-1}(G)$. By continuity of Φ, for every open set $\mathcal{U} \subset Y$, a set $\Phi^{-1}(\mathcal{U})$ is an open set of $\mathcal{X}$. Taking in particular in the above inequality $G = \Phi^{-1}(\mathcal{U})$, we obtain $\liminf_{n\to\infty} P(X^n)^{-1}(\Phi^{-1}(\mathcal{U})) \geq PX^{-1}(\Phi^{-1}(\mathcal{U}))$. But $P(X^n)^{-1}(\Phi^{-1}(\mathcal{U})) = P_n[(X^n)^{-1}(\Phi^{-1}(\mathcal{U}))] = P(\Phi \circ X^n)^{-1}(\mathcal{U})$ and $PX^{-1}(\Phi^{-1}(\mathcal{U})) = P[X^{-1}(\Phi^{-1}(\mathcal{U}))] = P(\Phi \circ X)^{-1}(\mathcal{U})$ for every open set $\mathcal{U} \subset Y$. Therefore, for every open set $\mathcal{U} \subset Y$ one has $\liminf_{n\to\infty} P(\Phi \circ X^n)^{-1}(\mathcal{U}) \geq P(\Phi \circ X)^{-1}(\mathcal{U})]$, which by Theorem 2.1 and the definition of weak convergence of sequences of random variables implies that $\Phi \circ X^n \Rightarrow \Phi \circ X$ as $n \to \infty$. □

3 Stochastic Processes

Throughout this section we assume that $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ is a complete filtered probability space with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ satisfying the usual conditions. We shall consider a family $X = (X_t)_{t\geq 0}$ of $\mathcal{X}$-random variables X_t on $\mathcal{P}_{\mathbb{F}}$ with $\mathcal{X} = \mathbb{R}$ or $\mathcal{X} = \mathbb{R}^d$. Such families are called one- or d-dimensional stochastic processes on $\mathcal{P}_{\mathbb{F}}$. It is easy to see that such stochastic processes can be regarded as functions $X : \mathbb{R}^+ \times \Omega \to \mathbb{R}$ and $X : \mathbb{R}^+ \times \Omega \to \mathbb{R}^d$, respectively, such that $X(t, \cdot)$ is an $\mathbb{R}$- or $\mathbb{R}^d$-random variable. We can also consider stochastic processes with the index set $I \subset \mathbb{R}^+$ instead of $\mathbb{R}^+$. If $I = \mathbb{N}$, we call X a discrete stochastic process on $\mathcal{P}_{\mathbb{F}}$. Given a d-dimensional stochastic process $X = (X_t)_{t\geq 0}$ on $\mathcal{P}_{\mathbb{F}}$ and fixed $\omega \in \Omega$, we call a mapping $\mathbb{R}^+ \ni t \to X_t(\omega) \in \mathbb{R}^d$ a trajectory or a path of X corresponding to $\omega \in \Omega$. We can characterize stochastic processes by properties of their trajectories. In particular, a process $X = (X_t)_{t\geq 0}$ defined on $\mathcal{P}_{\mathbb{F}}$ is said to be:

1. Continuous if almost all its paths are continuous on $\mathbb{R}^+$.
2. Right (left) continuous on $\mathbb{R}^+$ if almost all its paths are right (left) continuous on $\mathbb{R}^+$.
3. A càdlàg process if it is right continuous and almost all its paths have at every $t > 0$ a left limit $\lim_{s\to t, s<t} X_s$.

4. A càglàd process if it is left continuous and almost all its paths have at every $t \geq 0$ a right limit $\lim_{s \to t, s > t} X_s$.

Stochastic processes $X = (X_t)_{t \geq 0}$ and $Y = (Y_t)_{t \geq 0}$ defined on $\mathcal{P}_{\mathbb{F}}$ are called:

5. Indistinguishable if $P(\{X_t = Y_t \ : t \geq 0\}) = 1$.
6. Y is a modification of X if $P(\{X_t = Y_t\}) = 1$ for every $t \geq 0$.

The properties of the above types of "equivalence" of two stochastic processes are quite different. If X and Y are modifications, then for every $t \geq 0$, there exists a null set $\Omega_t \subset \Omega$ such that if $\omega \notin \Omega_t$, then $X_t(\omega) = Y_t(\omega)$. Since the interval $[0, \infty)$ is uncountable, the set $\Lambda = \bigcup_{t \geq 0} \Omega_t$ could have any probability between 0 and 1, and it could be even unmeasurable. If X and Y are indistinguishable, however, then there exists a null set $\Lambda \subset \Omega$ such that if $\omega \notin \Lambda$, then $X_t(\omega) = Y_t(\omega)$ for all $t \geq 0$. In other words, the paths of X and Y are the same for all $\omega \notin \Lambda$. We have $\Lambda \in \mathcal{F}_0 \subset \mathcal{F}_t$ for all $t \geq 0$. In some special cases, the above types of "equivalence" are equivalent.

Theorem 3.1. *Let X and Y be two stochastic processes, with X a modification of Y. If X and Y are right continuous, then they are indistinguishable.*

Proof. Let $\Omega_0 \subset \Omega$ be such that all paths of X and Y corresponding to $\omega \in \Omega \setminus \Omega_0$ are right continuous on $\mathbb{R}^+$ and $P(\Omega_0) = 0$. Let $\Lambda_t = \{X_t \neq Y_t\}$ and $\Lambda = \bigcup_{t \in Q} \Lambda_t$, where Q denotes the set of all rational numbers of $\mathbb{R}^+$. We have $P(\Lambda) = 0$ and $P(\Omega_0 \cup \Lambda) = 0$. Then $X_t(\omega) = Y_t(\omega)$ for $t \in Q$ and $\omega \notin \Omega_0 \cup \Lambda$. For fixed $t \in \mathbb{R}^+$, we can select a sequence $(t_n)_{n=1}^{\infty}$ of Q such that $t_n \to t$ as $n \to \infty$. We can assume that the t_n decrease to t through Q. Then we get $X_t(\omega) = \lim_{n \to \infty} X_{t_n}(\omega) = \lim_{n \to \infty} Y_{t_n}(\omega) = Y_t(\omega)$ for $\omega \notin \Omega_0 \cup \Lambda$ and every $t \geq 0$. □

A d-dimensional stochastic process $X = (X_t)_{t \geq 0}$ on $\mathcal{P}_{\mathbb{F}}$ is said to be:

(i) $\mathbb{F}$-adapted if X_t is $(\mathcal{F}_t, \beta(\mathbb{R}^d))$-measurable for every $t \geq 0$.
(ii) Measurable if a mapping $X : \mathbb{R}^+ \times \Omega \to \mathbb{R}^d$ defined by $X(t, \omega) = X_t(\omega)$ for $(t, \omega) \in \mathbb{R}^+ \times \Omega$ is $(\beta(\mathbb{R}^+) \otimes \mathcal{F}, \beta(\mathbb{R}^d))$-measurable.
(iii) $\mathbb{F}$-nonanticipative if it is measurable and $\mathbb{F}$-adapted.
(iv) $\mathbb{F}$-progressively measurable if for all $t \geq 0$, a restriction to $I_t \times \Omega$ of a mapping $X : \mathbb{R}^+ \times \Omega \to \mathbb{R}^d$ defined in (ii) with $I_t = [0, t]$ is $(\beta(I_t) \otimes \mathcal{F}_t, \beta(\mathbb{R}^d))$-measurable.
(v) $\mathbb{F}$-predictable or simply predictable if it is measurable with respect to a σ-algebra $\mathcal{P}(\mathbb{F})$ generated by all $\mathbb{F}$-adapted càglàd processes on $\mathcal{P}_{\mathbb{F}}$.
(vi) $\mathbb{F}$-optional or simply optional if it is measurable with respect to a σ-algebra $\mathcal{O}(\mathbb{F})$ generated by all $\mathbb{F}$-adapted càdlàg processes on $\mathcal{P}_{\mathbb{F}}$.

It can be verified that $\mathcal{P}(\mathbb{F}) \subset \mathcal{O}(\mathbb{F}) \subset \beta(\mathbb{R}^+) \otimes \mathcal{F}$. Therefore, each predictable process is optional, and both are measurable. It is clear that every $\mathbb{F}$-progressively measurable process is $\mathbb{F}$-nonanticipative. Let us note that for a given stochastic process $X = (X_t)_{t \geq 0}$ on $\mathcal{P}_{\mathbb{F}}$, we may identify each $\omega \in \Omega$ with its path $\mathbb{R}^+ \ni t \to X_t(\omega) \in \mathbb{R}^d$. Thus we may regard Ω as a subset of the space $\tilde{\Omega} = (\mathbb{R}^d)^{[0,\infty)}$

of all functions from $[0,\infty)$ into $\mathbb{R}^d$. Then the σ-algebra $\mathcal{F}$ will contain the σ-algebra $\mathcal{B}$, generated by sets $\{\omega \in \Omega : \omega(t_1) \in A_1, \dots, \omega(t_k) \in A_k\}$ for all $t_1, \dots, t_k \in \mathbb{R}^+$ and all Borel sets $A_i \subset \mathbb{R}^d$ for $i = 1, 2, \dots, k$ and $k \in \mathbb{N}$. The space $(\mathbb{R}^d)^{[0,\infty)}$ contains some important subspaces such as $\mathcal{C} = C(\mathbb{R}^+, \mathbb{R}^d)$, $\mathcal{C}_+ = C_+(\mathbb{R}^+, \mathbb{R}^d)$, and $\mathcal{C}_- = C_-(\mathbb{R}^+, \mathbb{R}^d)$ of respectively all continuous, right continuous, and left continuous functions $x : \mathbb{R}^+ \to \mathbb{R}^d$. A special role in such an approach to stochastic processes is played by an evaluation mapping defined for every fixed $t \geq 0$ by setting $e_t : (\mathbb{R}^d)^{[0,\infty)} \ni x \to x(t) \in \mathbb{R}^d$. We can define on the space $\mathcal{X} = (\mathbb{R}^d)^{[0,\infty)}$ a σ-algebra of cylindrical sets, denoted by $\mathcal{G}(\mathcal{X})$, as a σ-algebra generated by a family $\{e_t : t \geq 0\}$, i.e., $\mathcal{G}(\mathcal{X}) = \sigma(\{e_t : t \geq 0\})$. In a similar way, we can define a filtration $(\mathcal{G}_t)_{t\geq 0}$ by taking $\mathcal{G}_t = \sigma(\{e_s : 0 \leq s \leq t\})$. We have the following important result.

Theorem 3.2. *The σ-algebra $\mathcal{G}(\mathcal{C})$ of cylindrical sets of $\mathcal{C}$ coincides with the σ-algebra $\beta(\mathcal{C})$ of Borel sets of $\mathcal{C}$.*

Proof. We have only to verify that $\beta(\mathcal{C}) \subset \mathcal{G}(\mathcal{C})$. Let us observe that a family of sets $\{x \in \mathcal{C} : \max_{0\leq t\leq n} |x(t) - x_0(t)| \leq \varepsilon\}$ with fixed $x_0 \in \mathcal{C}$, $\varepsilon > 0$ and $n = 1, 2, \dots$ is a base of neighborhoods in $\mathcal{C}$. On the other hand, we have $\{x \in \mathcal{C} : \max_{0\leq t\leq n} |x(t) - x_0(t)| \leq \varepsilon\} = \bigcap_{r\in Q, 0\leq r\leq n}\{x \in \mathcal{C} : x(r) \in \mathcal{U}(x_0(r), \varepsilon)\}$, where $\mathcal{U}(a, \varepsilon) = \{x \in \mathbb{R}^d : |x - a| \leq \varepsilon\}$. Therefore, $\{x \in \mathcal{C} : \max_{0\leq t\leq n} |x(t) - x_0(t)| \leq \varepsilon\} \in \mathcal{G}(\mathcal{C})$, which implies that $\beta(\mathcal{C}) \subset \mathcal{G}(\mathcal{C})$. □

Remark 3.1. The above result is also true for the space $\mathcal{D}$ of all d-dimensional càdlàg functions on $[0,\infty)$, i.e., $\beta(\mathcal{D}) = \mathcal{G}(\mathcal{D})$, where $\mathcal{G}(\mathcal{D})$ denotes the σ-algebra of cylindrical sets of $\mathcal{D}$. □

Corollary 3.1. *A stochastic process $X = (X_t)_{t\geq 0}$ on $\mathcal{P}_{\mathbb{F}}$ can be regarded as an $(\mathbb{R}^d)^{[0,\infty)}$-random variable on $\mathcal{P}_{\mathbb{F}}$, i.e., as a mapping from Ω into $(\mathbb{R}^d)^{[0,\infty)}$ that is $(\mathcal{F}, \mathcal{G}(\mathcal{X}))$-measurable. In particular, by virtue of Theorem 3.2 and Remark 3.1, a d-dimensional continuous (càdlàg) process $X = (X_t)_{t\geq 0}$ on $\mathcal{P}_{\mathbb{F}}$ can be considered as a mapping from Ω into $\mathcal{C}$ ($\mathcal{D}$) that is $(\mathcal{F}, \beta(\mathcal{C}))$- ($(\mathcal{F}, \beta(\mathcal{D}))$)-measurable.* □

Remark 3.2. Given a d-dimensional continuous (càdlàg) stochastic process $X = (X_t)_{t\geq 0}$ on $\mathcal{P}_{\mathbb{F}}$ by PX^{-1}, we denote the distribution of $\mathcal{C}$-random ($\mathcal{D}$-random) variable $X : \Omega \to \mathcal{C}$ ($X : \Omega \to \mathcal{D}$), i.e., a probability measure defined by $(PX^{-1})(A) = P(X^{-1}(A))$ for $A \in \beta(\mathcal{C})$ ($A \in \beta(\mathcal{D})$). □

Corollary 3.2. *Let $X = (X_t)_{t\geq 1}$ and $\tilde{X} = (\tilde{X}_t)_{t\geq 1}$ be d-dimensional continuous stochastic processes on probability spaces $(\Omega, \mathcal{F}, P)$ and $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$, respectively, such that $PX^{-1} = P\tilde{X}^{-1}$. For every $(s, x) \in \mathbb{R}^+ \times \mathbb{R}^d$ such that $X_s = x$, P-a.s., one has $\tilde{X}_s = x$, $\tilde{P}$-a.s.*

Proof. The result follows immediately from Lemma 2.1. Indeed, assume that there is $(s, x) \in \mathbb{R}^+ \times \mathbb{R}^d$ such that $X_s = x$, P-a.s. Taking, in particular, $\mathcal{X} = C(\mathbb{R}^+, \mathbb{R}^d)$, $Y = \mathbb{R}^d$, and $\Phi = e_s$ in Lemma 2.1, where e_s is an evolution mapping corresponding to $s \geq 0$, we obtain $P(e_s \circ X)^{-1} = P(e_s \circ \tilde{X})^{-1}$. Hence, for $A_x = \{x\} \subset \mathbb{R}^d$, it follows that $P(e_s \circ X)^{-1}(A_x) = P(e_s \circ \tilde{X})^{-1}(A_x)$. Then $P(\{X_s = x\}) = \tilde{P}(\{\tilde{X}_s = x\})$, which implies that $\tilde{P}(\{\tilde{X}_s = x\}) = 1$. □

Corollary 3.3. *Let* $X^n = (X_t^n)_{t\geq 0}$ *and* $X = (X_t)_{t\geq 0}$ *be* d*-dimensional continuous stochastic processes on probability spaces* $(\Omega_n, \mathcal{F}_n, P_n)$ *and* $(\Omega, \mathcal{F}, P)$,, *respectively, for* $n = 1, 2, \ldots$. *If a sequence* $(X^n)_{n=1}^{\infty}$ *converges weakly in distribution to* X, *then* $X_s^n \Rightarrow X_s$ *as* $n \to \infty$ *for every* $s \geq 0$.

Proof. The result follows immediately from Lemma 2.2. Indeed, assume that a sequence $(X^n)_{n=1}^{\infty}$ converges weakly in distribution to X, and let $s \geq 0$. Taking, in particular, $\mathcal{X} = C(\mathbb{R}^+, \mathbb{R}^d)$, $Y = \mathbb{R}^d$, and $\Phi = e_s$ in Lemma 2.2, one obtains $e_s \circ X^n \Rightarrow e_s \circ X$ as $n \to \infty$. Then $X_s^n \Rightarrow X_s$ as $n \to \infty$. □

Remark 3.3. A finite-dimensional distribution of a d-dimensional stochastic process $X = (X_t)_{t\geq 0}$ on $\mathcal{P}_{\mathbb{F}}$ is defined as a probability measure $\mu_{t_1,\ldots,t_k}$ on $\beta(\mathbb{R}^{kd})$ for $k = 1, 2, \ldots$ defined by $\mu_{t_1,\ldots,t_k}(A_1 \times \cdots \times A_k) = P(\{X_{t_1} \in A_1, \ldots, X_{t_k} \in A_k\})$ for $t_i \in [0, \infty)$ and $A_i \in \beta(\mathbb{R}^d)$ for $i = 1, 2, \ldots, k$. □

Remark 3.4. If d-dimensional continuous (càdlàg) stochastic processes $X = (X_t)_{t\geq 0}$ and $\tilde{X} = (\tilde{X}_t)_{t\geq 0}$ on $\mathcal{P}_{\mathbb{F}}$ and $\tilde{\mathcal{P}}_{\tilde{\mathbb{F}}}$, respectively, have the same distributions, then $PX^{-1} = P\tilde{X}^{-1}$ is equivalent to $\mu_{t_1,\ldots,t_k}(A_1 \times \cdots \times A_k) = \tilde{\mu}_{t_1,\ldots,t_k}(A_1 \times \cdots \times A_k)$ for every $t_i \in [0, \infty)$ and $A_i \in \beta(\mathbb{R}^d)$ for $i = 1, 2, \ldots, k$. □

We have the following important theorems due to Kolmogorov.

Theorem 3.3 (Extension theorem). *Let* $\mu_{t_1,\ldots,t_k}$ *be for all* $t_1, \ldots, t_k \in [0, \infty)$ *and* $k \in \mathbb{N}$ *a probability measure on* $\beta(\mathbb{R}^{kd})$ *such that (i)* $\mu_{t_{\sigma(1)},\ldots,t_{\sigma(k)}}(A_{\sigma(1)} \times \cdots \times A_{\sigma(k)}) = \mu_{t_1,\ldots,t_k}(A_1 \times \cdots \times A_k)$ *for all permutations* $\sigma = (\sigma(1), \ldots, \sigma(k))$ *of* $\{1, 2, \ldots, k\}$ *and (ii)* $\mu_{t_1,\ldots,t_k}(A_1 \times \cdots \times A_k) = \mu_{t_1,\ldots,t_k,t_{k+1},\ldots,t_{k+m}}(A_1 \times \cdots \times A_k \times \underbrace{\mathbb{R}^d \times \cdots \times \mathbb{R}^d}_{m})$ *for all* $m \in \mathbb{N}$. *Then there exist a probability space* $(\Omega, \mathcal{F}, P)$ *and a* d*-dimensional stochastic process* $X = (X_t)_{t\geq 0}$ *on* $(\Omega, \mathcal{F}, P)$ *such that* $\mu_{t_1,\ldots,t_k}(A_1 \times \cdots \times A_k) = P(\{X_{t_1} \in A_1, \ldots, X_{t_k} \in A_k\})$ *for* $t_i \in [0, \infty)$ *and* $A_i \in \beta(\mathbb{R}^d)$ *with* $i = 1, 2, \ldots, k$ *and* $k \in \mathbb{N}$. □

Theorem 3.4 (Existence of continuous modification). *Suppose a* d*-dimensional stochastic process* $X = (X_t)_{t\geq 0}$ *on* $\mathcal{P}_{\mathbb{F}}$ *is such that for all* $T > 0$, *there exist positive constants* α, β, *and* γ *such that*

$$E[|X_t - X_s|^{\alpha}] \leq \gamma |t - s|^{1+\beta}$$

for $s, t \in [0, T]$. *Then there exists a continuous modification of* X. □

We shall now prove the following theorem.

Theorem 3.5. *Let* $(X^n)_{n=1}^{\infty}$ *be a sequence of* d*-dimensional continuous stochastic processes* $X^n = (X_t^n)_{t\geq 0}$ *on a probability space* $(\Omega_n, \mathcal{F}_n, P_n)$ *for* $n = 1, 2, \ldots$ *such that:*