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
$E[\Phi\|\mathcal{G}]$	– $\mathcal{G}$-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 $A_n = \{x \in \mathbb{R} : \operatorname{dist}(x, \Lambda) < 1/n\}$. 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 $\operatorname{cad}(\mathbb{F})$ 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 $\mathcal{F}_T = \sigma(\{X_T : X \in \operatorname{cad}(\mathbb{F})\})$.*

Proof. Let $\mathcal{G}_T = \sigma(\{X_T : X \in \text{cad}(\mathbb{F})\})$ 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 $X \in \text{cad}(\mathbb{F})$, which implies that $A \in \mathcal{G}_T$. Then $\mathcal{F}_T \subset \mathcal{G}_T$.

Let $X \in \text{cad}(\mathbb{F})$. 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 $X \in \text{cad}(\mathbb{F})$, 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 $f_k(x) = \psi(k \cdot \text{dist}(x, F))$ 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 $X_n \xrightarrow{\text{P}} X$ 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° $\mathrm{diam}(S_{i_1,\ldots,i_k}) \leq 2^{-k}$ 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 $n \geq 1$.

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 $k, n, p \geq 1$ we have $\rho(X_n^k(\omega), X_n^{k+p}(\omega)) \leq 1/2^k$ and $\rho(X^k(\omega), X^{k+p}(\omega)) \leq 1/2^k$. 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 $n \geq n_k$. Then $X_n^k(\omega) = X^k(\omega)$ and therefore,

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

for $n \geq n_k$. 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 $n = 1, 2, \ldots$ and $PX^{-1} = \mathcal{P}$. Let us first observe that $\tilde{P}(\{X_n^{k+p} \in \overline{S}_{i_1,\ldots,i_k}\}) = \tilde{P}(\{X_n^{k+p} \in \overset{\circ}{S}_{i_1,\ldots,i_k}\}) = P_n(S_{i_1,\ldots,i_k})$. Furthermore, every open set $O \subset \mathcal{X}$ can be defined as the union of a countable disjoint family of sets $S_{i_1,\ldots,i_k}$. 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 $n = 1, 2, \ldots$ 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 $k = 1, 2, \ldots$ on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ such that $PX_{n_k}^{-1} = P\tilde{X}_{n_k}^{-1}$ for $k = 1, 2, \ldots$ 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 $k = 1, 2, \ldots$ 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 $k = 1, 2, \ldots$ 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 $\tilde{P}(\{\tilde{X}_s = x\}) = 1$. $\square$

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 $n \to \infty$. $\square$

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 $i = 1, 2, \ldots, k$. $\square$

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 $i = 1, 2, \ldots, k$. $\square$

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 $k \in \mathbb{N}$.* $\square$

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. $\square$

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:*

(i) There exist positive numbers M *and* γ *such that* $E_n[|X_0^n|^\gamma] \le M$ *for* $n = 1, 2, \ldots$.

(ii) For every $T > 0$, *there exist* $M_T > 0$ *and positive numbers* α, β *independent of* $T > 0$ *such that* $E_n[|X_t^n - X_s^n|^\alpha] \le M_T|t-s|^{1+\beta}$ *for* $n = 1, 2, \ldots$ *and* $t, s \in [0, T]$.

Then $(X^n)_{n=1}^\infty$ *satisfies conditions (i) and (ii) of Theorem 2.4.*

Proof. By virtue of Chebyshev's inequality, we get $P(\{|X_0^n| > N\}) \le M/N^\gamma$ for $n = 1, 2, \ldots$. Therefore, condition (i) of Theorem 2.4 is satisfied. For simplicity, we assume now that $T > 0$ is a positive integer. By (ii), the process $Y = (Y(t))_{t\ge 0}$ defined by $Y(t) = X_t^n$ for fixed $n = 1, 2, \ldots$ satisfies $E_n[|Y(t) - Y(s)|^\alpha] \le M_T|t-s|^{1+\beta}$ for $t, s \in [0, T]$. Hence, by Chebyshev's inequality applied to every $a > 0$, it follows that

$$\begin{aligned} P_n(\{|Y((i+1)/2^m) - Y(i/2^m)| > 1/2^{ma}\}) &\le M_T 2^{-m(1+\beta)}2^{ma\alpha} \\ &= M_T 2^{-m(1+\beta-a\alpha)} \end{aligned}$$

for $i = 0, 1, 2, \ldots, 2^m T - 1$. Taking now a number a such that $0 < a < \beta/\alpha$, one obtains

$$P_n\left(\left\{\max_{0\le i\le 2^m T-1} |Y((i+1)/2^m) - Y(i/2^m)| > 1/2^{ma}\right\}\right) \le TM_T 2^{-m(\beta-a\alpha)}.$$

Let $\varepsilon > 0$ and $\delta > 0$, and select $\nu = \nu(\delta, \varepsilon)$ such that $(1 + 2/(2^a - 1))/2^{a\nu} \le \varepsilon$ and $\sum_{m=\nu}^\infty 2^{-m(\beta-a\alpha)} < \delta/TM_T$. We get

$$\begin{aligned} P_n\left(\bigcup_{m=\nu}^\infty \left\{\max_{0\le i\le 2^m T-1} |Y((i+1)/2^m) - Y(i/2^m)| > 1/2^{ma}\right\}\right) \\ \le TM_T \sum_{m=\nu}^\infty 2^{-m(\beta-a\alpha)} < \delta. \end{aligned}$$

Put $\Omega_\nu = \bigcup_{m=\nu}^\infty \{\max_{0\le i\le 2^m T-1} |Y((i+1)/2^m) - Y(i/2^m)| > 1/2^{ma}\}$. We have $P_n(\Omega_\nu) < \delta$ and if $\omega \notin \Omega_\nu$ then $|Y((i+1)/2^m) - Y(i/2^m)| \le 1/2^{ma}$ for $m \ge \nu$ and all $i = 0, 1, 2\ldots$ such that $(i+1)/2^m \le T$. Let D_T be the set of all dyadic rational numbers of $[0, T]$. If $s \in D_T \cap [i/2^\nu, (i+1)/2^\nu)$, then it can be expressed by the formula $s = i/2^\nu + \sum_{l=1}^j \alpha_l/2^{\nu+1}$ with $\alpha_l \in \{0, 1\}$. Therefore, for such s and $\omega \notin \Omega_\nu$, one has

$$\begin{aligned} |Y(s) - Y(i/2^\nu)| &\le \sum_{k=1}^j \left| Y\left(i/2^\nu + \sum_{l=1}^k \alpha_l/2^{\nu+l}\right) - Y\left(i/2^\nu + \sum_{l=1}^{k-1} \alpha_l/2^{\nu+l}\right)\right| \\ &\le \sum_{k=1}^j \alpha_l/2^{(\nu+k)a} \le \sum_{k=1}^\infty \alpha_l/2^{(\nu+k)a} = 1/(2^a-1)2^{a\nu}. \end{aligned}$$

Therefore, for $\omega \notin \Omega_\nu$ and $s, t \in D_T$ satisfying $|s-t| \leq 1/2^\nu$, we get

$$|Y(s) - Y(t)| \leq \left(1 + \frac{2}{2^a - 1}\right)/2^{a\nu} \leq \varepsilon.$$

Indeed, if $t \in [(i-1)/2^\nu, i/2^\nu)$ and $s \in [i/2^\nu, (i+1)/2^\nu)$, then

$$|Y(s) - Y(t)| \leq |Y(s) - Y(i/2^\nu)| + |Y(t) - Y((i-1)/2^\nu)|$$
$$+|Y(i/2^\nu) - Y((i-1)/2^\nu)| \leq \left(1 + \frac{2}{2^a - 1}\right)/2^{a\nu}.$$

If $t, s \in [i/2^\nu, (i+1)/2^\nu)$, then

$$|Y(s) - Y(t)| \leq |Y(s) - Y(i/2^\nu)| + |Y(t) - Y(i/2^\nu)| \leq \frac{2}{(2^a - 1)2^{a\nu}}.$$

But D_T is dense in $[0, T]$ and $|Y(s) - Y(t)| \leq \varepsilon$ for every $s, t \in D_T$. Then for every $s, t \in [0, T]$ satisfying $|s-t| \leq 1/2^\nu$, we also have $P_n(\{\max_{t,s\in[0,T],|t-s|\leq 1/2^\nu} |Y(s) - Y(t)| > \varepsilon\}) \leq P_n(\Omega_\nu) < \delta$. But $\nu = \nu(\delta, \varepsilon)$ does not depend on n. Therefore, this implies that condition (ii) of Theorem 2.4 is also satisfied. □

There are some weaker sufficient conditions for relative weak compactness of sequences of continuous stochastic processes. We shall show here that for a given sequence $(X^n)_{n=1}^\infty$ of d-dimensional continuous stochastic processes $(X_t^n)_{t\geq 0}$ defined on a probability space $(\Omega, \mathcal{F}, P)$ such that the sequence $(\mu_n)_{n=1}^\infty$ of probability distributions μ_n of X_0^n is tight, then the sequence $(P(X^n)^{-1})_{n=1}^\infty$ of distributions of X^n is tight if there are numbers $\gamma \geq 0$, $\alpha > 1$ and a real-valued continuous nondecreasing stochastic process $(\Gamma(t))_{t\geq 0}$ such that $E[\Gamma(T) - \Gamma(0)] < \infty$ and $P(\{|X_t^n - X_s^n| \geq \lambda\} \leq 1/\lambda^\gamma E[|\Gamma(t) - \Gamma(s)|^\alpha]$ for every $T > 0$, $s, t \in [0, T]$, and $\lambda > 0$. To begin with, let us introduce the following notation. Given a probability space $\mathcal{P} = (\Omega, \mathcal{F}, P)$ and random variables $\xi_i : \Omega \to \mathbb{R}^d$ for $i = 1, 2, \ldots, n$, let us define $S_k = \xi_1 + \cdots + \xi_k$ for $k = 1, \ldots, n$ and $S_0 = 0$. Then let $M_n = \max_{0\leq k\leq n} |S_k|$ and $M_n' = \max_{0\leq k\leq n}(\min\{|S_k|, |S_n - S_k|\})$. It is easy to see that $M_n' \leq M_n$ and $M_n \leq M_n' + |S_n|$ a.s. Therefore, for every $\lambda > 0$, we have

$$P(\{M_n \geq \lambda\}) \leq P(\{M_n' \geq \lambda/2\}) + P(\{|S_n| \geq \lambda/2\}). \tag{3.1}$$

In what follows, we shall need the following auxiliary results.

Lemma 3.1. *Let $\gamma \geq 0$ and $\alpha > 1/2$ be given and suppose there are positive random variables $u_1, \ldots, u_n$ such that $E\left(\sum_{l=1}^n u_l\right)^{2\alpha} < \infty$ and*

$$P\left(\{|S_j - S_i| \geq \lambda, |S_k - S_j| \geq \lambda\}\right) \leq \frac{1}{\lambda^{2\gamma}} E\left(u_{i+1} + \cdots + u_k\right)^{2\alpha} \tag{3.2}$$

is satisfied for $0 \le i \le j \le k \le n$ *and every* $\lambda > 0$. *Then there exists a number* $K_{\gamma,\alpha}$ *such that for every positive* λ, *one has*

$$P\left(\left\{M_n' \ge \lambda\right\}\right) \le \frac{K_{\gamma,\alpha}}{\lambda^{2\gamma}} E\left(u_1 + \cdots + u_n\right)^{2\alpha}. \tag{3.3}$$

Proof. Let $\delta = 1/(2\gamma + 1)$. We have $2^{\delta}\left[1/2^{2\alpha\delta} + 1/K^{\delta}\right] \le 1$ for sufficiently large $K > 0$. We shall show that (3.3) is satisfied if K satisfies the above inequality and $K \ge 1$. It can be verified ([21], Theorem 2.12.1) that the minimal number K satisfying the above inequalities is given by

$$K_{\gamma,\alpha} = \left[\frac{1}{2^{1/(2\gamma+1)}} - \left(\frac{1}{2^{1/(2\gamma+1)}}\right)^{2\alpha}\right]^{-(2\gamma+1)}.$$

The proof of (3.3) we get by induction on n. For $n = 1$, the inequality (3.3) is trivial. Let $n = 2$. Immediately from (3.2) for $K \ge 1$, it follows that

$$\begin{aligned} P\left(\left\{M_2' \ge \lambda\right\}\right) &= P\left(\{\min\left[|S_1|, |S_2 - S_1|\right] \ge \lambda\}\right) \\ &\le \frac{1}{\lambda^{2\gamma}} E\left(u_1 + u_2\right)^{2\alpha} \le \frac{K}{2^{2\gamma}} E\left(u_1 + u_2\right)^{2\alpha} \end{aligned}$$

for $\lambda > 0$. Assume now that (3.3) is satisfied for every positive integer $k < n$. We shall show that it is also satisfied for $k = n$. Let $\upsilon = E\left(u_1 + \cdots + u_n\right)^{2\alpha}$, $\upsilon_0 = 0$, and $\upsilon_h = E\left(u_1 + \cdots + u_h\right)^{2\alpha}$, with $1 \le h \le n$. We can assume that $\upsilon > 0$. We have $\upsilon_{h-1} \le \upsilon_h$. Then $0 \le \upsilon_1/\upsilon \le \upsilon_2/\upsilon \le \cdots \le \upsilon_{n-1}/\upsilon \le 1$. Therefore, $[0, 1] = \bigcup_{h=1}^{n}\left[\upsilon_{h-1}/\upsilon, \upsilon_h/\upsilon\right]$. By virtue of the assumption $\alpha > 1/2$, we have $1/2^{2\alpha} \in [0, 1]$. Therefore, there is $1 \le h \le n$ such that $\upsilon_{h-1}/\upsilon \le 1/2^{2\alpha} \le \upsilon_h/\upsilon$. Define U_1, U_2, D_1, and D_2 by setting

$$U_1 = \max_{0 \le j \le h-1} \min\left\{|S_j|, |S_{h-1} - S_j|\right\}, \quad U_2 = \max_{h \le j \le n} \min\left\{|S_j - S_h|, |S_n - S_j|\right\},$$

$$D_1 = \min\left\{|S_{h-1}|, |S_n - S_{h-1}|\right\}, \quad \text{and} \quad D_2 = \min\left\{|S_h|, |S_n - S_h|\right\}.$$

Let us observe that for $1 \le h \le n$ and α taken as above, we have $\upsilon_{h-1} \le (2^{2\alpha} - 1)\upsilon/2^{2\alpha}$ and $z_{h+1} \le (2^{2\alpha} - 1)\upsilon/2^{2\alpha}$, where $z_{h+1} = E(u_{h+1} + \cdots + u_n)^{2\alpha}$. Indeed, we have $\upsilon_{h-1} \le \upsilon/2^{2\alpha} \le (2^{2\alpha} - 1)\upsilon/2^{2\alpha}$. Furthermore,

$$\begin{aligned} \frac{\upsilon_h}{\upsilon} + \frac{z_{h+1}}{\upsilon} &= \frac{E\left[\left(u_1 + \cdots + u_h\right)^{2\alpha} + \left(u_{h+1} + \cdots + u_n\right)^{2\alpha}\right]}{\upsilon} \\ &\le \frac{E\left[\left(u_1 + \cdots + u_h\right) + \left(u_{h+1} + \cdots + u_n\right)\right]^{2\alpha}}{\upsilon} = 1 \end{aligned}$$

and $1-\upsilon_h/\upsilon \leq 1-1/2^{2\alpha} = (2^{2\alpha}-1)/2^{2\alpha}$. Therefore, $z_{h+1} \leq (2^{2\alpha}-1)\upsilon/2^{2\alpha}$. Let us observe that (3.2) will be satisfied if we take $h-1$ instead of n. Since $h-1 < n$, we can assume that (3.3) is satisfied for random variables $\xi_1, \ldots, \xi_{h-1}$ and $u_1, \ldots, u_{h-1}$. Consequently, the above inequalities imply

$$P\left(\{U_1 \geq \lambda\}\right) \leq \frac{K}{\lambda^{2\gamma}} E\left(u_1 + \cdots + u_{h-1}\right)^{2\alpha} \leq \frac{K(2^{2\alpha}-1)}{2^{2\alpha}\lambda^{2\gamma}}\,\upsilon.$$

Similarly, taking indices $h \leq i \leq j \leq n$ in (3.2), we shall consider only random variables $\xi_{h+1}, \ldots, \xi_n$ and $u_{h+1}, \ldots, u_n$, and we can assume that (3.3) is satisfied for these random variables because $n-h < n$. With this and the above inequalities, we obtain

$$P\left(\{U_2 \geq \lambda\}\right) \leq \frac{K}{\lambda^{2\gamma}} E\left(u_{h+1} + \cdots + u_n\right)^{2\alpha} \leq \frac{K(2^{2\alpha}-1)}{2^{2\alpha}\lambda^{2\gamma}}\,\upsilon.$$

Next, by (3.2), we have

$$P\left(\{D_1 \geq \lambda\}\right) \leq \frac{1}{\lambda^{2\gamma}} E\left(u_1 + \cdots + u_n\right)^{2\alpha} = \frac{\upsilon}{\lambda^{2\gamma}} \quad \text{and} \quad P\left(\{D_2 \geq \lambda\}\right) \leq \frac{\upsilon}{\lambda^{2\gamma}}.$$

Let us observe that in the particular cases $h = 1$ and $h = n$, the above inequalities are trivial. Similarly as in ([21], Theorem 2.12.1), we can verify that $M_n^{'} \leq \max[U_1 + D_1, U_2 + D_2]$ and therefore,

$$P\left(\left\{M_n^{'} \geq \lambda\right\}\right) \leq P\left(\{U_1 + D_1 \geq \lambda\}\right) + P\left(\{U_2 + D_2 \geq \lambda\}\right). \tag{3.4}$$

On the other hand, we have

$$\begin{aligned} P\left(\{U_1 + D_1 \geq \lambda\}\right) &\leq P\left(\{U_1 \geq \lambda_0\}\right) + P\left(\{D_1 \geq \lambda_1\}\right) \\ &\leq \left[\frac{1}{\lambda_0^{2\alpha}} \frac{K(2^{2\alpha}-1)}{2^{2\alpha}} + \frac{1}{\lambda_1^{2\alpha}}\right]\upsilon \end{aligned} \tag{3.5}$$

for positive numbers λ_0 and λ_1 such that $\lambda = \lambda_0 + \lambda_1$. It can be verified ([21], Theorem 2.12.1) that for positive numbers $C_0, C_1, \lambda, \delta$, and γ such that $\delta = 1/(2\gamma + 1)$, we have

$$\min_{\lambda_0+\lambda_1=\lambda}\left[\frac{C_0}{\lambda_0^{2\gamma}} + \frac{C_1}{\lambda_1^{2\gamma}}\right] = \frac{1}{\lambda_1^{2\gamma}}\left[C_0^{\delta} + C_1^{\delta}\right]^{1/\delta},$$

where the minimum is taken over all positive numbers λ_0 and λ_1 such that $\lambda_0 + \lambda_1 = \lambda$. Therefore, (3.5) implies

$$P\left(\{U_1 + D_1 \geq \lambda\}\right) \leq \frac{\upsilon}{\lambda^{2\gamma}}\left[\left(\frac{K(2^{2\alpha}-1)}{2^{2\alpha}}\right)^{\delta} + 1\right]^{1/\delta}.$$

In a similar way, we obtain

$$P\left(\{U_2 + D_2 \geq \lambda\}\right) \leq \frac{\upsilon}{\lambda^{2\gamma}}\left[\left(\frac{K(2^{2\alpha}-1)}{2^{2\alpha}}\right)^{\delta} + 1\right]^{1/\delta}.$$

Therefore, (3.4) implies

$$P\left(\left\{M_n' \geq \lambda\right\}\right) \leq \frac{2\upsilon}{\lambda^{2\gamma}}\left[\left(\frac{K(2^{2\alpha}-1)}{2^{2\alpha}}\right)^{\delta} + 1\right]^{1/\delta}.$$

For $\alpha > 1/2$ and sufficiently large $K \geq 1$ satisfying $2^{\delta}[1/2^{2\alpha\delta} + 1/K^{\delta}] \leq 1$, we have

$$\left[\left(\frac{K(2^{2\alpha}-1)}{2^{2\alpha}}\right)^{\delta} + 1\right]^{1/\delta} \leq K.$$

Indeed, we have

$$\left[\left(\frac{2^{2\alpha}-1}{2^{2\alpha}}\right)^{\delta} + \frac{1}{K^{\delta}}\right] \to \left(\frac{2^{2\alpha}-1}{2^{2\alpha}}\right)^{\delta}$$

as $K \to \infty$. Therefore, for sufficiently large $K \geq 1$, we get

$$\left[\left(\frac{K(2^{2\alpha}-1)}{2^{2\alpha}}\right)^{\delta} + 1\right]^{1/\delta} = K\left[\left(\frac{2^{2\alpha}-1}{2^{2\alpha}}\right)^{\delta} + \frac{1}{K^{\delta}}\right]^{1/\delta} \leq K.$$

Then for sufficiently large $K \geq 1$, we get

$$P\left(\left\{M_n' \geq \lambda\right\}\right) \leq \frac{K_{\gamma,\alpha}}{\lambda^{2\gamma}} E\left(u_1 + \cdots + u_n\right)^{2\alpha}$$

with $K_{\gamma,\alpha} = 2K$. □

Lemma 3.2. *Let $\gamma \geq 1$ and an integer $\alpha > 1$ be given and suppose there are random variables $\xi_i : \Omega \to \mathbb{R}^m$ and $u_i : \Omega \to \mathbb{R}^+$ for $i = 1, \ldots, n$ such that $E(u_1 + \cdots + u_n)^{\alpha} < \infty$ and*

$$P\left(\{|S_j - S_i| \geq \lambda\}\right) \leq \frac{1}{\lambda^{\gamma}} E(u_{i+1} + \cdots + u_j)^{\alpha} \tag{3.6}$$

for every $\lambda > 0$ and $0 \leq i < j \leq n$. Then there is a positive number $K_{\gamma,\alpha}'$ such that

$$P\left(\{M_n \geq \lambda\}\right) \leq \frac{K_{\gamma,\alpha}'}{\lambda^{\gamma}} E\left(u_1 + \cdots + u_n\right)^{\alpha}. \tag{3.7}$$

Proof. Taking into account the inequalities $P(E_1 \cap E_2) \leq [P(E_1)]^{1/2}[P(E_2)]^{1/2}$ and $xy \leq (x+y)^2$ for $E_1, E_2 \in \mathcal{F}$ and $x, y \in \mathbb{R}$, we can easily see that (3.6) implies

$$
\begin{aligned}
&P\left(\left\{|S_j - S_i| \geq \lambda, |S_k - S_j| \geq \lambda\right\}\right) \\
&\leq \left[P\left(\left\{|S_j - S_i| \geq \lambda\right\}\right)\right]^{1/2} \left[P\left(\left\{|S_k - S_j| \geq \lambda\right\}\right)\right]^{1/2} \\
&\leq \frac{1}{\lambda^{\gamma/2}} \left[E\left(\sum_{i<l\leq j} u_l\right)^{\alpha}\right]^{1/2} \cdot \frac{1}{\lambda^{\gamma/2}} \left[E\left(\sum_{j<l\leq k} u_l\right)^{\alpha}\right]^{1/2} \\
&\leq \frac{1}{\lambda^{\gamma}} E\left[\left(\sum_{i<l\leq j} u_l\right)^{\alpha} + \left(\sum_{j<l\leq k} u_l\right)^{\alpha}\right] \\
&\leq \frac{2}{\lambda^{\gamma}} E\left[\sum_{i<l\leq j} u_l + \sum_{j<l\leq k} u_l\right]^{\alpha} = \frac{1}{\lambda^{\gamma}} \left(u_{i+1} + \cdots + u_k\right)^{\alpha}.
\end{aligned}
$$

Then the assumption (3.2) of Lemma 3.1 with $\gamma/2$ and $\alpha/2$ instead of γ and α, respectively, is satisfied. Therefore, by virtue of Lemma 3.1, we obtain

$$P\left(\left\{M_n' \geq \lambda\right\}\right) \leq \frac{\tilde{K}}{\lambda^{\gamma}} E\left(u_1 + \cdots + u_n\right)^{\alpha}$$

with $\tilde{K} = K_{\gamma/2,\alpha/2}$. On the other hand, (3.6) implies

$$P\left(\{|S_n| \geq \lambda\}\right) \leq \frac{1}{\lambda^{\gamma}} E\left(u_1 + \cdots + u_n\right)^{\alpha}.$$

With this and inequality (3.1), we obtain

$$P\left(\{M_n \geq \lambda\}\right) \leq \frac{K_{\gamma,\alpha}'}{\lambda^{\gamma}} E\left(u_1 + \cdots + n_n\right)^{\alpha}$$

with $K_{\gamma,\alpha}' = 2^{\gamma}\left(\hat{K} + 1\right)$. Then (3.7) is satisfied. □

We can prove now the following result.

Theorem 3.6. *A sequence $(X^n)_{n=1}^{\infty}$ of continuous m-dimensional stochastic processes $X^n = (X^n(t))_{0\leq t\leq T}$ on a probability space $\mathcal{P} = (\Omega, \mathcal{F}, P)$ is tight if for every $\epsilon > 0$ there is a number $a > 0$ such that $P(|X^n(0)| > a) \leq \epsilon$ for $n \geq 1$ and there are $\gamma \geq 0$, an integer $\alpha > 1$, and a continuous nondecreasing stochastic process $\Gamma = (\Gamma(t))_{0\leq t\leq T}$ on $\mathcal{P}$ such that $E[\Gamma(T) - \Gamma(0)]^{\alpha} < \infty$ and*

$$P\left(\{|X^n(t) - X^n(s)| \geq \lambda\}\right) \leq \frac{1}{\lambda^{\gamma}} E\left|\Gamma(t) - \Gamma(s)\right|^{\alpha} \tag{3.8}$$

for every $n \geq 1$, $\lambda > 0$, and $s, t \in [0, T]$.

Proof. For simplicity, assume that $T = 1$ and $\Gamma(0) = 0$ a.s. It is clear that $(X^n)_{n=1}^{\infty}$ satisfies condition (i) of Theorem 2.4. Therefore, by virtue of Remark 2.1, it is enough only to verify that for every $\varepsilon > 0$, one has

$$\lim_{\delta\to 0}\sup_{n\geq 1}\sum_{j<\delta^{-1}} P\left(\left\{\sup_{j\delta\leq s\leq (j+1)\delta}|X^n(s)-X^n(j\delta)|\geq\varepsilon\right\}\right)=0. \tag{3.9}$$

Fix $n \geq 1$ and $j \geq 1$. For a positive integer k, consider m-dimensional random variables $\xi_1^j,\ldots,\xi_k^j$ defined by

$$\xi_i^j = X^n\left(j\delta+\frac{i}{k}\delta\right)-X^n\left(j\delta+\frac{i-1}{k}\delta\right)$$

for $i = 1,\ldots,k.$ Immediately from (3.8), it follows that (3.6) is satisfied with

$$u_l^j = \Gamma\left(j\delta+\frac{l}{k}\delta\right)-\Gamma\left(j\delta+\frac{l-1}{k}\delta\right)$$

for $l = 1,2,\ldots,k,$ because

$$\begin{aligned}P\left(\left\{\left|S_j-S_i\right|\geq\lambda\right\}\right) &= P\left(\left\{\left|X^n\left(j\delta+\frac{i}{k}\delta\right)-X^n\left(j\delta+\frac{j}{k}\delta\right)\right|\geq\lambda\right\}\right)\\ &\leq \frac{1}{\lambda^\gamma}E\left|\Gamma\left(j\delta+\frac{i}{k}\delta\right)-\Gamma\left(j\delta+\frac{j}{k}\delta\right)\right|^\alpha\end{aligned}$$

$$\begin{aligned}&= \frac{1}{\lambda^\gamma}E\left(\sum_{i<l\leq j}\left[\Gamma\left(j\delta+\frac{l}{k}\delta\right)-\Gamma\left(j\delta+\frac{l-1}{k}\delta\right)\right]\right)^\alpha\\ &= \frac{1}{\lambda^\gamma}E\left(u_{i+1}^j+\cdots+u_j^j\right)^\alpha.\end{aligned}$$

Therefore, by virtue of Lemma 3.2, there is $K'_{\gamma,\alpha} > 0$ such that

$$\begin{aligned}&P\left(\left\{\max_{0\leq i\leq k}\left|X^n\left(j\delta+\frac{i}{k}\delta\right)-X^n\left(j\delta\right)\right|\geq\lambda\right\}\right)\leq\frac{K'_{\gamma,\alpha}}{\lambda^\gamma}E\left(u_1+\cdots+u_k\right)^\alpha\\ &=\frac{K'_{\gamma,\alpha}}{\lambda^\gamma}E\left[\Gamma\left((j+1)\delta\right)-\Gamma\left(j\delta\right)\right]^\alpha.\end{aligned}$$

Similarly as in [21], Theorem 2.12.3, by continuity of X^n, it follows that

$$P\left(\left\{\sup_{j\delta\leq s\leq (j+1)\delta}|X^n(s)-X^n(j\delta)|\geq\lambda\right\}\right)\leq\frac{K'_{\gamma,\alpha}}{\lambda^\gamma}E\left[\Gamma((j+1)\delta)-\Gamma(j\delta)\right]^\alpha.$$

Therefore, for every $n\geq 1$ one has

$$\sum_{j<\delta^{-1}}P\left(\left\{\sup_{j\delta\leq s\leq (j+1)\delta}|X^n(s)-X^n(j\delta)|\geq\lambda\right\}\right)$$
$$\leq\frac{K'_{\gamma,\alpha}}{\lambda^\gamma}E\left\{\Lambda_\delta\sum_{j<\delta^{-1}}\left[\Gamma((j+1)\delta)-\Gamma(j\delta)\right]\right\},$$

where

$$\Lambda_\delta=\left[\max_{j<\delta^{-1}}|\Gamma((j+1)\delta)-\Gamma(j\delta)|\right]^{\alpha-1}.$$

Hence it follows that

$$\sum_{j<\delta^{-1}}P\left(\left\{\sup_{j\delta\leq s\leq (j+1)\delta}|X^n(s)-X^n(j\delta)|\geq\lambda\right\}\right)\leq\frac{K'_{\gamma,\alpha}}{\lambda^\gamma}E\left[\Lambda_\delta\Gamma(1)\right],$$

because $\sum_{j<\delta^{-1}}\left[\Gamma((j+1)\delta)-\Gamma(j\delta)\right]\leq\Gamma(1)$ a.s. By continuity of the stochastic process $\Gamma=(\Gamma(t))_{0\leq t\leq 1}$ and the assumption $\alpha>1$, we get $\lim_{\delta\to 0}H_\delta(\omega)=0$ for a.e. $\omega\in\Omega$, where $H_\delta(\omega)=\sup_{0\leq t\leq 1}\left[\Gamma(t+\delta)(\omega)-\Gamma(t)(\omega)\right]^{\alpha-1}$ for $\omega\in\Omega$. Hence, by the properties of Γ, it follows that $\lim_{\delta\to 0}E\left[H_\delta\Gamma(1)\right]=0$ for every $n\geq 1$. But $\Lambda_\delta\leq H_\delta$ a.s. Then

$$\sum_{j<\delta^{-1}}P\left(\left\{\sup_{j\delta\leq s\leq (j+1)\delta}|X^n(s)-X^n(j\delta)|\geq\lambda\right\}\right)\leq\frac{K'_{\gamma,\alpha}}{\lambda^\gamma}E\left[H_\delta\Gamma(1)\right].$$

Therefore, (3.9) is satisfied, which together with the property of the sequence $(X^n(0))_{n=1}^\infty$ implies that $(X^n)_{n=1}^\infty$ is tight. □

4 Special Classes of Stochastic Processes

There are two important classes of stochastic processes: martingales and Markov processes. We characterize them by giving their most important properties. Similarly as above, we shall denote by $\mathcal{P}_\mathbb{F}$ a complete filtered probability space $(\Omega,\mathcal{F},\mathbb{F},P)$ with a filtration $\mathbb{F}=(\mathcal{F}_t)_{t\geq 0}$ satisfying the usual conditions.

A real-valued $\mathbb{F}$-adapted stochastic process $X = (X_t)_{t\geq 0}$ is said to be an $\mathbb{F}$-martingale or simply martingale (supermartingale, submartingale) on $\mathcal{P}_{\mathbb{F}}$ if (1) $E|X_t| < \infty$ for $t \geq 0$ and (2) $E[X_t|\mathcal{F}_s] = X_s$ ($E[X_t|\mathcal{F}_s] \leq X_s$, $E[X_t|\mathcal{F}_s] \geq X_s$) a.s. for $0 \leq s \leq t < \infty$. A martingale $X = (X_t)_{t\geq 0}$ on $\mathcal{P}_{\mathbb{F}}$ is said to be closed by a random variable Y on $\mathcal{P}_{\mathbb{F}}$ if $E|Y| < \infty$ and $X_t = E[Y|\mathcal{F}_t]$ a.s. for $t \geq 0$.

We shall present some properties of martingales.

Theorem 4.1. *Let $X = (X_t)_{t\geq 0}$ be a supermartingale on $\mathcal{P}_{\mathbb{F}}$. The function $\mathbb{R}^+ \ni t \to E[X_t] \in \mathbb{R}$ is right continuous if and only if there exists a unique modification Y of X that is càdlàg.* □

Corollary 4.1. *If $X = (X_t)_{t\geq 0}$ is a martingale on $\mathcal{P}_{\mathbb{F}}$, then there exists a modification Y of X that is càdlàg.*

Proof. If X is a martingale, then the function $\mathbb{R}^+ \ni t \to E[X_t] \in \mathbb{R}$ is constant, and hence it is continuous. By Theorem 4.1, there exists a unique modification Y of X that is càdlàg. □

Theorem 4.2. *Let $X = (X_t)_{t\geq 0}$ be a right continuous supermartingale (martingale) on $\mathcal{P}_{\mathbb{F}}$, and S and T bounded $\mathbb{F}$-stopping times such that $S \leq T$ a.s. Then X_S and X_T are integrable and $X_S \geq E[X_T|\mathcal{F}_S]$ ($X_S = E[X_T|\mathcal{F}_S]$) a.s.* □

If T is an $\mathbb{F}$-stopping, time then so is $t \wedge T$ for each $t \geq 0$. Given a stochastic process $X = (X_t)_{t\geq 0}$ then the process $(X_{t\wedge T})_{t\geq 0}$ is denoted by X^T and said to be the process stopped at T.

Corollary 4.2. *If $X = (X_t)_{t\geq 0}$ is an $\mathbb{F}$-adapted and càdlàg process and T is an $\mathbb{F}$-stopping time, then $X_t^T = X_t \mathbb{1}_{\{t<T\}} + X_T \mathbb{1}_{\{t\geq T\}}$ and X^T is also $\mathbb{F}$-adapted.* □

We shall show now that if X is a right continuous and uniformly integrable martingale, then the stopped process X^T is also a martingale. Recall that a family $(X_\alpha)_{\alpha\in\Lambda}$ of random variables on $\mathcal{P}_{\mathbb{F}}$ is said to be uniformly integrable if $\lim_{n\to\infty} \sup_{\alpha\in\Lambda} \int_{\{|X_\alpha|\geq n\}} |X_\alpha| \mathrm{d}P = 0$.

Theorem 4.3. *Let $X = (X_t)_{t\geq 0}$ be a uniformly integrable right continuous martingale on $\mathcal{P}_{\mathbb{F}}$ and let T be an $\mathbb{F}$-stopping time. Then X^T is also a uniformly integrable right continuous martingale.*

Proof. It is clear that X^T is right continuous. By Corollary 4.2, it is also $\mathbb{F}$-adapted. Hence, by the equality $X_T \mathbb{1}_{\{t\geq T\}} = X_t^T - X_t \mathbb{1}_{\{t<T\}}$ and properties of stopping times, it follows that $X_T \mathbb{1}_{\{t\geq T\}}$ is $\mathcal{F}_t$-measurable for every $t \geq 0$. Let $0 \leq s < t \leq \infty$ be fixed. We have $\mathbb{1}_{\{t<T\}} = \mathbb{1}_{\{s<T\}} - \mathbb{1}_{\{s<T\}} \cdot \mathbb{1}_{\{t\geq T\}}$ and $\mathbb{1}_{\{t\geq T\}} = \mathbb{1}_{\{s\geq T\}} + \mathbb{1}_{\{s<T\}} \cdot \mathbb{1}_{\{t\geq T\}}$. Therefore,

$$
\begin{aligned}
E[X_t^T|\mathcal{F}_s] &= E[\mathbb{1}_{\{t<T\}}X_t + \mathbb{1}_{\{t\geq T\}}X_T|\mathcal{F}_s] \\
&= E[\mathbb{1}_{\{s<T\}}X_t|\mathcal{F}_s] - E[\mathbb{1}_{\{s<T\}}\cdot\mathbb{1}_{\{t\geq T\}}X_t|\mathcal{F}_s] \\
&\quad + E[\mathbb{1}_{\{s\geq T\}}X_T|\mathcal{F}_s] + E[\mathbb{1}_{\{s<T\}}\cdot\mathbb{1}_{\{t\geq T\}}X_T|\mathcal{F}_s] \\
&= \mathbb{1}_{\{s<T\}}E[X_t|\mathcal{F}_s] + E[\mathbb{1}_{\{s\geq T\}}X_T|\mathcal{F}_s].
\end{aligned}
$$

But $X_T\mathbb{1}_{\{s\geq T\}}$ is $\mathcal{F}_s$-measurable and $E[X_t|\mathcal{F}_s] = X_s$ for every $t \geq 0$. Therefore, $E[X_t^T|\mathcal{F}_s] = \mathbb{1}_{\{s<T\}}X_s + \mathbb{1}_{\{s\geq T\}}X_T = X_s^T$. □

In what follows, we shall need the following results.

Lemma 4.1 (Jensen's inequality). *Assume that $\varphi : \mathbb{R} \to \mathbb{R}$ is convex, and let X and $\varphi(X)$ be integrable random variables on $\mathcal{P}_{\mathbb{F}}$. For every σ-algebra $\mathcal{G} \subset \mathcal{F}$, one has $\varphi\left(E[X|\mathcal{G}]\right) \leq E[\varphi(X)|\mathcal{G}]$.*

Corollary 4.3. *Let $X = (X_t)_{t\geq 0}$ be a martingale and let $\varphi : \mathbb{R} \to \mathbb{R}$ be convex such that $\varphi(X_t)$ is integrable for $0 \leq t < \infty$. Then $\varphi(X)$ is a submartingale. In particular, $|X|$ and X^2 are submartingales.* □

Theorem 4.4 (Doob's martingale inequality). *If $X = (X_t)_{t\geq 0}$ is a continuous martingale on $\mathcal{P}_{\mathbb{F}}$, then for all $p \geq 1$, $T \geq 0$, and $\lambda > 0$, one has $P\left(\{\sup_{0\leq t\leq T}|X_t| \geq \lambda\}\right) \leq \frac{1}{\lambda^p}E[|X_T|^p]$.* □

Theorem 4.5. *Let $X = (X_t)_{t\geq 0}$ be a positive submartingale on $\mathcal{P}_{\mathbb{F}}$. For all $p > 1$, one has $\|\sup_{t\geq 0}|X_t|\| \leq q\sup_{t\geq 0}\|X_t\|$, where q is such that $1/p + 1/q = 1$, and $\|\cdot\|$ denotes the norm on the space $\mathbb{L}^p(\Omega, \mathcal{F}, P, \mathbb{R})$.* □

Corollary 4.4. *If X is as in Theorem 4.5 with $p = 2$, then $E\left(\sup_{t\geq 0}|X_t|\right)^2 \leq 4\sup_{t\geq 0}E|X_t|^2$.* □

Theorem 4.6. *Let $X = (X_t)_{t\geq 0}$ be $\mathbb{F}$-adapted càdlàg process on $\mathcal{P}_{\mathbb{F}}$ such that $E|X_T| < \infty$ and $EX_T = 0$ for any $\mathbb{F}$-stopping time T. Then X is a uniformly integrable martingale on $\mathcal{P}_{\mathbb{F}}$.*

Proof. Let $0 \leq s \leq t < \infty$ and $\Lambda \in \mathcal{F}_s$. For fixed $u \geq 0$, let $u_\Lambda = u$ if $\omega \in \Lambda$ and $u_\Lambda = \infty$ if $\omega \notin \Lambda$. It can be verified that for $u \geq s$, the random variable $u_\Lambda : \Omega \to \mathbb{R} \cup \{\infty\}$ is an $\mathbb{F}$-stopping time. Moreover,

$$
\int_\Lambda X_{u_\Lambda}\mathrm{d}P = \int_\Omega X_{u_\Lambda}\mathrm{d}P - \int_{\Omega\setminus\Lambda} X_\infty \mathrm{d}P = -\int_{\Omega\setminus\Lambda} X_\infty\mathrm{d}P,
$$

because $E[X_{u_\Lambda}] = 0$ for $u \geq s$. Thus for $\Lambda \in \mathcal{F}_s$ and $s < t$, one has $E[X_t\mathbb{1}_\Lambda] = E[X_s\mathbb{1}_\Lambda] = -E[X_\infty\mathbb{1}_{\Omega\setminus\Lambda}]$. Then $E[X_t|\mathcal{F}_s] = X_s$ for $0 \leq s \leq t < \infty$. □

We can also consider discrete-time martingales. Given a probability space $(\Omega, \mathcal{F}, P)$ and an increasing sequence $(\mathcal{F}_n)_{n=1}^\infty$ of sub-σ-algebras $\mathcal{F}_n$ of $\mathcal{F}$, we define a discrete-time martingale as a sequence $(X_n)_{n=0}^\infty$ of random variables on $(\Omega, \mathcal{F}, P)$ adapted to $(\mathcal{F}_n)_{n=0}^\infty$ such that $E|X_n| < \infty$ and $E[X_{n+1}|\mathcal{F}_n] = X_n$

for $n \geq 0$. We can also consider the discrete martingales $(X_n)_{n=-\infty}^{n=0}$ with respect to a discrete filtration $(\mathcal{F}_n)_{n=-\infty}^{n=0}$. For such martingales, we have the following backward convergence theorem.

Theorem 4.7. *Let $(X_n)_{n\leq 0}$ be a uniformly integrable discrete-time martingale on a probability space $(\Omega, \mathcal{F}, P)$ with respect to a discrete filtration $(\mathcal{F}_n)_{n\leq 0}$. Then $X_n \to E[X_0|\mathcal{F}_{-\infty}]$ a.s. and $X_n \to E[X_0|\mathcal{F}_{-\infty}]$ in the $\mathbb{L}$-norm topology, where $\mathcal{F}_{-\infty} = \bigcap_{n\leq 0} \mathcal{F}_n$.* □

As a consequence of this theorem, we obtain the following result.

Lemma 4.2. *Let $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ with $\mathbb{F} = (\mathcal{F}_t)_{0\leq t\leq T}$ be such that $\mathcal{F}_T = \mathcal{F}$ and let $(t_k)_{k=1}^{\infty}$ be a decreasing sequence of $[0, T]$ converging to t^* as $k \to \infty$. Then for every $X \in \mathbb{L}(\Omega, \mathbb{R})$, one has $E[X|\mathcal{F}_{t_k}] \to E[X|\mathcal{F}_{t^*}]$ a.s. and $E[X|\mathcal{F}_{t_k}] \to E[X|\mathcal{F}_{t^*}]$ in the $\mathbb{L}$-norm topology as $k \to \infty$.*

Proof. Let $\mathcal{F}_n = \mathcal{F}_{t_{(-n)}}$ and $X_n = E[X|\mathcal{F}_n]$ for $n = -k$ with $k = 1, 2, \ldots$. Put $t_0 = T$ and $X_0 = E[X|\mathcal{F}_{t_0}]$. We have $\sup_{n\leq 0} E|X_n|^2 \leq E|X|^2 < \infty$, and $(X_n)_{n\leq 0}$ is a uniformly integrable discrete martingale with respect to the discrete filtration $(\mathcal{F}_n)_{n\leq 0}$. Then by virtue of Theorem 4.7, we have $X_n \to E[X_0|\mathcal{F}_{-\infty}]$ a.s. and $X_n \to E[X_0|\mathcal{F}_{-\infty}]$ in the $\mathbb{L}$-norm topology as $n \to -\infty$. But

$$X_n = E[X|\mathcal{F}_{t_{(-n)}}] = E[X|\mathcal{F}_{t_k}], \quad X_0 = E[X|\mathcal{F}_T] = E[X|\mathcal{F}] = X$$

and $\mathcal{F}_{-\infty} = \bigcap_{n\leq 0} \mathcal{F}_{t_{(-n)}} = \bigcap_{k\geq 0} \mathcal{F}_{t_k} = \mathcal{F}_{t^*}$. Therefore, $E[X|\mathcal{F}_{t_k}] \to E[X|\mathcal{F}_{t^*}]$ a.s. and $E[X|\mathcal{F}_{t_k}] \to E[X|\mathcal{F}_{t^*}]$ in the $\mathbb{L}$-norm topology as $k \to \infty$. □

Remark 4.1. It can be verified that if $Y \in \mathbb{L}(\Omega, \mathbb{R}^d)$ and $(\mathcal{F}_k)_{k\geq 1}$ is a filtration of $\mathcal{F}$ and $\mathcal{F}_\infty$ is a σ-algebra generated by $\{\mathcal{F}_1, \mathcal{F}_2, \ldots\}$, then $E[Y|\mathcal{F}_k] \to E[Y|\mathcal{F}_\infty]$ a.s. and $E[Y|\mathcal{F}_k] \to E[Y|\mathcal{F}_\infty]$ in the $\mathbb{L}$-norm topology as $k \to \infty$.

Proof. Let $M_k =: E[Y|\mathcal{F}_k]$ for every $k \geq 1$. It is clear that a discrete martingale $(M_k)_{k\geq 1}$ is uniformly integrable. Then there exists $M \in \mathbb{L}(\Omega, \mathbb{R}^d)$ such that $M_k \to M$ a.s. in the $\mathbb{L}$-norm topology as $k \to \infty$. It remains to prove that $M = E[Y|\mathcal{F}_\infty]$. To see this, let us observe that

$$\|M_k - E[M|\mathcal{F}_k]\| = \|E[M_k|\mathcal{F}_k] - E[M|\mathcal{F}_k]\| \leq \|M_k - M\|$$

for every $k \geq 1$, which implies that $\|M_k - E[M|\mathcal{F}_k]\| \to 0$ as $k \to \infty$. Hence it follows that for every $\bar{k} \geq 1$ and every $A \in \mathcal{F}_{\bar{k}}$, we have $\int_A (Y - M) dP = 0$, because

$$\int_A (Y - M) dP = \int_A E[(Y - M)|\mathcal{F}_k] dP = \int_A (M_k - E[M|\mathcal{F}_k]) dP$$

for $k \geq \bar{k}$ and $\int_A (M_k - E[M|\mathcal{F}_k]) dP \to 0$ as $k \to \infty$. This implies that $\int_A (Y - M) dP = 0$ for every $A \in \bigcup_{k=1}^{\infty} \mathcal{F}_k$. Therefore, $E[Y|\mathcal{F}_\infty] = E[M|\mathcal{F}_\infty] = M$. □

An $\mathbb{F}$-adapted càdlàg process $X = (X_t)_{t\geq 0}$ on $\mathcal{P}_{\mathbb{F}}$ is said to be a local $\mathbb{F}$-martingale if there exists an increasing sequence $(T_n)_{n=1}^{\infty}$ of $\mathbb{F}$-stopping times T_n with $\lim_{n\to\infty} T_n = \infty$ a.s. such that the process $(X_{t\wedge T_n}\mathbb{1}_{\{T_n>0\}})_{t\geq 0}$ is a uniformly integrable martingale for each $n \geq 1$. Such a sequence $(T_n)_{n=1}^{\infty}$ of $\mathbb{F}$-stopping times is called a fundamental sequence of a local martingale X. It can be verified that if X and Y are continuous real-valued local martingales, then there exists a unique (up to indistinguishability) $\mathbb{F}$-adapted continuous process of bounded variation $\langle X, Y\rangle$ with $\langle X, Y\rangle_0 = 0$ a.s. such that $XY - \langle X, Y\rangle$ is a continuous local martingale. The process $\langle X, Y\rangle$ is called the cross-variation of X and Y. If $X = Y$, we write $\langle X\rangle = \langle X, X\rangle$ and call this process the quadratic variation of X. It is nondecreasing and $\mathbb{F}$-adapted. For a continuous process $X = (X_t)_{t\in[0,T]}$ on a probability space $(\Omega, \mathcal{F}, P)$, its quadratic variation can be also defined in the following way. Given a partition $\Delta = \{0 = t_0 < t_1 < \cdots < t_r = T\}$, we can define the process $\langle X\rangle_t^{\Delta}$ by setting $\langle X\rangle_t^{\Delta} = \sum_{k=0}^{r-1}(X_{t\wedge t_{k+1}} - X_{t\wedge t_k})^2$. If a sequence $(\Delta_n)_{n=1}^{\infty}$ of partitions $0 = t^n < t_1^n < \cdots < t_{r^n}^n$ of $[0, T]$ is such that the sequence $(|\Delta_n|)_{n=1}^{\infty}$ defined by $|\Delta_n| = \max_{0\leq k\leq r^n-1}|t_{k+1}^n - t_k^n|$ converging to zero as $n \to \infty$, then we can consider for every $t \in [0, T]$, the limit in probability of the sequence $(\langle X\rangle_t^{\Delta_n})_{n=1}^{\infty}$. If such a limit exists and is independent of the choice of sequence $(\Delta_n)_{n=1}^{\infty}$, then it is equal to the quadratic variation of X. It can be proved that if $X = (X_t)_{t\in[0,T]}$ is a continuous bounded $\mathbb{F}$-martingale, then for every sequence of partitions $(\Delta_n)_{n=1}^{\infty}$ such that $|\Delta_n| \to 0$ as $n \to \infty$, the sequence $(\langle X\rangle_t^{\Delta_n})_{n=1}^{\infty}$ converges uniformly in the $\mathbb{L}^2$-norm topology to $(\langle X\rangle_t)_{0\leq t\leq T}$.

From many viewpoints, very interesting applications have stochastic processes $X = (X_t)_{t\geq 0}$ that are representable (not necessarily in a unique way) as sums $X = X_0 + A + M$, where $A = (A_t)_{t\geq 0}$ is a càdlàg, $\mathbb{F}$-adapted process with paths of finite variation on compacts and $M = (M_t)_{t\geq 0}$ is a local $\mathbb{F}$-martingale on a given filtered probability space $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ satisfying the usual conditions. Such processes are said to be semimartingales on $\mathcal{P}_{\mathbb{F}}$. Similarly as above, we can define semimartingales that are measurable, $\mathbb{F}$-adapted, continuous, and right and left continuous. The class of semimartingales is stable with respect to many transformations, such as absolute changes of measure, time changes, localization, and changes of filtration.

By the definition of martingales and the interpretation of conditional expectations of random variables, it follows that the martingale property means that for a given present time s, the process has no tendency in future times $t \geq s$, that is, the average over all future possible states of X_t gives just the present state X_s. In contrast, the Markov property, which will follow in the next definition, means that the present has no memory, that is, that the average of X_t knowing the past is the same as the average of X_t knowing the present. Let $X = (X_t)_{t\geq 0}$ be an n-dimensional $\mathbb{F}$-adapted process on $\mathcal{P}_{\mathbb{F}}$. It is called an $\mathbb{F}$-Markov process if for every $0 \leq s \leq t < \infty$ and every bounded Borel measurable function $f : \mathbb{R}^d \to \mathbb{R}$, one has $E[f(X_t)|\mathcal{F}_s] = E[f(X_t)|\sigma(X_s)]$ a.s.

Remark 4.2. Equivalently, the above process $X = (X_t)_{t\geq 0}$ is an $\mathbb{F}$-Markov process if for each $t \geq 0$, the σ-algebras $\mathcal{F}_t$ and $\sigma(\{X_u : 0 \leq u \leq t\})$ are conditionally independent given X_t. □

Remark 4.3. In particular, for $f(x) = x$, the Markov property defined above takes the form $E[X_t|\mathcal{F}_s] = E[X_t|\sigma(X_s)]$ a.s. for every $0 \leq s \leq t < \infty$. □

Using the Markov property, one can define a transition function for a Markov process on $\mathcal{P}_{\mathbb{F}}$ in the following way: for every $0 \leq s \leq t < \infty$ and bounded and Borel measurable function $f : \mathbb{R}^d \to \mathbb{R}$, we take $P_{s,t}(X_s, f) = E[f(X_t)|\mathcal{F}_s]$. In particular, if $f = \mathbb{1}_A$, the indicator function of a measurable set $A \subset \mathbb{R}^d$, then the preceding equality reduces to $P(X_t \in A|\mathcal{F}_s) = P_{s,t}(X_t, \mathbb{1}_A)$. Usually, we write $P_{s,t}(X_t, A)$ instead of $P_{s,t}(X_t, \mathbb{1}_A)$. A Markov process X on $\mathcal{P}_{\mathbb{F}}$ is said to be time-homogeneous if its transition function $P_{s,t}$ satisfies $P_{s,t} = P_{t-s}$ for every $0 \leq s \leq t < \infty$.

Remark 4.4. If a Markov process X is time-homogeneous, then the family $(P_t)_{t\geq 0}$ of its transition functions $P_t = P_{t-0}$ is a semigroup of operators known as the transition semigroup $(P_t)_{t\geq 0}$. □

Corollary 4.5. *If X is a time-homogeneous Markov process on $\mathcal{P}_{\mathbb{F}}$, then for every $A \subset \mathbb{R}^d$ and $0 \leq s \leq t < \infty$, one has $E[\mathbb{1}_{\{X_{t+s}\in A\}}|\mathcal{F}_t] = P_s(X_t, A)$.* □

In contrast to the Markov property, we can define the strong Markov property if we require that the Markov property hold for every $\mathbb{F}$-stopping time. More precisely, a time-homogeneous $\mathbb{F}$-Markov process X is said to be a strong $\mathbb{F}$-Markov process if for every $\mathbb{F}$-stopping time T with $P(T < \infty) = 1$, every measurable set $A \subset \mathbb{R}^d$, and $s \geq 0$, one has $E[\mathbb{1}_{\{X_{T+s}\in A\}}|\mathcal{F}_T] = P_s(X_T, A)$.

Remark 4.5. The strong Markov property for a process X on $\mathcal{P}_{\mathbb{F}}$ can be equivalently written as follows: $E[f(X_{T+s})|\mathcal{F}_T] = P_s(X_T, f)$ for every $\mathbb{F}$-stopping time T and every bounded Borel measurable function $f : \mathbb{R}^d \to \mathbb{R}$. □

In the next section, we define two special processes known as the Poisson process and the Brownian motion. They are important examples of strong Markov processes with respect to their natural filtrations. The Brownian motion belongs to both classes of processes presented above.

5 Poisson Processes and Brownian Motion

Poisson processes and Brownian motion are the two most important examples in the theory of stochastic processes. Assume that we are given a filtered probability space $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ satisfying the usual conditions. Let $(T_n)_{n=0}^{\infty}$ be a strictly increasing sequence of positive random variables such that $T_0 = 0$ a.s. The process $N = (N_t)_{t\geq 0}$ defined by $N_t = \sum_{n\geq 1} \mathbb{1}_{\{t\geq T_n\}}$ with values in $\mathbb{N} \cup \{\infty\}$, where $\mathbb{N} = \{0, 1, 2, \ldots\}$, is called the counting process associated to the sequence $(T_n)_{n=0}^{\infty}$.

Remark 5.1. If $T = \sup_{n \geq 1} T_n$, then we have $[T_n, \infty) = \{N \geq n\} = \{(t, \omega) : N_t(\omega) \geq n\}$, $[T_n, T_{n+1}) = \{N = n\}$ and $[T, \infty) = \{N = \infty\}$. □

If $\sup_{n \geq 1} T_n < \infty$ a.s., then the random variable $T = \sup_{n \geq 1} T_n$ is called the explosion time of N. If $T = \infty$ a.s., then N is said to be a counting process without explosions.

Corollary 5.1. *For every* $0 \leq s \leq t < \infty$, *one has* $N_t - N_s = \sum_{n \geq 1} \mathbb{1}_{\{s < T_n \leq t\}}$. □

The increments $N_t - N_s$ count the number of random times T_n that occur between the fixed times s and t. Immediately from the definition of a counting process, it follows that it is not necessarily adapted to the filtration $\mathbb{F}$. In particular, if T_n are $\mathbb{F}$-stopping times for $n = 1, 2, \ldots$, then the counting process is $\mathbb{F}$-adapted. This follows from the following theorem.

Theorem 5.1. *A counting process* N *is* $\mathbb{F}$*-adapted if and only if the associated random variables* T_n *are* $\mathbb{F}$*-stopping times for* $n = 1, 2, \ldots$.

Proof. If T_n are $\mathbb{F}$-stopping times for $n = 1, 2, \ldots$, with $T_0 = 0$ a.s., then $\{N_t = n\} = \{T_n \leq t\} \cap \{T_{n+1} > t\} \in \mathcal{F}_t$ for $n = 1, 2, \ldots$. Thus N_t is $\mathcal{F}_t$-measurable. If N is $\mathbb{F}$-adapted, then $\{T_n \leq t\} = \{N_t \geq n\} \in \mathcal{F}_t$ for every t, and therefore, T_n is an $\mathbb{F}$-stopping time. □

Remark 5.2. A counting process without explosions has right continuous paths with left limits. □

An $\mathbb{F}$-adapted counting process $N = (N_t)_{t \geq 0}$ without explosion is said to be a Poisson process if the following conditions are satisfied:

(i) For every $0 \leq s \leq t < \infty$, the random variable $N_t - N_s$ is independent of $\mathcal{F}_s$.
(ii) For every $0 \leq s \leq t < \infty$ and $0 \leq u \leq v < \infty$ such that $t - s = v - u$, the random variables $N_t - N_s$ and $N_v - N_u$ have the same distributions.

Remark 5.3. Properties (i) and (ii) are known as increments independent of the past and stationary increments property, respectively. □

Theorem 5.2. *Let* N *be a Poisson process on* $\mathcal{P}_{\mathbb{F}}$. *Then* $P(\{N_t = n\}) = \mathrm{e}^{-\lambda t} \cdot (\lambda t)^n / n!$ *for* $n = 0, 1, 2, \ldots$ *for some* $\lambda \geq 0$.

Proof (Sketch of proof). The proof runs into the following four steps.

Step 1. By the properties of the Poisson process from $\{N_t = 0\} = \{N_s = 0\} \cap \{N_t - N_s = 0\}$, it follows that $P(\{N_t = 0\}) = P(\{N_s = 0\}) \cdot P(\{N_t - N_s = 0\}) = P(\{N_s = 0\}) \cdot P(\{N_{t-s} = 0\}$ for $0 \leq s \leq t < \infty$. Taking $\alpha(t) = P(\{N_t = 0\})$, we get $\alpha(t) = \alpha(s)\alpha(t-s)$ for all $0 \leq s \leq t < \infty$. By the right continuity of α, we deduce that either $\alpha(t) = 0$ for $t \geq 0$ or $\alpha(t) = \mathrm{e}^{-\lambda t}$ for some $\lambda \geq 0$. If $\alpha(t) = 0$, it would follow that $N_t(\omega) = 0$ a.s. for all $t \geq 0$, which would contradict that N is a counting process.

Step 2. It is verified that $\lim_{t\to 0}(1/t)P(\{N_t \geq 2\}) = 0$.

Step 3. By Step 2, we have $P(\{N_t \geq 2\}) = o(t)$, which together with the equality $P(\{N_t = 1\}) = 1 - P(\{N_t = 0\} - P(\{N_t \geq 2\})$ gives

$$\lim_{t\to 0} \frac{1}{t} P(\{N_t = 1\}) = \lim_{t\to 0} \frac{1 - \mathrm{e}^{-\lambda t} + o(t)}{t} = \lambda .$$

Step 4. Let $\varphi(t) = E[\alpha^{N_t}]$ for $0 \leq \alpha \leq 1$. By the properties of the Poisson process, for every $0 \leq s < t \leq \infty$, we get $\varphi(t+s) = \varphi(t) \cdot \varphi(s)$, which in turn implies that $\varphi(t) = \mathrm{e}^{t\psi(\alpha)}$, where $\psi(t) = \lim_{t\to 0}[(\varphi(t) - 1)/t]$. But $\varphi(t) = \sum_{n=0}^{\infty} \alpha^n P(\{N_t = n\})$, $\psi(\alpha) = \varphi'(0)$ and $\psi(t) = \lim_{t\to 0}(\varphi(t) - 1)/t = -\lambda + \lambda\alpha$. Therefore,

$$\varphi(t) = \sum_{n=0}^{\infty} \alpha^n P(\{N_t = n\}) = \mathrm{e}^{-\lambda t} \sum_{n=0}^{\infty} \frac{(\lambda t)^n \alpha^n}{n!},$$

which implies that $P(\{N_t = n\}) = \mathrm{e}^{-\lambda t} \cdot (\lambda t)^n / n!$. □

Remark 5.4. The parameter λ associated to a Poisson process N such that $P(\{N_t = n\}) = \mathrm{e}^{-\lambda t} \cdot (\lambda t)^n / n!$ is called the intensity or the arrival rate of the process N. □

Theorem 5.3. *A Poisson process N with intensity λ satisfies: (i) $E[N_t] = \lambda t$ and (ii) $D[N_t] = \mathrm{Var}(N_t) = \lambda t$.* □

Remark 5.5. A counting process N without explosion is a Poisson process if and only if there is $\lambda \geq 0$ such that $E[N_t] < \infty$ and $E[N_t - N_s | \mathcal{F}_s] = \lambda(t - s)$ for every $0 \leq s \leq t < \infty$. □

Theorem 5.4. *Let N be a Poisson process with intensity λ. Then $N_t - \lambda t$ and $(N_t - \lambda t)^2 - \lambda t$ are martingales.*

Proof. By Theorem 5.3, we have $E[N_t - \lambda t] = E[(N_t - \lambda t)^2 - \lambda t] = 0$. By the independence of $N_t - N_s$ on $\mathcal{F}_s$, for every $0 \leq s \leq t < \infty$, we get $E[(N_t - \lambda t) - (N_s - \lambda s) | \mathcal{F}_s] = E[(N_t - \lambda t) - (N_s - \lambda s)] = 0$. Similarly, for every $0 \leq s \leq t < \infty$, we also get $E[(N_t - \lambda t)^2 - \lambda t | \mathcal{F}_s] = (N_s - \lambda s)^2 - \lambda s$. □

An m-dimensional $\mathbb{F}$-adapted process $B = (B_t)_{t\geq 0}$ on $\mathcal{P}_{\mathbb{F}}$ is called an m-dimensional $\mathbb{F}$-Brownian motion or a Brownian motion on $\mathcal{P}_{\mathbb{F}}$ if (1) for every $0 \leq s \leq t < \infty$, $B_t - B_s$ is independent of $\mathcal{F}_s$ and (2) for every $0 \leq s \leq t < \infty$, $B_t - B_s$ is a Gaussian random variable with mean zero and variance matrix $(t-s)C$ for a given nonrandom matrix C.

Remark 5.6. A Brownian motion starts at $x \in \mathbb{R}^d$ if $P(\{B_0 = x\}) = 1$. □

Remark 5.7. The existence of an $\mathbb{F}$-Brownian motion can be proved using a path-space construction together with Kolmogorov's extension theorem. But it is not true that there exists a Brownian motion on every complete filtered probability space $\mathcal{P}_{\mathbb{F}}$. Sometimes, the underlying probability space $\mathcal{P}_{\mathbb{F}}$ is just too small. Nevertheless, it can be proved that there exists a complete filtered probability space such that there exists a Brownian motion on that space. □

Corollary 5.2. *If $B = (B_t)_{t\geq 0}$ is an $\mathbb{F}$-Brownian motion on $\mathcal{P}_{\mathbb{F}}$ and $E|B_0| < \infty$, then it is a martingale on $\mathcal{P}_{\mathbb{F}}$.*

Proof. For every $t \geq 0$, one has $E|B_t| \leq E|B_t - B_0| + E|B_0| \leq \sqrt{E|B_t - B_0|^2} + E|B_0| < \infty$. Furthermore, for every $0 \leq s \leq t < \infty$, we have $E[B_t - B_s|\mathcal{F}_s] = 0$ a.s., which implies $E[B_t|\mathcal{F}_s] = B_s$. □

Theorem 5.5. *Let $B = (B_t)_{t\geq 0}$ be a Brownian motion on $\mathcal{P}_{\mathbb{F}}$. Then there exists a modification of B that has continuous paths a.s.*

Proof. It can be verified that if X is a random variable with normal distribution $N(0, \sigma^2)$, then

$$EX^{2n} = \frac{(2n)!}{2^n n!}\sigma^{2n} \qquad \text{and} \qquad EX^{2n+1} = 0 \qquad \text{for} \qquad n = 0, 1, 2, \ldots .$$

Then in particular, it follows that $E|B_t - B_s|^{2n} = C_n|t-s|^n$ with any constant C_n. The result now follows by Kolmogorov's continuity theorem. □

In what follows, we shall always assume that we are using the version of a Brownian motion with continuous paths. We shall also assume that we always have to deal with a Brownian motion with a matrix C equal to the identity matrix. We have the following results dealing with some properties of Brownian motions.

Theorem 5.6. *(i) For every $\alpha < 1/2$, almost all paths of Brownian motions are Hölder continuous with exponent α. (ii) For every $\alpha > 1/2$, almost all paths of Brownian motions are nowhere Hölder continuous with exponent α.* □

Corollary 5.3. *(i) Almost all sample paths of a Brownian motion are nowhere differentiable. (ii) Almost all sample paths of a Brownian motion have infinite variation on any finite interval.*

Proof. (i) If the function $\mathbb{R}^+ \ni t \to B_t(\omega) \in \mathbb{R}^m$ were differentiable at a point $t_0 \in (0, \infty)$ for $\omega \in \Omega_0 \subset \Omega$ with $\Omega_0 \in \mathcal{F}$ such that $P(\Omega_0) > 0$, then it would be Lipschitz continuous at that point, which is a contradiction to (ii) of Theorem 5.6.

(ii) Since a function $f : \mathbb{R}^+ \to \mathbb{R}^m$ with finite variation is almost everywhere differentiable, then (ii) is a consequence of (i). □

Theorem 5.7. *Let $B = (B_t)_{t\geq 0}$ be a one-dimensional $\mathbb{F}$-Brownian motion on $\mathcal{P}_{\mathbb{F}}$ with $B_0 = 0$ a.s. Then the process $M = (M_t)_{t\geq 0}$ with $M_t = B_t^2 - t$ is a martingale.*

Proof. We have $E[M_t] = E[B_t^2 - t] = 0$, $E[M_t - M_s|\mathcal{F}_s] = E[B_t^2 - B_s^2 - (t - s)|\mathcal{F}_s]$, and $E[B_t B_s|\mathcal{F}_s] = B_s E[B_t|\mathcal{F}_s] = B_s^2$. Hence it follows that $E[M_t - M_s|\mathcal{F}_s] = E[(B_t - B_s)^2 - (t-s)|\mathcal{F}_s] = E[(B_t - B_s)^2] - (t-s) = 0$ for every $0 \le s \le t < \infty$. Therefore, $E[M_t|\mathcal{F}_s] = M_s$ a.s. for every $0 \le s \le t < \infty$. □

Theorem 5.8. *Let $X = (X_t)_{t\ge 0}$ be an m-dimensional continuous $\mathbb{F}$-adapted process on $\mathcal{P}_\mathbb{F}$ such that (i) $E[X_t - X_s|\mathcal{F}_s] = 0$ a.s. and (ii) $E[(X_t^i - X_s^i)(X_t^j - X_s^j)] = \delta_{ij}(t-s)$ a.s. for every $0 \le s \le t < \infty$, where $\delta_{ij} = 0$ for $i \ne j$ and $\delta_{ij} = 1$ for $i = j$. Then X is an $\mathbb{F}$-Brownian motion on $\mathcal{P}_\mathbb{F}$.* □

Remark 5.8. It can be verified that an m-dimensional $\mathbb{F}$-Brownian motion on $\mathcal{P}_\mathbb{F}$ satisfies the strong Markov property with the stationary transition function

$$P_t(x, A) = \frac{1}{(2\pi t)^{n/2}} \int_A \exp\left[-\frac{|x-y|^2}{2t}\right] \mathrm{d}y. \qquad \square$$

Remark 5.9. It can be verified that if $B = (B_t)_{t\ge 0}$ is an m-dimensional Brownian motion on $\mathcal{P}_\mathbb{F}$ and $t_0 \ge 0$, then the process $\tilde{B} = (\tilde{B}_t)_{t\ge 0}$ with $\tilde{B}_t = B_{t_0+t} - B_t$ for $t \ge 0$ is an $\mathbb{F}$-Brownian motion on $\mathcal{P}_\mathbb{F}$. □

Remark 5.10. A real Brownian motion can be defined on a given probability space $(\Omega, \mathcal{F}, P)$ as a continuous stochastic process $\beta = (\beta_t)_{t\ge 0}$ such that $\beta_0 = 0$ and β is a stationary process with independent Gaussian increments such that $E[\beta_t - \beta_s] = 0$ and $E[(\beta_t - \beta_s)^2] = \sigma^2(t-s)$ for every $0 \le s < t < \infty$. In such a case, we can define a filtration $\mathbb{F}^\beta = (F_t^\beta)_{t\ge 0}$ with an augmented σ-algebra F_t^β defined for every $t \ge 0$ by a family $\{\beta_s : 0 \le s \le t\}$ of random variables, i.e., $\mathcal{F}_t^\beta = \bigcap_{s>t} \sigma(F_s^\beta \cup \mathcal{N})$, where $F_t^\beta = \sigma\{\beta_s : s \le t\}$ and $\mathcal{N}$ is the collection of all P-null sets in $\mathcal{F}$. It can be verified that β is a real $\mathbb{F}^\beta$-Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}^\beta, P)$. □

We shall prove that the above-defined filtration $(F_t^\beta)_{t\ge 0}$ is continuous, i.e., that

$$F_{t-}^\beta = F_t^\beta = F_{t+}^\beta \text{ where } F_{t-}^\beta = \sigma(\bigcup_{s<t} F_s^\beta) \;\; F_{t+}^\beta = \bigcap_{s>t} \sigma(F_s^\beta) \text{ and } F_{0-}^\beta = F_0^\beta.$$

Theorem 5.9. *Let $(\Omega, \mathcal{F}, P)$ be a probability space such that a real Brownian motion $\beta = (\beta_t)_{t\ge 0}$ can be defined on this space. The filtration $\mathbb{F}^\beta = (F_t^\beta)_{t\ge 0}$ defined in Remark 5.10 is continuous.*

Proof. Let us observe that $F_{t-}^\beta = F_t^\beta$ follows immediately from continuity of the Brownian motion β. Indeed, we have

$$F_{t-}^\beta = \sigma\left(\bigcup_{s>t} F_s^\beta\right) \quad \text{and} \quad F_t^\beta = \sigma\left(\bigcup_{s>t} F_s^\beta \cup F^\beta(t)\right),$$

where $F^{\beta}(t) = \sigma(\beta_t)$. But $\beta_t = \lim_{r\uparrow t}\beta_r$, where $r \in Q$. Then $F^{\beta}(t) \subset \sigma(\bigcup_{s>t} F_s^{\beta})$. Thus $F_{t-}^{\beta} = F_t^{\beta}$.

To verify that $F_{t+}^{\beta} = F_t^{\beta}$, let us observe that for every $t > s$, one has $E[\exp\{i\lambda(\beta_t - \beta_s)\}|F_s^{\beta}] = \exp\{-\lambda^2/2(t-s)\}$. Indeed, it suffices to observe that a differential equation $z'(t) = -\lambda^2/2z(t)$ possesses on the interval $[s,\infty)$ exactly one solution satisfying the initial condition $z(s) = \exp\{i\lambda\beta_s\}$. It is defined by $z(t) = z(s)\exp\{-\lambda^2/2t\}$. In particular, we can verify that $z(t) = E[\exp\{i\lambda\beta_t\}|F_s^{\beta}]$ satisfies the above differential equation. Therefore, $E[\exp\{i\lambda(\beta_t - \beta_s)\}|F_s^{\beta}] = \exp\{-\lambda^2/2(t-s)\}$. Hence it follows that

$$E[\exp\{i\lambda\beta_t\}|F_s^{\beta}] = E[E[\exp\{i\lambda\beta_t\}|F_s^{\beta}]|F_s^{\beta}] = \exp\{i\lambda\beta_s - \lambda^2/2(t-s)\}.$$

Let $\varepsilon > 0$ be such that $0 < \varepsilon < t-s$. Then

$$\begin{aligned} E[\exp\{i\lambda\beta_t\}|F_{s+}^{\beta}] &= E[E[\exp\{i\lambda\beta_t\}|F_{s+\varepsilon}^{\beta}]|F_{s+}^{\beta}] \\ &= E[\exp\{i\lambda\beta_{s+\varepsilon} - \lambda^2/2(t-s-\varepsilon)\}|F_{s+}^{\beta}]. \end{aligned}$$

Hence, in the limit $\varepsilon \downarrow 0$, it follows that

$$E[\exp\{i\lambda\beta_t\}|F_{s+}^{\beta}] = E[\exp\{i\lambda\beta_s - \lambda^2/2(t-s)\}|F_{s+}^{\beta}] = \exp\{i\lambda\beta_s - \lambda^2/2(t-s)\}.$$

Then $E[\exp\{i\lambda\beta_t\}|F_s^{\beta}] = E[\exp\{i\lambda\beta_t\}|F_{s+}^{\beta}]$. Therefore, for every measurable and bounded function $f : \mathbb{R} \to \mathbb{R}$, it follows that $E[f(\beta_t)|F_s^{\beta}] = E[f(\beta_t)|F_{s+}^{\beta}]$. Let $s < t_1 < t_2$ and $f_1, f_2 : \mathbb{R} \to \mathbb{R}$ be measurable and bounded. The above equalities imply $E[f_2(\beta_{t_2})f_1(\beta_{t_1})|F_s^{\beta}] = E[f_2(\beta_{t_2})f_1(\beta_{t_1})|F_{s+}^{\beta}]$. In a similar way, for $s < t_1 < \cdots < t_n$ and measurable bounded functions $f_1, \ldots, f_n : \mathbb{R} \to \mathbb{R}$, we obtain $E[\{f_1(\beta_{t_1})\cdot\cdots\cdot f_n(\beta_{t_n})\}|F_s^{\beta}] = E[\{f_1(\beta_{t_1})\cdot\cdots\cdot f_n(\beta_{t_n})\}|F_{s+}^{\beta}]$. Therefore, for every $t > s$ and F_t^{β}-measurable bounded random variable η, we have $E[\eta|F_s^{\beta}] = E[\eta|F_{s+}^{\beta}]$ a.s. Taking in particular the F_{s+}^{β}-measurable random variable η,, we get $E[\eta|F_s^{\beta}] = \eta$ a.s. Hence, by the properties of the filtration $\mathbb{F}^{\beta}$, it follows that η is F_s^{β}-measurable. Then $F_{s+}^{\beta} \subset F_s^{\beta}$. It is clear that we also have $F_s^{\beta} \subset F_{s+}^{\beta}$. Then $F_s^{\beta} = F_{s+}^{\beta}$. □

6 Stochastic Integrals

Stochastic integrals with respect to finite-variation stochastic processes can be thought of as an extension of path-by-path Lebesgue–Stieltjes integration. Unfortunately, in practical applications we have to deal with processes with almost all paths of infinite variation on compacts. The most important example of such processes is Brownian motion. Therefore, it was important to define stochastic integrals in a way other than via Lebesgue–Stieltjes integration. A new idea for such a definition was introduced by K. Itô. We shall present it for some special classes of stochastic processes with respect to a Brownian motion.

Given a filtered probability space $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ with $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$, we say that a d-dimensional process $X = (X_t)_{t\geq 0}$ on $\mathcal{P}_{\mathbb{F}}$ is $\mathbb{F}$-nonanticipative if it is measurable and $\mathbb{F}$-adapted. In what follows, we shall denote by $\mathcal{M}^2_{\mathbb{F}}(a,b)$ the family of restrictions of all $\mathbb{F}$-nonanticipative processes to the interval $[a,b]$ such that $P(\{\int_a^b |X_t|^2 \mathrm{d}t < \infty\}) = 1$. By $\mathcal{L}^2_{\mathbb{F}}(a,b)$, we denote the subset of $\mathcal{M}^2_{\mathbb{F}}(a,b)$ of all $X \in \mathcal{M}^2_{\mathbb{F}}(a,b)$ such that $E[\int_a^b |X_t|^2 \mathrm{d}t] < \infty$.

A stochastic process $X \in \mathcal{M}^2_{\mathbb{F}}(a,b)$ is called simple if there exists a partition $a = t_0 < t_1 < \cdots < t_r = b$ of $[a,b]$ such that $X_t = X_{t_i}$ for $t_i \leq t < t_{i+1}$ with $i = 0,1,\ldots,r-2$ and $X_t = X_{t_{r-1}}$ for $t_{r-1} \leq t \leq b$. The class of all simple processes of $\mathcal{M}^2_{\mathbb{F}}(a,b)$ is denoted by $\mathcal{S}_{\mathbb{F}}(a,b)$.

Corollary 6.1. *Every $F \in \mathcal{S}_{\mathbb{F}}(a,b)$ can be presented by $F = \sum_{i=0}^{r-2} \mathbb{1}_{[t_i,t_{i+1})}\varphi_i + \mathbb{1}_{[t_{r-1},b]}\varphi_{r-1}$, where φ_i is an $\mathcal{F}_{t_i}$-measurable $\mathbb{R}^d$-random variable on $\mathcal{P}_{\mathbb{F}}$ for $i = 0,1,\ldots,r-1$.* □

Let $B = (B_t)_{t\geq 0}$ be a one-dimensional $\mathbb{F}$-Brownian motion on $\mathcal{P}_{\mathbb{F}}$ such that $B_a = 0$. By a stochastic Itô integral of $F \in \mathcal{S}_{\mathbb{F}}(a,b)$ with respect to a Brownian motion B we mean an $\mathbb{R}^d$-random variable on $\mathcal{P}_{\mathbb{F}}$, denoted by $\int_a^b F_t \mathrm{d}B_t$ and defined by $\int_a^b F_t \mathrm{d}B_t = \sum_{i=0}^{r-1} \varphi_i (B_{t_i+1} - B_{t_i})$, where for $i = 0,1,\ldots,r-1$, the $\mathcal{F}_{t_i}$-measurable random variables φ_i are such that $F = \sum_{i=0}^{r-2} \mathbb{1}_{[t_i,t_{i+1})}\varphi_i + \mathbb{1}_{[t_{r-1},b]}\varphi_{r-1}$.

Lemma 6.1. *Let F, F,^{1} $F^2 \in \mathcal{S}_{\mathbb{F}}(a,b)$, $\lambda_1, \lambda_2 \in \mathbb{R}$, $\varepsilon > 0$, and $N > 0$. Then:*

(i) $\int_a^b (\lambda_1 F_t^1 + \lambda_2 F_t^2)\mathrm{d}B_t = \lambda_1 \int_a^b F_t^1 \mathrm{d}B_t + \lambda_2 \int_a^b F_t^2 \mathrm{d}B_t$ *a.s.*
(ii) If $F \in \mathcal{S}_{\mathbb{F}}(a,b) \cap \mathcal{L}^2_{\mathbb{F}}(a,b)$, then $E[\int_a^b F_t \mathrm{d}B_t] = 0$.
(iii) $P(\{|\int_a^b F_t \mathrm{d}B_t| > \varepsilon\}) \leq P(\{\int_a^b |F_t|^2 \mathrm{d}t > N\}) + N/\varepsilon^2$.
(iv) If $F \in \mathcal{S}_{\mathbb{F}}(a,b) \cap \mathcal{L}^2_{\mathbb{F}}(a,b)$, then $E|\int_a^b F_t \mathrm{d}B_t|^2 = E\int_a^b |F_t|^2 \mathrm{d}t$.

Proof. Conditions (i) and (ii) follow immediately from the definition of the Itô integral. Assume $F \in \mathcal{S}_{\mathbb{F}}(a,b) \cap \mathcal{L}^2_{\mathbb{F}}(a,b)$. By the above definition of the Itô integral, it follows that

$$
\begin{aligned}
E\left|\int_a^b F_t \mathrm{d}B_t\right|^2 &= \sum_{i=0}^{r-1} E[|F_{t_i}|^2 (B_{t_{i+1}} - B_{t_i})^2] \\
&= \sum_{i=0}^{r-1} E|F_{t_i}|^2 E[(B_{t_{i+1}} - B_{t_i})^2] \\
&= \sum_{i=0}^{r-1} E|F_{t_i}|^2 (t_{i+1} - t_i) = E\int_a^b |F_t|^2 \mathrm{d}t.
\end{aligned}
$$

For the proof of (iii), let $\Phi_N(t)$ be defined by

$$\Phi_N(t) = \begin{cases} F_t & if \quad t_k \leq t < t_{k+1} \text{ and } \sum_{j=0}^{k} |F_{t_j}|^2 (t_{j+1} - t_j) \leq N \\ 0 & if \quad t_k \leq t < t_{k+1} \text{ and } \sum_{j=0}^{k} |F_{t_j}|^2 (t_{j+1} - t_j) > N \end{cases}$$

for $k = 0, 1, 2, \ldots, r-1$, where $a = t_0 < t_1 < \cdots < t_r = b$. The process $\Phi_N = (\Phi_N(t)_{a \leq t \leq b}$ belongs to $\mathcal{S}_{\mathbb{F}}(a,b) \cap \mathcal{L}^2_{\mathbb{F}}(a,b)$ and $\int_a^b |\Phi_N(t)|^2 \mathrm{d}t = \sum_{j=0}^{\nu} |F_{t_j}|^2 (t_{j+1} - t_j)$, where ν is the largest integer such that $\sum_{j=0}^{k} |F_{t_j}|^2 (t_{j+1} - t_j) \leq N$, $\nu \leq r-1$. Hence it follows that $E \int_a^b |\Phi_N(t)|^2 \mathrm{d}t \leq N$. Further, $F_t - \Phi_N(t) = 0$ for all $t \in [a,b]$ if $\int_a^b |\Phi_N(t)|^2 \mathrm{d}t < N$. Therefore,

$$P\left(\left\{\left|\int_a^b F_t \mathrm{d}B_t\right| > \varepsilon\right\}\right) \leq P\left(\left\{\left|\int_a^b \Phi_N(t) \mathrm{d}B_t\right| > \varepsilon\right\}\right) + P\left(\left\{\int_a^b |F_t|^2 \mathrm{d}t > N\right\}\right).$$

By Chebyshev's inequality, the first integral on the right-hand side is bounded by $(1/\varepsilon^2) E |\int_a^b \Phi_N(t) \mathrm{d}B_t|^2 \leq N/\varepsilon^2$. Therefore, (iii) is satisfied. □

To extend the above definition of stochastic integrals on the whole space $\mathcal{M}^2_{\mathbb{F}}(a,b)$, we need the following results.

Lemma 6.2. *Let* $F \in \mathcal{M}^2_{\mathbb{F}}(a,b)$. *Then:*

(i) There exists a sequence $(G^n)_{n=1}^{\infty}$ *of continuous processes* $G^n \in \mathcal{M}^2_{\mathbb{F}}(a,b)$ *such that* $\lim_{n\to\infty} \int_a^b |G_t^n - F_t|^2 \mathrm{d}t = 0$ *a.s.*

(ii) There exists a sequence $(F^k)_{k=1}^{\infty}$ *of* $\mathcal{S}_{\mathbb{F}}(a,b)$ *such that* $\int_a^b |F_t^n - F_t|^2 \mathrm{d}t \xrightarrow{\mathrm{P}} 0$ *as* $k \to \infty$.

Proof (Sketch of proof). Let

$$\rho(t) = \begin{cases} c \exp[-1/(1-t^2)] & if \quad |t| \leq 1 \\ 0 & if \quad |t| > 1 \end{cases}$$

with $c > 0$ be such that $\int_{-\infty}^{+\infty} \rho(t) \mathrm{d}t = 1$. For every $\varepsilon \in (0, 1/2)$, we define

$$(J_\varepsilon F)(t) = \frac{1}{\varepsilon} \int_a^b \rho\left(\frac{t - s - \varepsilon}{\varepsilon}\right) \tilde{F}_s \mathrm{d}s,$$

where $\tilde{F}_s = F_s$ for $s \in [a,b]$ and $\tilde{F}_s = 0$ for $s \in \mathbb{R} \setminus [a,b]$. For every fixed $\omega \in \Omega$, we can select nonrandom functions u_n such that $u_n(t) = 0$ for $t \in \mathbb{R} \setminus [a,b]$ and $\int_a^b |u_n(t) - F_t(\omega)|^2 \mathrm{d}t \to 0$ as $n \to \infty$. It can be verified that $(J_\varepsilon u_n)(t) \to u_n(t)$ uniformly in $t \in [a,b]$ as $\varepsilon \to 0$ and hence that

$\limsup_{\varepsilon\to 0}\int_a^b |(J_\varepsilon F)(t) - F_t|^2 \mathrm{d}t = 0$ a.s. Taking $G^n = J_{1/n}F$, we obtain (i). To prove (ii), we take $h^k_{n,m}(t) = G^n(k/m)$ if $a + k/m \le t \le a + (k+1)/m$ for $k \ge 1$. For every $\delta > 0$, there are $n = n_0$ and $m = m_0$ such that

$$P\left(\left\{\int_a^b |F_t - G_t^{n_0}|^2 \mathrm{d}t > \frac{\delta}{2}\right\}\right) < \frac{\delta}{2} \quad \text{and}$$

$$P\left(\left\{\int_a^b |G_t^{n_0} - h^k_{n_0,m_0}|^2 \mathrm{d}t > \frac{\delta}{2}\right\}\right) < \frac{\delta}{2}$$

for $k \ge 1$. Hence it follows that $P(\{\int_a^b |F_t - h^k_{n_0,m_0}|^2 \mathrm{d}t > \delta\}) < \delta$ for $k \ge 1$. Taking $\delta = 1/k$ and denoting correspondingly $h^k_{n_0,m_0}$ by F^k, we obtain $F^k \in \mathcal{S}_{\mathbb{F}}(a,b)$ and $\int_a^b |F_t^k - F_t|^2 \mathrm{d}t \xrightarrow{P} 0$ as $k \to \infty$. □

Lemma 6.3. *Let $F \in \mathcal{L}^2_{\mathbb{F}}(a,b)$. Then:*

(i) There exists a sequence $(H^n)_{n=1}^\infty$ of continuous processes $H^n \in \mathcal{L}^2_{\mathbb{F}}(a,b)$ such that $E\int_a^b |H_t^n - F_t|^2 \mathrm{d}t \to 0$ as $n \to \infty$.

(ii) There exists a sequence $(h^n)_{n=1}^\infty$ of $\mathcal{S}_{\mathbb{F}}(a,b) \cap \mathcal{L}^2_{\mathbb{F}}(a,b)$ such that $E\int_a^b |h_t^n - F_t|^2 \mathrm{d}t \to 0$ as $n \to \infty$.

Proof. Let G^n be as in Lemma 6.2 and let $N \ge 1$. Put

$$\Phi_N(t) = \begin{cases} t & if \quad |t| \le N \\ Nt/|t| & if \quad |t| > N\,. \end{cases}$$

We obtain $|\Phi_N(t) - \Phi_N(s)| \le |t - s|$. Therefore, $\int_a^b |\Phi_N(F_t) - \Phi_N(G_t^n)|^2 \mathrm{d}t \to 0$ a.s. as $n \to \infty$. Hence, by the Lebesgue dominated convergence theorem, it follows that $E\int_a^b |\Phi_N(G_t^n) - F_t|^2 \mathrm{d}t \to 0$ as $n \to \infty$. Then for every $k = 1, 2, \ldots$, there are $N = N(k)$ and $n = n(k, N)$ such that $E\int_a^b |\Phi_N(G_t^n) - F_t|^2 \mathrm{d}t < 1/k$. Taking $H^k = \Phi_N(G^n)$ with $N = N(k)$ and $n = n(k, N)$, we can see that (i) is satisfied. The proof of (ii) is similar to thr of (ii) of Lemma 6.2. The h^n are of the form $\Phi_N(F^n)$, where F^n are as in Lemma 6.2. □

Lemma 6.4. *For every $F \in \mathcal{M}^2_{\mathbb{F}}(a,b)$ and every sequence $(F^n)_{n=1}^\infty$ of $\mathcal{S}_{\mathbb{F}}(a,b)$ such that $\int_a^b |F_t^n - F_t|^2 \mathrm{d}t \xrightarrow{P} 0$ as $n \to \infty$, there is an $\mathbb{R}^d$-random variable $J(F)$ on $\mathcal{P}_{\mathbb{F}}$, independent of the particular choice of sequence $(F^n)_{n=1}^\infty$, such that $\int_a^b F_t^n \mathrm{d}B_t \xrightarrow{P} J(F)$ as $n \to \infty$.*

Proof. Let $(F^n)_{n=1}^\infty$ be a sequence of $\mathcal{S}_{\mathbb{F}}(a,b)$ such that $\int_a^b |F_t^n - F_t|^2 \mathrm{d}t \xrightarrow{P} 0$ as $n \to \infty$. Hence it follows that $\int_a^b |F_t^n - F_t^m|^2 \mathrm{d}t \xrightarrow{P} 0$ as $n, m \to \infty$. By virtue of Lemma 6.1, for every $\varepsilon > 0$ and $\rho > 0$, we get

$$P\left(\left\{\left|\int_a^b F_t^n \mathrm{d}B_t - \int_a^b F_t^m \mathrm{d}B_t\right| > \varepsilon\right\}\right) \leq \rho + P\left(\left\{\int_a^b |F_t^n - F_t^m|^2 \mathrm{d}t > \varepsilon^2\rho\right\}\right).$$

Then the sequence $(\int_a^b F_t^n \mathrm{d}B_t)_{n=1}^{\infty}$ is a Cauchy sequence with respect to convergence in probability. By completeness with respect to convergence in probability of the space of all $\mathbb{R}^d$-random variables on $\mathcal{P}_{\mathbb{F}}$, there is an $\mathbb{R}^d$-random variable $J(F)$ on $\mathcal{P}_{\mathbb{F}}$ such that $\int_a^b F_t^n \mathrm{d}B_t \xrightarrow{P} J(F)$ as $n \to \infty$.

Suppose $(G^n)_{n=1}^{\infty}$ is a sequence of $\mathcal{S}_{\mathbb{F}}(a,b)$ such that $\int_a^b |G_t^n - F_t|^2 \mathrm{d}t \xrightarrow{P} 0$ as $n \to \infty$. The sequence $(H^n)_{n=1}^{\infty}$ defined by $H^{2n} = F^n$ and $H^{2n+1} = G^n$ satisfies $\int_a^b |H_t^n - F_t|^2 \mathrm{d}t \xrightarrow{P} 0$ as $n \to \infty$. Hence it follows that the sequence $(\int_a^b H_t^n \mathrm{d}B_t)_{n=1}^{\infty}$ converges in probability to a random variable $K(F)$. Therefore, its subsequence $(\int_a^b H_t^{2n} \mathrm{d}B_t)_{n=1}^{\infty}$ also converges in probability to $K(F)$. By the definition of H^{2n}, it follows that $J(F) = K(F)$ a.s. □

The random variable $J(F)$ defined in Lemma 6.4 is denoted by $\int_a^b F_t \mathrm{d}B_t$ and said to be an Itô integral of $F \in \mathcal{M}_{\mathbb{F}}^2(a,b)$ with respect to the $\mathbb{F}$-Brownian motion $B = (B_t)_{t \geq 0}$. In particular, $\int_a^b F_t \mathrm{d}B_t \in \mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{R}^d)$ for $F \in \mathcal{L}_{\mathbb{F}}^2(a,b)$.

Theorem 6.1. *Let $F, F^1, F^2 \in \mathcal{M}_{\mathbb{F}}^2(a,b)$, $\lambda_1, \lambda_2 \in \mathbb{R}$, $\varepsilon > 0$, and $N > 0$. Then $\lambda_1 F_t^1 + \lambda_2 F_t^2 \in \mathcal{M}_{\mathbb{F}}^2(a,b)$, and the following relations are satisfied:*

(i) $\int_a^b (\lambda_1 F_t^1 + \lambda_2 F_t^2) \mathrm{d}B_t = \lambda_1 \int_a^b F_t^1 \mathrm{d}B_t + \lambda_2 \int_a^b F_t^2 \mathrm{d}B_t$ *a.s.*

(ii) $P(\{|\int_a^b F_t \mathrm{d}B_t| > \varepsilon\}) \leq P(\{\int_a^b |F_t|^2 \mathrm{d}t > N\}) + N/\varepsilon^2$.

Proof. The equality (i) is a consequence of the definition of the Itô integral and Lemma 6.1. For the proof of (ii), let us assume that $(F^n)_{n=1}^{\infty}$ is a sequence of $\mathcal{S}_{\mathbb{F}}(a,b)$ such that $\lim_{n\to\infty} \int_a^b |F_t^n - F_t|^2 \mathrm{d}t = 0$ a.s. By the definition of $\int_a^b F_t \mathrm{d}B_t$, we have $\int_a^b F_t^n \mathrm{d}B_t \xrightarrow{P} \int_a^b F_t \mathrm{d}B_t$ as $n \to \infty$. By virtue of Lemma 6.1, we have

$$P\left(\left\{\left|\int_a^b F_t^n \mathrm{d}B_t\right| > \varepsilon'\right\}\right) \leq P\left(\left\{\int_a^b |F_t^n|^2 \mathrm{d}t > N'\right\}\right) + \frac{N'}{(\varepsilon')^2}$$

for $\varepsilon > \varepsilon'$ and $N < N'$. Passing to the limit $n \to \infty$, using the above property of the sequence $(\int_a^b F_t^n \mathrm{d}B_t)_{n=1}^{\infty}$, and taking $\varepsilon' \uparrow \varepsilon$ and $N' \downarrow N$, we obtain

$$P\left(\left\{\left|\int_a^b F_t \mathrm{d}B_t\right| > \varepsilon\right\}\right) \leq P\left(\left\{\int_a^b |F_t|^2 \mathrm{d}t > N\right\}\right) + \frac{N}{\varepsilon^2}.$$

□

Theorem 6.2. *Let* $F \in \mathcal{M}^2_{\mathbb{F}}(a,b)$ *and let* $(F^n)_{n=1}^{\infty}$ *be a sequence of* $\mathcal{M}^2_{\mathbb{F}}(a,b)$ *such that* $\int_a^b |F_t^n - F_t|^2 \mathrm{d}t \xrightarrow{P} 0$ *as* $n \to \infty$. *Then* $\int_a^b F_t^n \mathrm{d}B_t \xrightarrow{P} \int_a^b F_t \mathrm{d}B_t$ *as* $n \to \infty$.

Proof. By Theorem 6.1, for every $\varepsilon > 0$ and $\rho > 0$, one has

$$P\left(\left\{\left|\int_a^b (F_t^n - F_t)\mathrm{d}B_t\right| > \varepsilon\right\}\right) \le \rho + P\left(\left\{\int_a^b |F_t^n - F_t|^2 \mathrm{d}t > \varepsilon^2 \rho\right\}\right)$$

for $n = 1, 2, \ldots$. From this and the properties of the sequence $(F^n)_{n=1}^{\infty}$, the result follows. □

Theorem 6.3. *If* $F \in \mathcal{L}^2_{\mathbb{F}}(a,b)$, *then (i)* $E\int_a^b F_t \mathrm{d}B_t = 0$ *and (ii)* $E|\int_a^b F_t \mathrm{d}B_t|^2 = E\int_a^b |F_t|^2 \mathrm{d}t$.

Proof. Let $(F^n)_{n=1}^{\infty}$ be a sequence of $\mathcal{S}_{\mathbb{F}}(a,b) \cap \mathcal{L}^2_{\mathbb{F}}(a,b)$ such that $E\int_a^b |F_t^n - F_t|^2 \mathrm{d}t \to 0$ as $n \to \infty$. This implies that $E\int_a^b |F_t^n|^2 \mathrm{d}t \to E\int_a^b |F_t|^2 \mathrm{d}t$ as $n \to \infty$. By virtue of Lemma 6.1, we get $E\int_a^b F_t^n \mathrm{d}B_t = 0$ and $E|\int_a^b F_t^n \mathrm{d}B_t|^2 = E\int_a^b |F_t^n|^2 \mathrm{d}t$ for every $n = 1, 2, \ldots$. Hence in particular, it follows that $(\int_a^b F^n \mathrm{d}B_t)_{n=1}^{\infty}$ is a Cauchy sequence of $\mathbb{L}^2(\Omega, \mathcal{F}, P, \mathbb{R}^d)$. By virtue of Theorem 6.2, it converges in probability to $\int_a^b F_t \mathrm{d}B_t$, which implies that $E|\int_a^b F_t^n \mathrm{d}B_t|^2 \to E|\int_a^b F_t \mathrm{d}B_t|^2$ as $n \to \infty$. Then $E|\int_a^b F_t \mathrm{d}B_t|^2 = \lim_{n\to\infty} E|\int_a^b F_t^n \mathrm{d}B_t|^2 = \lim_{n\to\infty} E\int_a^b |F_t^n|^2 \mathrm{d}t = E\int_a^b |F_t|^2 \mathrm{d}t$. □

Remark 6.1. For every $F \in \mathcal{L}^2_{\mathbb{F}}(a,b)$, we can define the integral $E\int_a^b F_t \mathrm{d}t$ as the integral of a $\Sigma_{\mathbb{F}}$-measurable function on $[a,b] \times \Omega$ with respect to the product measure $\mathrm{d}t \times P$ a.s., where $\sum_{\mathbb{F}}$ denotes a σ-algebra of $\mathbb{F}$-nonanticipative subsets of $[a,b] \times \Omega$. □

Corollary 6.2. *For every* $G \in \mathbb{L}^2([a,b] \times \Omega, \Sigma_{\mathbb{F}}, \mathbb{R}^d)$, $F \in \mathcal{M}^2_{\mathbb{F}}(a,b)$, *and* $\psi \in \mathbb{L}^2(\Omega, \mathcal{F}_a, \mathbb{R})$, *one has* $E\int_a^b (\psi \cdot G)_t \mathrm{d}t = E[\psi \int_a^b G_t \mathrm{d}t]$ *and* $\int_a^b (\psi \cdot F)_t \mathrm{d}B_t = \psi \int_a^b F_t \mathrm{d}B_t$.

Proof. It is clear that $\psi \cdot G \in \mathbb{L}^2([a,b] \times \Omega, \Sigma_{\mathbb{F}}, \mathbb{R}^d)$ and $\psi \cdot F \in \mathcal{M}^2_{\mathbb{F}}(a,b)$. Immediately from Fubini's theorem and properties of the integral $\int_a^b (\psi \cdot G)_t \mathrm{d}t$, one obtains

$$E\int_a^b (\psi \cdot G)_t \mathrm{d}t = E\left[\int_a^b (\psi \cdot G)_t \mathrm{d}t\right] = E\left[\psi \int_a^b G_t \mathrm{d}t\right].$$

Let $(F^n)_{n=1}^{\infty}$ be a sequence of $\mathcal{S}_{\mathbb{F}}(a,b)$ such that $\int_a^b |F_t^n - F_t|^2 \mathrm{d}t \xrightarrow{\mathrm{P}} 0$ as $n \to \infty$. It is clear that $\psi \cdot F^n \in \mathcal{S}_{\mathbb{F}}(a,b)$ for every $n \ge 1$ and $\int_a^b |(\psi \cdot F^n)_t - (\psi \cdot F)_t|^2 \mathrm{d}t \xrightarrow{\mathrm{P}} 0$, because $\int_a^b |(\psi \cdot F^n)_t - (\psi \cdot F)_t|^2 \mathrm{d}t = \psi^2 \int_a^b |F_t^n - F_t|^2 \mathrm{d}t$ and

$\int_a^b |F_t^n - F_t|^2 \mathrm{d}t \xrightarrow{\mathrm{P}} 0$ as $n \to \infty$. Therefore, $\int_a^b (\psi \cdot F^n)_t \mathrm{d}B_t \xrightarrow{\mathrm{P}} \int_a^b (\psi \cdot F)_t \mathrm{d}B_t$ as $n \to \infty$. Immediately from the definition of $\int_a^b (\psi \cdot F^n)_t \mathrm{d}B_t$, it follows that $\int_a^b (\psi \cdot F^n)_t \mathrm{d}B_t = \psi \int_a^b F_t^n \mathrm{d}B_t$. Furthermore, we have $\int_a^b F_t^n \mathrm{d}B_t \xrightarrow{\mathrm{P}} \int_a^b F_t \mathrm{d}B_t$ as $n \to \infty$. Therefore, $\psi \int_a^b F_t^n \mathrm{d}B_t \xrightarrow{\mathrm{P}} \psi \int_a^b F_t \mathrm{d}B_t$ as $n \to \infty$. But for every $n \geq 1$, we have

$$\begin{aligned}
\left| \int_a^b (\psi \cdot F)_t \mathrm{d}B_t - \psi \int_a^b F_t \mathrm{d}B_t \right| &\leq \left| \int_a^b (\psi \cdot F)_t \mathrm{d}B_t - \int_a^b (\psi \cdot F^n)_t |\mathrm{d}B_t \right| \\
&\quad + \left| \psi \int_a^b F_t^n \mathrm{d}B_t - \psi \int_a^b F_t \mathrm{d}B_t \right| \\
&= \left| \int_a^b (\psi \cdot F^n)_t \mathrm{d}B_t - \int_a^b (\psi \cdot F)_t |\mathrm{d}B_t \right| \\
&\quad + |\psi| \left| \int_a^b F_t^n \mathrm{d}B_t - \int_a^b F_t \mathrm{d}B_t \right| .
\end{aligned}$$

Therefore, $\left| \int_a^b (\psi \cdot F)_t \mathrm{d}B_t - \psi \int_a^b F_t \mathrm{d}B_t \right| = 0$ a.s., because $\int_a^b (\psi \cdot F^n)_t \mathrm{d}B_t \xrightarrow{\mathrm{P}} \int_a^b (\psi \cdot F)_t \mathrm{d}B_t$ and $\int_a^b F_t^n \mathrm{d}B_t \xrightarrow{\mathrm{P}} \int_a^b F_t \mathrm{d}B_t$ as $n \to \infty$. □

7 The Indefinite Itô Integral

Given the above filtered probability space $\mathcal{P}_{\mathbb{F}}$, by $\mathcal{L}^2_{\mathbb{F}}$ we shall denote the space of all $\mathbb{F}$-nonanticipative processes $f = (f_t)_{t \geq 0}$ such that $f \in \mathcal{L}^2_{\mathbb{F}}(0, T)$ for every $T > 0$. For $f \in \mathcal{L}^2_{\mathbb{F}}$ and a one-dimensional Brownian motion $B = (B_t)_{t \geq 0}$ on $\mathcal{P}_{\mathbb{F}}$, a stochastic process $(\int_0^t f_\tau \mathrm{d}B_\tau)_{t \geq 0}$ is called an indefinite Itô integral corresponding to the pair (f, B).

Corollary 7.1. *For the pair (f, B) given above, the indefinite Itô integral $(\int_0^t f_\tau \mathrm{d} B_\tau)_{t \geq 0}$ is $\mathbb{F}$-adapted.*

Proof. Let $T > 0$ and suppose $f \in \mathcal{S}_{\mathbb{F}}(0, T) \cap \mathcal{L}^2_{\mathbb{F}}(0, T)$. For every $t \in [0, T]$, one has $\int_0^t f_\tau \mathrm{d}B_\tau = \sum_{i=1}^{k-1} f_{t_i}(B_{t_{i+1}} - B_{t_i})$, where $t_k = t$. Hence it follows that $\int_0^t f_\tau \mathrm{d}B_\tau$ is $\mathcal{F}_t$-measurable, because $f_{t_i}(B_{t_{i+1}} - B_{t_i})$ is $\mathcal{F}_{t_k}$-measurable for $i = 1, 2, \ldots, k-1$. If $(f^n)_{n=1}^\infty$ is a sequence of $\mathcal{S}_{\mathbb{F}}(0, T) \cap \mathcal{L}^2_{\mathbb{F}}(0, T)$ such that $E \int_0^T |f_t^n - f_t|^2 \mathrm{d}t \to 0$ as $n \to \infty$, then $\int_0^t f_\tau^n \mathrm{d}B_\tau$ is $\mathcal{F}_t$-measurable for every fixed $t \in [0, T]$ and $n \geq 1$. Hence it follows that $\int_0^t f_\tau \mathrm{d}B_\tau$ is $\mathcal{F}_t$-measurable for every $0 \leq t \leq T$ and every $T > 0$, because $\int_0^t f_\tau^{n_k} \mathrm{d}B_\tau \to \int_0^t f_\tau \mathrm{d}B_\tau$ a.s. as $k \to \infty$ for every increasing subsequence $(n_k)_{k=1}^\infty$ of $(n)_{n=1}^\infty$. □

Theorem 7.1. *For every $T > 0$ and $f \in \mathcal{L}^2_{\mathbb{F}}$, there exists a continuous modification $(J_t)_{0\le t\le T}$ of $(\int_0^t f_\tau \mathrm{d}B_\tau)_{0\le t\le T}$.*

Proof. Let $(f^n)_{n=1}^\infty$ be a sequence of $\mathcal{S}_{\mathbb{F}}(0,T) \cap \mathcal{L}^2_{\mathbb{F}}(0,T)$ such that $f^n = \sum_{i=1}^{k-2} \varphi_i^n \mathbb{1}_{[t_i,t_{i+1})} + \varphi_{k-1}^n \mathbb{1}_{[t_{k-1},T]}$ and $E\int_0^T |f_t^n - f_t|^2 \mathrm{d}t \to 0$ as $n \to \infty$. Put $I_n(t) = \int_0^t f_\tau^n \mathrm{d}B_\tau$ and $I(t) = \int_0^t f_\tau \mathrm{d}B_\tau$ for $t \in [0,T]$. Immediately from the definition of $I_n(t)$, it follows that for every $0 \le s < t \le T$, one has $\int_0^t f_\tau^n \mathrm{d}B_\tau - \int_0^s f_\tau^n \mathrm{d}B_\tau = \int_s^t f_\tau^n \mathrm{d}B_\tau$ a.s. Hence continuity of $I_n = (I_n(t))_{0\le t\le T}$ for every $n = 1,2,\ldots$ follows. Furthermore, for every $0 \le s < t \le T$ and $n = 1,2,\ldots$, one has

$$
\begin{aligned}
E[I_n(t)|\mathcal{F}_s] &= E\left[\int_0^s f_\tau^n \mathrm{d}B_\tau + \int_s^t f_\tau^n \mathrm{d}B_\tau | \mathcal{F}_s\right] \\
&= \int_0^s f_\tau^n \mathrm{d}B_\tau + E\left[\sum_{s\le t_j^n < t_{j+1}^n \le t} \varphi_j^n (B_{t_{j+1}} - B_{t_j})|\mathcal{F}_s\right] \\
&= \int_0^s f_\tau^n \mathrm{d}B_\tau + \sum_j E[\varphi_j^n E[(B_{t_{j+1}} - B_{t_j})|\mathcal{F}_{t_j}]|\mathcal{F}_s] \\
&= \int_0^s f_\tau^n \mathrm{d}B_\tau = I_n(s),
\end{aligned}
$$

because $B = (B_t)_{t\ge 0}$ is an $\mathbb{F}$-martingale. Then $I_n = (I_n(t))_{0\le t\le T}$ is for every $n = 1,2,\ldots$ an $\mathbb{F}$-martingale. Thus $I_n - I_m$ is also an $\mathbb{F}$-martingale for each $n,m = 1,2,\ldots$. Therefore, by Doob's inequality, we get

$$
\begin{aligned}
P\left(\left\{\sup_{0\le t\le T} |I_n(t) - I_m(t)| > \varepsilon\right\}\right) &\le \frac{1}{\varepsilon^2} E[|I_n(T) - I_m(T)|^2] \\
&= \frac{1}{\varepsilon^2} E\int_0^T |f_t^n - f_t^m|^2 \mathrm{d}t,
\end{aligned}
$$

which by the properties of the sequence $(f^n)_{n=1}^\infty$, implies that $P(\{\sup_{0\le t\le T} |I_{n_{k+1}}(t) - I_{n_k}(t)| > 2^{-k}\}) \le 2^{-k}$ for every $k = 1,2,\ldots$ and every increasing subsequence $(n_k)_{k=1}^\infty$ of $(n)_{n=1}^\infty$. By the Borel–Cantelli lemma, we obtain

$$
P(\{\sup_{0\le t\le T} |I_{n_{k+1}}(t) - I_{n_k}(t)| > 2^{-k} \text{ for infinitely many } k\}) =: 0.
$$

Therefore, for a.e. $\omega \in \Omega$, there exists $k_1(\omega)$ such that $\sup_{0\le t\le T} |I_{n_{k+1}}(t) - I_{n_k}(t)| > 2^{-k}$ for $k \le k_1(\omega)$. Then the sequence $(I_{n_k}(t))_{k=1}^\infty$ is uniformly convergent for $t \in [0,T]$ a.s. Let $J = (J(t)_{0\le t\le T}$ be an a.s. limit of the sequence $(I_{n_k})_{k=1}^\infty$ of continuous processes $I_{n_k} = (I_{n_k}(t))_{0\le t\le T}$. It is a continuous stochastic process on $\mathcal{P}_{\mathbb{F}}$. Since $I_{n_k}(t) \to I(t)$ for every $t \in [0,T]$ as $k \to \infty$ in the $\mathbb{L}^2$-norm topology, we must have $I(t) = J(t)$ a.s. for all $t \in [0,T]$. □

Corollary 7.2. *For every* $T > 0$ *and* $f \in \mathcal{L}^2_{\mathbb{F}}$, *the process* $I = (\int_0^t f_\tau \mathrm{d}B_\tau)_{0\le t\le T}$ *is an* $\mathbb{F}$*-martingale and*

$$P\left(\left\{\sup_{0\le t\le T} |I(t)| \ge \lambda\right\}\right) \le \frac{1}{\lambda^2} E\int_0^T |f_t|^2 \mathrm{d}t \tag{7.1}$$

for every $\lambda > 0$, *where* $I(t) = \int_0^t f_\tau \mathrm{d}B_\tau$.

Proof. We can assume that I is a continuous process. For every $n = 1, 2, \dots$, let I_n be the stochastic process defined in the proof of Theorem 7.1. It is an $\mathbb{F}$-martingale. Therefore, by Doob's inequality, it follows that there exists an increasing subsequence $(n_k)_{k=1}^\infty$ of $(n)_{n=1}^\infty$ such that $I_{n_k}(t) \to I(t)$ in the $\mathbb{L}^2$-norm topology for all $t \in [0, T]$ as $k \to \infty$. Then the process $I = (I(t)_{0\le t\le T}$ is also an $\mathbb{F}$-martingale. The inequality (7.1) now follows immediately from Doob's martingale inequality. □

From the above results, it follows that for every $T > 0$ and $f \in \mathcal{L}^2_{\mathbb{F}}$, the process $I = (\int_0^t f_\tau \mathrm{d}B_\tau)_{0\le t\le T}$ is a continuous $\mathbb{F}$-martingale such that $E|I(t)|^2 < \infty$ for $0 \le t \le T$. This is not true in the general case for $f \in \mathcal{M}^2_{\mathbb{F}}(0, \infty)$. But it can be verified that in such a case, the process $\left(\int_0^t f_\tau \mathrm{d}B_\tau\right)_{0\le t\le T}$ is a local $\mathbb{F}$-martingale. It is enough to define for every $n = 1, 2, \dots$ an $\mathbb{F}$-stopping time T_n by setting $T_n = \inf\{t > 0 : \int_0^t |f_\tau|^2 \mathrm{d}\tau \ge n\} \wedge n$. Then $P(\{T_n \le n\}) = 1$, $P(\{T_n \le T_{n+1}\}) = 1$, and $P(\{\lim_{n\to\infty} T_n = \infty\}) = 1$. For every $n = 1, 2, \dots$, we have $I(t \wedge T_n) = \int_0^{t\wedge T_n} f_\tau \mathrm{d}B_\tau = \int_0^t \mathbb{1}_{\{\tau\le T_n\}} f_\tau \mathrm{d}B_\tau$ and $\int_0^\infty E[\mathbb{1}_{\{\tau\le T_n\}}|f_\tau|^2]\mathrm{d}\tau = \int_0^n E[\mathbb{1}_{\{\tau\le T_n\}}|f_\tau|^2]\mathrm{d}\tau \le n$. Then the process $\{I(t \wedge T_n) : t \ge 0\}$ is a square integrable $\mathbb{F}$-martingale for every $n = 1, 2, \dots$. It can be verified that for every $n = 1, 2, \dots$, a family $\{I(t \wedge T_n) : t \ge 0\}$ is uniformly integrable.

Let us note that the above-defined Itô integral can be defined for $\mathbb{F}$-nonanticipative matrix-valued processes with respect to vector-valued $\mathbb{F}$-Brownian motions $B = (B^1, \dots, B^m)$, where $B^1, \dots, B^m$ denote real-valued $\mathbb{F}$-Brownian motions on $\mathcal{P}_{\mathbb{F}}$ such that B^i and B^j are independent for $i \ne j$. In such a case, we consider a matrix-valued stochastic process $F = (f^{ij})_{n\times m}$ with $f^{ij} \in \mathcal{M}^2_{\mathbb{F}}(0, \infty)$ and define for every $T > 0$, a multidimensional Itô integral $\int_0^T F_t \mathrm{d}B_t$ to be an $n \times 1$ matrix of the form

$$\int_0^T F_t \mathrm{d}B_t =: \left(\sum_{j=1}^m \int_0^T f_t^{1j} \mathrm{d}B_t^j, \dots, \sum_{j=1}^m \int_0^T f_t^{nj} \mathrm{d}B_t^j\right)^*,$$

where x^* denotes the transpose of $x \in \mathbb{R}^n$. It can be verified that all properties of the Itô integral presented above can be extended to the multidimensional case.

Remark 7.1. Similarly as above, we can define stochastic integrals with respect to continuous local martingales. In particular, if M is a continuous local martingale and Φ is an $\mathbb{F}$-predictable process on a filtered probability space $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ such that $E[\int_0^T \Phi_t^2 \mathrm{d}\langle M\rangle_t] < \infty$ for every $T > 0$, then a stochastic integral of Φ with respect to M on the interval $[0, t]$ is denoted by $\int_0^t \Phi_s \mathrm{d}M_s$. It can be verified that a family $I(\Phi, M) = (\int_0^t \Phi_s \mathrm{d}M_s)_{t\geq 0}$ is a continuous local martingale on $\mathcal{P}_{\mathbb{F}}$ and that for every continuous local martingale N on $\mathcal{P}_{\mathbb{F}}$, one has $\langle I(\Phi, M), N\rangle_t = \int_0^t \Phi_s \mathrm{d}\langle M, N\rangle_s$ for every $t \geq 0$. □

8 Itô's Formula and the Martingale Representation Theorem

In the theory of stochastic processes, we have no differentiation theory in the classical sense, only an integration theory. Nevertheless, it turns out that it is possible to establish an Itô-integral version of the chain rule, called Itô's formula. It is very useful for applications and is connected with Itô processes. Let $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ with $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ satisfying the usual conditions, and let $B = (B^1, \ldots, B^m)$ be an m-dimensional $\mathbb{F}$-Brownian motion on $\mathcal{P}_{\mathbb{F}}$. Assume that $F = (f^1, \ldots, f^n)^*$ and $G = (g^{ij})_{n\times m}$ are $\mathbb{F}$-nonanticipative processes with f^i and g^{ij} such that $P(\{\int_0^\infty |f_t^i|\mathrm{d}t < \infty\}) = 1$ and $P(\{\int_0^\infty |g_t^{ij}|^2\mathrm{d}t < \infty\}) = 1$ for $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, m$.

An n-dimensional stochastic process $X = (X_t)_{t\geq 0}$ on $\mathcal{P}_{\mathbb{F}}$ defined by $X_t = X_0 + \int_0^t F_\tau \mathrm{d}\tau + \int_0^t G_\tau \mathrm{d}B_\tau$ a.s. for $t \geq 0$ is said to be an n-dimensional Itô process starting at X_0 with stochastic differential $\mathrm{d}X$ on $[0, \infty)$ denoted by $\mathrm{d}X_t = F_t\mathrm{d}t + G_t\mathrm{d}B_t$ for $t \geq 0$. We have the following theorem, known as Itô's lemma.

Theorem 8.1. *Let $X = (X_t)_{t\geq 0}$ be an n-dimensional Itô process on $\mathcal{P}_{\mathbb{F}}$ having a stochastic differential $\mathrm{d}X_t = F_t\mathrm{d}t + G_t\mathrm{d}B_t$ for $t \geq 0$ with $F = (f^1, \ldots, f^n)^*$ and $G = (g^{ij})_{n\times m}$ such as above. Assume that $g : [0, \infty) \times \mathbb{R}^n \to \mathbb{R}^p$ is a $C^{1,2}$-map. Then the process $Y = (Y_t)_{t\geq 0}$ defined by $Y_t = g(t, X_t)$ for $t \geq 0$ is a p-dimensional Itô process having a stochastic differential $\mathrm{d}Y = (\mathrm{d}Y_t)_{t\geq 0}$ with $\mathrm{d}Y_t = (\mathrm{d}Y_t^1, \ldots, \mathrm{d}Y_t^p)$ and $\mathrm{d}Y_t^k$ defined by*

$$\mathrm{d}Y_t^k = \frac{\partial g_k}{\partial t}(t, X_t)\mathrm{d}t + \sum_{i=1}^n \frac{\partial g_k}{\partial x_i}(t, X_t)\mathrm{d}X_t^i + \frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n \frac{\partial^2 g_k}{\partial x_i \partial x_j}(t, X_t)\mathrm{d}X_t^i\mathrm{d}X_t^j,$$

for $k = 1, 2, \ldots, p$, where $\mathrm{d}B_t^i\mathrm{d}B_t^j = \delta_{ij}\mathrm{d}t$ and $\mathrm{d}B_t^i\mathrm{d}t = \mathrm{d}t\mathrm{d}B_t^i = 0$ for $i, j = 1, 2, \ldots, m$. □

Example 8.1. Let $r, \alpha \in \mathbb{R}$ and let $X = (X_t)_{t\geq 0}$ be a stochastic process on $\mathcal{P}_{\mathbb{F}}$ such that $\mathrm{d}X_t = r X_t\mathrm{d}t + \alpha X_t\mathrm{d}B_t$ for $t \geq 0$, where $B = (B_t)_{t\geq 0}$ is a given $\mathbb{F}$-Brownian motion on $\mathcal{P}_{\mathbb{F}}$. Using Itô formula, we can determine the process X. To do this, let us rewrite the above equation in the form $\mathrm{d}X_t/X_t = r\mathrm{d}t + \alpha\mathrm{d}B_t$. Taking $g(t, x) = \ln(x)$ for $x > 0$, immediately from Itô's formula we obtain

$$\mathrm{d}(\ln(X_t)) = \frac{1}{X_t}\cdot\mathrm{d}X_t + \frac{1}{2}\left(-\frac{1}{X_t^2}\right)(\mathrm{d}X_t)^2 = \frac{\mathrm{d}X_t}{X_t} - \frac{1}{2X_t^2}\alpha^2X_t^2\mathrm{d}t = \frac{\mathrm{d}X_t}{X_t} - \frac{1}{2}\alpha^2\mathrm{d}t.$$

Therefore,

$$rt + \alpha B_t = \int_0^t \frac{\mathrm{d}X_t}{X_t} = \int_0^t \mathrm{d}(\ln(X_s)) + \frac{1}{2}\alpha^2 t\,.$$

Assuming that $X_0 \neq 0$ a.s., we get

$$\ln\left(\frac{X_t}{X_0}\right) = \left(r - \frac{1}{2}\alpha^2\right) t + \alpha B_t$$

a.s. for $t \geq 0$. Then

$$X_t = X_0 \exp\left[\left(r - \frac{1}{2}\alpha^2\right) t + \alpha B_t\right] \quad \text{a.s.}$$

Thus $X = (X_t)_{t\geq 0}$ is defined by $X_t = X_0 \exp(\mu t + \alpha B_t)$ a.s. for $t \geq 0$ with $\mu = (r - \frac{1}{2}\alpha^2)$.

A process $X = (X_t)_{t\geq 0}$ of the form $X_t = X_0 \exp(\mu t + \alpha B_t)$ with $\alpha, \mu \in \mathbb{R}$ is called a geometric Brownian motion. Such processes are important as models for stochastic prices in mathematical economics.

Remark 8.1. As an application of Itô's formula, it follows that for every $p \geq 2$, there exist positive constants $K_1 = K_1(p)$ and $K_2 = K_2(p)$ such that

$$K_1 E[\langle M\rangle_t^{p/2}] \leq E\left[\sup_{0\leq t\leq T} |M_t|^p\right] \leq K_2 E[\langle M\rangle_T^{p/2}$$

for $0 \leq t \leq T$, every $T > 0$ and every continuous local martingale M such that $E[|M_T|^p] < \infty$. □

Immediately from the properties of stochastic processes defined by indefinite Itô integrals, it follows that for a given matrix-valued process $G = (g^{ij})_{n\times m}$ with $g^{ij} \in \mathcal{M}^2_{\mathbb{F}}(0,\infty)$ and an m-dimensional $\mathbb{F}$-Brownian motion $B = (B_t)_{t\geq 0}$, the process $X = (X_t)_{t\geq 1}$ with $X_t = X_0 + \int_0^t G_\tau \mathrm{d}B_\tau$ for $t \geq 0$ is a continuous n-dimensional local $\mathbb{F}$-martingale. It can be proved that for local martingales of certain types, the converse is also true. We precede the presentation of such a theorem by notions dealing with extensions of filtered probability spaces. Given a filtered probability space $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$, we will say that a filtered probability space $\tilde{\mathcal{P}}_{\tilde{\mathbb{F}}} = (\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{F}}, \tilde{P})$ with $\tilde{\mathbb{F}} = (\tilde{\mathcal{F}}_t)_{t\geq 0}$ is an extension of $\mathcal{P}_{\mathbb{F}}$ if there exists an $(\tilde{\mathcal{F}}, \mathcal{F})$-measurable mapping $\pi : \tilde{\Omega} \to \Omega$ such that $\pi^{-1}(\mathcal{F}_t) \subset \tilde{\mathcal{F}}_t$ for $t \geq 0$, $P = \tilde{P} \circ \pi^{-1}$, and for every $Z \in \mathbb{L}^\infty(\Omega, \mathcal{F}, P, \mathbb{R}^d)$, an $\mathbb{R}^d$-random variable $\tilde{Z}$ on $\tilde{\mathcal{P}}_{\tilde{\mathbb{F}}}$ defined by setting $\tilde{Z}(\tilde{\omega}) = Z(\pi(\tilde{\omega}))$ for $\tilde{\omega} \in \tilde{\Omega}$ satisfies $\tilde{E}[\tilde{Z}|\tilde{\mathcal{F}}_t](\tilde{\omega}) = E[Z||\mathcal{F}_t](\pi(\tilde{\omega}))$ for every $\tilde{\omega} \in \tilde{\Omega}$. There is a more general extension, called the standard extension, of a probability space $\mathcal{P}_{\mathbb{F}}$. It is connected with the following problem: given an $\mathbb{F}$-adapted stochastic process $X = (X_t)_{t\geq 0}$ on a filtered probability space $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$, we may need an m-dimensional $\mathbb{F}$-Brownian motion independent of X. But because $\mathcal{P}_{\mathbb{F}}$ may not be rich enough to support such a Brownian motion, we must extend the probability space in order to construct this. To do this, suppose $(\Omega', \mathcal{F}', P')$ is a another probability space on which we have given an m-dimensional Brownian motion $B' = (B')_{t\geq 0}$, and let

$$\tilde{\Omega} = \Omega \times \Omega', \ \ \tilde{\mathcal{F}} = \mathcal{F} \otimes \mathcal{F}', \ \ \tilde{P} = P \times P', \ \ \pi\tilde{\omega} = \omega \ \text{ for } \ \tilde{\omega} = (\omega, \omega') \in \tilde{\Omega}\,.$$

If $\tilde{\mathbb{F}} = (\tilde{\mathcal{F}}_t)_{t\geq 0}$ is a filtration on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ such that $\mathcal{F}_t \otimes \mathcal{F}' \supset \tilde{\mathcal{F}}_t \supset \mathcal{F}_t \otimes \{\Omega', \emptyset\}$, then $\tilde{\mathcal{P}}_{\tilde{\mathbb{F}}} = (\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{F}}, \tilde{P})$ with $\tilde{\mathbb{F}} = (\tilde{\mathcal{F}}_t)_{t\geq 0}$ is called a standard extension of the filtered probability space $\mathcal{P}_{\mathbb{F}}$. It can be verified that a standard extension of a filtered probability space $\mathcal{P}_{\mathbb{F}}$ is an extension of this space. Let us observe that the filtration $\tilde{\mathbb{F}}$ defined above may not satisfy the usual conditions, so we augment it and make it right continuous by defining $\tilde{\mathcal{F}}_t = \bigcap_{s>t} \sigma(\tilde{\mathcal{F}}_s \cup \mathcal{N})$, where $\mathcal{N}$ is the collection of all $\tilde{\mathcal{P}}$-null sets in $\tilde{\mathcal{F}}$. We also complete $\tilde{\mathcal{F}}$ by defining $\tilde{\mathcal{F}} = \sigma(\tilde{\mathcal{F}} \cup \mathcal{N})$. We may extend X and B to $\tilde{\mathbb{F}}$-adapted processes on $\tilde{\mathcal{P}}$ by defining $\tilde{X}_t(\tilde{\omega}) = X_t(\omega)$ and $\tilde{B}_t(\tilde{\omega}) = B_t(\omega')$ for $\tilde{\omega} = (\omega, \omega') \in \tilde{\Omega}$. Then $\tilde{B} = (\tilde{B}_t)_{t\geq 0}$ is an m-dimensional Brownian motion, independent of $\tilde{X} = (\tilde{X}_t)_{t\geq 0}$.

Remark 8.2. If $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ is separable, then there is a separable standard extension of this space.

Proof. Let us take in the above definition of the standard extension of $\mathcal{P}_{\mathbb{F}}$ a probability space $(\Omega', \mathcal{F}', P')$ that is separable and denote by $(A_n)_{n=1}^{\infty}$ and $(A'_m)_{m=1}^{\infty}$ sequences dense in $\mathcal{F}$ and $\mathcal{F}'$, respectively. Let $A \in \mathcal{F}$, $A' \in \mathcal{F}'$ and denote by $(A_{n_k})_{k=1}^{\infty}$ and $(A'_{m_k})_{k=1}^{\infty}$ subsequences of $(A_n)_{n=1}^{\infty}$ and $(A'_m)_{m=1}^{\infty}$, respectively such that $P(A \triangle A_{n_k}) \to 0$ and $P'(A' \triangle A'_{m_k}) \to 0$ as $k \to \infty$. We obtain $A_{n_k} \times A'_{m_k} \in \mathcal{F} \times \mathcal{F}'$ and $\tilde{P}[(A \times A') \triangle (A_n \times A'_{m_k})] = (P \times P')[(A \times A') \triangle (A_{n_k} \times A'_{m_k})] \leq 3P(A \triangle A_{n_k}) \cdot P'(A' \triangle A'_{m_k})$. Therefore, $\tilde{P}[(A \times A') \triangle (A_{n_k} \times A'_{m_k})] \to 0$ as $k \to \infty$. Hence it follows that for every $\tilde{A} \in \tilde{\mathcal{F}}$, there is a subsequence $(C_k)_{k=1}^{\infty}$ of $(A_n \times A'_m)_{n,m=1}^{\infty}$ such that $\tilde{P}(\tilde{A} \triangle C_k) \to 0$ as $k \to \infty$. □

Now we can formulate the following representation theorem.

Theorem 8.2. *Suppose $M = (M^1, \ldots, M^d)$, with $M^i = (M^i_t)_{t\geq 0}$ for $i = 1, 2, \ldots, d$, is a d-dimensional continuous local $\mathbb{F}$-martingale on $\mathcal{P}_{\mathbb{F}}$ such that for every $i, j = 1, 2, \ldots, d$, the function $\mathbb{R}^+ \ni t \to \langle M^i, M^j \rangle_t (\omega) \in \mathbb{R}$ is absolutely continuous for a.e. $\omega \in \Omega$. Then there are a standard extension $\tilde{\mathcal{P}}_{\tilde{\mathbb{F}}} = (\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{F}}, \tilde{P})$ of $\mathcal{P}_{\mathbb{F}}$, a d-dimensional $\tilde{\mathbb{F}}$-Brownian motion $\tilde{B} = (\tilde{B}_t)_{t\geq 0}$ on $\tilde{\mathcal{P}}_{\tilde{\mathbb{F}}}$, and a matrix-valued process $\rho = (\rho^{ij})_{d\times d}$ with $\rho^{ij} \in \mathcal{M}^2_{\tilde{\mathbb{F}}}(0, \infty)$ for $i, j = 1, 2, \ldots, d$ such that:*

(i) $M_t = \int_0^t \rho_\tau \mathrm{d}\tilde{B}_\tau$ *for* $t \geq 0$;
(ii) $\langle M^i, M^j \rangle_t = \sum_{k=1}^n \int_0^t \rho^{ik}_\tau \rho^{jk}_\tau \mathrm{d}\tau$ *a.s. for* $t \geq 0$ *and* $i, j = 1, 2, \ldots, d$. □

9 Stochastic Differential Equations and Diffusions

There are several approaches to the study of diffusions, running from the purely analytic to the purely probabilistic. We present the stochastic differential equations approach. It was suggested by P. Lèvy and was carried out in a masterly way by K. Itô. The stochastic differential equations approach to diffusion processes provides a powerful methodology and useful representation for a very large class of such processes.

Given Borel-measurable vector- and matrix-valued mappings $f : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^d$ and $g : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^{d\times m}$, by a stochastic differential equation $SDE(f, g)$ we mean a relation

$$X_t = X_0 + \int_0^t f(\tau, X_\tau)\mathrm{d}\tau + \int_0^t g(\tau, X_\tau)\mathrm{d}B_\tau , \tag{9.1}$$

written usually in the differential form

$$\mathrm{d}X_t = f(t, X_t)\mathrm{d}t + g(t, X_t)\mathrm{d}B_t , \tag{9.2}$$

which has to be satisfied a.s. for every $t \geq 0$ by a system $(\mathcal{P}_{\mathbb{F}}, X, B)$ consisting of a complete filtered probability space $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ satisfying the usual conditions, a d-dimensional $\mathbb{F}$-adapted continuous stochastic process $X = (X_t)_{t\geq 0}$, and an m-dimensional $\mathbb{F}$-Brownian motion $B = (B_t)_{t\geq 0}$ on $\mathcal{P}_{\mathbb{F}}$ such that

$$P\left[\int_0^t \{|f^i(\tau, X_\tau)| + |g^{ij}(\tau, X_\tau)|^2\}\mathrm{d}\tau < \infty\right] = 1$$

holds for every $1 \leq i \leq d$, $1 \leq j \leq m$ and $t \geq 0$. Such system is said to be a weak solution of the stochastic differential equation (9.1). A weak solution $(\mathcal{P}_{\mathbb{F}}, X, B)$ of (9.1) is said to be unique in law if for every other weak solution $(\tilde{\mathcal{P}}_{\tilde{\mathbb{F}}}, \tilde{X}, \tilde{B})$ of (9.1), one has $PX^{-1} = P\tilde{X}^{-1}$.

Corollary 9.1. *If* $(\mathcal{P}_{\mathbb{F}}, X, B)$ *is a weak solution of* (9.1)*, then* $P(\{\int_0^t |f(s, X_s)| \mathrm{d}s + \int_0^t \|g(s, X_s)\|^2 \mathrm{d}s < \infty\}) = 1$ *for every* $t \geq 0$*, where* $|\cdot|$ *and* $\|\cdot\|$ *denote norms of* $\mathbb{R}^d$ *and* $\mathbb{R}^{d\times m}$*, respectively.* □

Given a probability measure $\mu : \beta(\mathbb{R}^d) \to [0, 1]$, a weak solution $(\mathcal{P}_{\mathbb{F}}, X, B)$ of (9.1) such that $PX_0^{-1} = \mu$ is called a weak solution of (9.1) (or equivalently of (9.2)) with an initial distribution μ.

Remark 9.1. For a given $(s, x) \in \mathbb{R}^+ \times \mathbb{R}^d$, we can also define a weak solution $(\mathcal{P}_{\mathbb{F}}, X, B)$ of (9.1) with the initial condition $X_s = x$ a.s. □

If apart from the above mappings f and g, we are also given a complete filtered probability space $\mathcal{P}_{\mathbb{F}}$ and an m-dimensional $\mathbb{F}$-Brownian motion $B = (B_t)_{t\geq 0}$ on

$\mathcal{P}_{\mathbb{F}}$, then we can look for a d-dimensional $\mathbb{F}$-adapted continuous stochastic process $X = (X_t)_{t\geq 0}$ such that a system $(\mathcal{P}_{\mathbb{F}}, X, B)$ satisfies (9.2) a.s. for every $t \geq 0$. Such a process X is called a strong solution of (9.2). It is clear that every strong solution of (9.2) is also a weak solution; more precisely, a strong solution determines a weak one. There are, however, stochastic differential equations having weak solutions that do not admit strong ones. One such example is Tanaka's equation, of the form $\mathrm{d}X_t = \mathrm{sgn}(X_t)\mathrm{d}B_t$ for $t \geq 0$.

In what follows, a weak solution $(\mathcal{P}_{\mathbb{F}}, X, B)$ of (9.2) can be identified with a pair (X, B) of stochastic processes defined on $\mathcal{P}_{\mathbb{F}}$. Many properties of weak solutions of (9.2) are represented by properties of the process X. Therefore, a weak solution $(\mathcal{P}_{\mathbb{F}}, X, B)$ of (9.2) is often identified with the process X.

Let $f = (f^1, \ldots, f^d)^*$ and $g = (g^{ij})_{d\times m}$ be as above. Associate to the pair (f, g) a linear operator $\mathbb{L}_{fg}$ defined on the space $C_0^2(\mathbb{R}^d)$ of all continuous functions $h : \mathbb{R}^d \to \mathbb{R}$ with compact support and having continuous and bounded derivatives h'_{x_i} and $h''_{x_i x_j}$ for $i, j = 1, 2, \ldots, d$, by setting

$$(\mathbb{L}_{fg}h)(t,x) = \sum_{i=1}^{d} f^i(t,x)h'_{x_i}(x) + \frac{1}{2}\sum_{i,j=1}^{d} \sigma^{ij}(t,x)h''_{x_i x_j}(x) \tag{9.3}$$

for $h \in C_0^2(\mathbb{R}^d)$, $t \geq 0$, and $x \in \mathbb{R}^d$, where $(\sigma^{ij})_{d\times d} = g \cdot g^*$. We shall prove the following theorem.

Theorem 9.1. *Let $f : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^d$ and $g : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^{d\times m}$ be Borel-measurable mappings and let μ be a probability measure on $\beta(\mathbb{R}^d)$. The stochastic differential equation (9.2) possesses at least one weak solution with an initial distribution μ if and only if there exist a filtered probability space $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ satisfying the usual conditions and a d-dimensional continuous $\mathbb{F}$-adapted process $X = (X_t)_{t\geq 0}$ on $\mathcal{P}_{\mathbb{F}}$ with $PX_0^{-1} = \mu$ and such that for every $h \in C_0^2(\mathbb{R}^d)$, the process $\varphi^h = (\varphi_t^h)_{t\geq 0}$ defined by $\varphi_t^h = h(X_t) - h(X_0) - \int_0^t (\mathbb{L}_{fg}h)(s, X_s)\mathrm{d}s$ a.s. for $t \geq 0$ is a continuous local $\mathbb{F}$-martingale on $\mathcal{P}_{\mathbb{F}}$.*

Proof. Let $(\mathcal{P}_{\mathbb{F}}, X, B)$ be a weak solution of (9.2) such that $PX_0^{-1} = \mu$. By Itô's formula, for every $h \in C_0^2(\mathbb{R}^d)$, one gets

$$h(X_t) - h(X_0) = \int_0^t (\mathbb{L}_{fg}h)(s, X_s)\mathrm{d}s + \sum_{i=1}^{d}\sum_{j=1}^{m}\int_0^t h'_{x_i}(X_s) \cdot g^{ij}(s, X_s)\mathrm{d}B_s^j \tag{9.4}$$

a.s. for $t \geq 0$. Therefore, for every $h \in C_0^2(\mathbb{R}^d)$, we have $\varphi_t^h = \sum_{i=1}^d \sum_{j=1}^m \int_0^t h'_{x_i}(X_s) \cdot g^{ij}(s, X_s)\mathrm{d}B_s^j$ a.s. for $t \geq 0$, where $\mathbb{L}_{fg}$ is defined by (9.3). Hence, by the properties of X and Itô integrals, it follows that a process $\varphi^h = (\varphi_t^h)_{t\geq 0}$ is a continuous local $\mathbb{F}$-martingale. Indeed, let $T_n = \inf\{t \geq 0 : |X_t| \geq n \text{ or } \int_0^t \|g(s, X_s)\|^2 \mathrm{d}s \geq n\}$. We have $T_n \leq T_{n+1}$ a.s., and by Corollary 9.1, $\lim_{n\to\infty} T_n = \infty$. Immediately from (9.4) and the properties of Itô integrals, we obtain that $(\varphi^h_{t\wedge T_n})_{t\geq 0}$ is for every $n = 1, 2, \ldots$ a continuous square integrable $\mathbb{F}$-martingale. Then $(\varphi_t^h)_{t\geq 0}$ is a continuous local $\mathbb{F}$-martingale.

Assume that there exist $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ satisfying the usual conditions and a d-dimensional continuous $\mathbb{F}$-adapted process $X = (X_t)_{t\geq 0}$ on $\mathcal{P}_{\mathbb{F}}$ with $PX_0^{-1} = \mu$ and such that for every $h \in C_0^2(\mathbb{R}^d)$, the process $\varphi^h = (\varphi_t^h)_{t\geq 0}$ defined above is a continuous local $\mathbb{F}$-martingale on $\mathcal{P}_{\mathbb{F}}$. In particular, the process $\varphi_l^{h_i} = (\varphi^{h_i}_{t\wedge T_l})_{t\geq 0}$ defined by

$$\varphi^{h_i}_{t\wedge T_l} = X^i_{t\wedge T_l} - X^i_0 - \int_0^{t\wedge T_l} f^i(s, X_s)\mathrm{d}s\,,$$

for $h_i \in C_0^2(\mathbb{R}^d)$ such that $h_i(x) = x_i$ for $x \in K_l := \{x \in \mathbb{R}^d : |x| \leq l\}$ for $l = 1, 2, \ldots$ with $T_l = \inf\{t > 0 : X_t \notin K_l\}$ for $l = 1, 2, \ldots$ is a continuous local $\mathbb{F}$-martingale on $\mathcal{P}_{\mathbb{F}}$. Therefore, for every $l = 1, 2, \ldots$, there is an increasing sequence $(S_n^l)_{n=1}^\infty$ of finite $\mathbb{F}$-stopping times S_n^l such that $\lim_{n\to\infty} S_n^l = \infty$ a.s. Taking $\tau_l = T_l \wedge S_l^l$ for $l = 1, 2, \ldots$, we obtain that $(\varphi^{h_i}_{t\wedge\tau_l})_{t\geq 0}$ is an $\mathbb{F}$-martingale for $l = 1, 2, \ldots$. Then a process $(\varphi_t^{h_i})_{t\geq 0}$ is a continuous local martingale and

$$\varphi_t^{h_i} = X_t^i - X_0^i - \int_0^t f^i(s, X_s)\mathrm{d}s$$

for $i = 1, 2, \ldots, d$ and $t \geq 0$.

Similarly, for every $h_{ij} \in C_0^2(\mathbb{R}^d)$ such that $h_{ij}(x) = x_i x_j$ for $x \in K_l$, the process $\varphi^{h_{ij}} = (\varphi_t^{h_{ij}})_{t\geq}$ is a continuous local $\mathbb{F}$-martingale and

$$\varphi_t^{h_{ij}} = X_t^i X_t^j - X_0^i X_0^j - \int_0^t [X_s^i f^j(s, X_s) + X_s^j f^i(s, X_s) + \sigma^{ij}(s, X_s)]\mathrm{d}s$$

for $i, j = 1, 2, \ldots, d$ and $t \geq 0$. Hence it follows that

$$\varphi_t^{h_i} \cdot \varphi_t^{h_j} - \int_0^t \sigma^{ij}(s, X_s)\mathrm{d}s = \varphi_t^{h_{ij}} - X_0^i \varphi_t^{h_j} - X_0^j \varphi_t^{h_i} + M_t^{ij} \tag{9.5}$$

a.s., where

$$M_t^{ij} = \int_0^t (X_s^i - X_t^i) f^j(s, X_s)\mathrm{d}s + \int_0^t (X_s^j - X_t^j) f^i(s, X_s)\mathrm{d}s + \left(\int_0^t f^i(s, X_s)\mathrm{d}s\right) \cdot \left(\int_0^t f^j(s, X_s)\mathrm{d}s\right)$$

for $i, j = 1, 2, \ldots, d$ and $t \geq 0$, which by Itô's formula, can be written in the form

$$M_t^{ij} = \int_0^t (\varphi_s^{h_i} - \varphi_t^{h_i}) f^j(s, X_s)\mathrm{d}s + \int_0^t (\varphi_s^{h_j} - \varphi_t^{h_j}) f^i(s, X_s)\mathrm{d}s$$
$$= -\int_0^t \left(\int_0^s f^j(u, X_u)\mathrm{d}u\right)\mathrm{d}\varphi_s^{h_i} - \int_0^t \left(\int_0^s f^i(u, X_u)\mathrm{d}u\right)\mathrm{d}\varphi_s^{h_j}.$$

But X is $\mathbb{F}$-adapted. Then by virtue of Remark 7.1, $(M_t^{ij})_{t\geq 0}$ is a continuous local $\mathbb{F}$-martingale. It is clear that $(\varphi_t^{h_{ij}} - X_0^i \varphi_t^{h_j} - X_0^j \varphi_t^{h_i})_{t\geq 0}$ is a continuous local $\mathbb{F}$-martingale, too. Therefore, immediately from (9.5), it follows that

$$\left\langle \varphi^{h_i}, \varphi^{h_j} \right\rangle_t = \int_0^t \sigma^{ij}(s, X_s)\mathrm{d}s \quad \text{a.s. for } t \geq 0.$$

By virtue of Theorem 8.2, there exist a standard extension $\tilde{\mathcal{P}}_{\tilde{\mathbb{F}}} = (\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{F}}, \tilde{P})$ of $\mathcal{P}_{\mathbb{F}}$, a d-dimensional $\tilde{\mathbb{F}}$-Brownian motion $\tilde{B} = (\tilde{B}_t)_{t\geq 0}$ on $\tilde{\mathcal{P}}_{\tilde{\mathbb{F}}}$, and a matrix-valued process $\rho = (\rho^{ij})_{d\times d}$ with $\rho^{ij} \in \mathcal{M}_{\tilde{\mathbb{F}}}^2(0, \infty)$ satisfying $P[\int_0^t (\rho^{ij})_\tau^2 \mathrm{d}\tau < \infty] = 1$ for $i, j = 1, 2, \ldots, d$ and such that

$$\tilde{\varphi}_t^{h_i} = \sum_{j=1}^d \int_0^t \rho_s^{ij} \mathrm{d}\tilde{B}_s^j \quad \tilde{P} - \text{a.s. for } t \geq 0,$$

with $\tilde{\varphi}_t^{h_i}(\tilde{\omega}) = \varphi_t^{h_i}(\pi(\tilde{\omega}))$ for $\tilde{\omega} \in \tilde{\Omega}$, where $\pi : \tilde{\Omega} \to \Omega$ is the $(\tilde{\mathcal{F}}, \mathcal{F})$-measurable mapping described in the definition of the extension of $\mathcal{P}_{\mathbb{F}}$ because its standard extension $\tilde{\mathcal{P}}_{\tilde{\mathbb{F}}}$ is also its extension. But

$$\tilde{\varphi}_t^{h_i} = \tilde{X}_t^i - \tilde{X}_0^i - \int_0^t f^i(s, \tilde{X}_s)\mathrm{d}s \quad \tilde{P} - \text{a.s. for } t \geq 0 \text{ and } i = 1, 2, \ldots, d,$$

where $\tilde{X}_t^i(\tilde{\omega}) = X_t^i(\pi(\tilde{\omega}))$ for $\tilde{\omega} \in \tilde{\Omega}$. Therefore,

$$\tilde{X}_t = \tilde{X}_0 + \int_0^t f(s, \tilde{X}_s)\mathrm{d}s + \int_0^t \rho_s \mathrm{d}\tilde{B}_s \quad \tilde{P} - \text{a.s. for } t \geq 0. \tag{9.6}$$

Furthermore, for every $A \in \beta(\mathbb{R}^d)$, we have $(P\tilde{X}_0^{-1})(A) = \tilde{P}[\tilde{X}_0^{-1}(A)] = \tilde{P}[(X_0 \circ \pi)^{-1}(A)] = (\tilde{P} \circ \pi^{-1})[X_0^{-1}(A)] = P[X_0^{-1}(A)] = (PX_0^{-1})(A) = \mu$. Hence it follows that if we are able to establish the existence on $\tilde{\mathcal{P}}_{\tilde{\mathbb{F}}}$ of an m-dimensional $\tilde{\mathbb{F}}$-Brownian motion $\hat{B} = (\hat{B}_t)_{t\geq 0}$ such that $\int_0^t \rho_s \mathrm{d}\tilde{B}_s = \int_0^t g(s, \tilde{X}_s)\mathrm{d}\hat{B}_s$ a.s. for $t \geq 0$, then the system $(\tilde{\mathcal{P}}_{\tilde{\mathbb{F}}}, \tilde{X}, \hat{B})$ will be a weak solution of (9.2) with an initial distribution μ. Indeed, by (9.6), it follows that (9.2) will be satisfied. Furthermore, by the identities

$$\left\langle \varphi^{h_i}, \varphi^{h_j} \right\rangle_t = \int_0^t \sigma^{ij}(s, X_s)\mathrm{d}s, \qquad \tilde{\varphi}_t^{h_i} = \sum_{j=1}^d \int_0^t \rho_s^{ij} \mathrm{d}\tilde{B}_s^j$$

and the definition of σ , it follows that

$$\tilde{P}\left[\sum_{j=1}^{d}\rho_t^{ij}\rho_t^{kj}=\sigma^{ik}(t,X_t)\ \text{ for a.e. }\ t\geq 0\right]=1$$

for $i\leq i,k\leq d$, which implies that

$$\tilde{P}\left[\int_0^t (g^{ij})^2(\tau,\tilde{X}_\tau)\mathrm{d}\tau<\infty\right]=1\ \text{ for }\ 1\leq i\leq d,\ 1\leq j\leq m\ \text{ and }\ t\geq 0\,.$$

Let us observe first that it suffices to construct a process $\hat{B}$ with $m=d$. Indeed, if $d<m$, we may augment $\tilde{X}$, $f\circ\tilde{X}$ and $g\circ\tilde{X}$ by setting $\tilde{X^i}=f^i(t,\tilde{X}_t)=g^{ij}(t,\tilde{X}_t)=0$ for $d+1\leq i\leq m$ and $1\leq j\leq m$. This d-dimensional process $\tilde{X}$ satisfies the conditions presented in the second part of the proof, and we may proceed as before except now we shall obtain a matrix ρ, which, like $g\circ\tilde{X}$, will be of dimension $d\times d$. On the other hand, if $m<d$, we have only to augment $g\circ\tilde{X}$ by setting $g^{ij}(t,\tilde{X}_t)=0$ for $1\leq i\leq d$ and $m+1\leq j\leq d$, and nothing else is affected. Both ρ and $g\circ\tilde{X}$ are then $d\times d$ matrices.

By diagonalization, $\rho\cdot\rho^*=\tilde{g}\cdot\tilde{g}^*$, where $\tilde{g}=g\circ\tilde{X}$, and studying the effect of the diagonalization transformation on ρ and $\tilde{g}$, we can show that there exists a Borel-measurable $d\times d$-matrix-valued function $\mathcal{R}(\rho,\tilde{g})$ defined on the set

$$\mathrm{D}=\{(\rho,\tilde{g}):\rho\ \text{ and }\ \tilde{g}\ \text{ are }\ d\times d-\text{matrices with }\rho\cdot\rho^*=\tilde{g}\cdot\tilde{g}^*\}$$

such that $\tilde{g}=\rho\mathcal{R}(\rho,\tilde{g})$ and $\mathcal{R}(\rho,\tilde{g})\cdot\mathcal{R}^*(\rho,\tilde{g})=I$, the identity $d\times d$ matrix. We set $\hat{B}_t=\int_0^t\mathcal{R}^*(\rho_s,g(s,\tilde{X}_s))\mathrm{d}\tilde{B}_s$ for $t\geq 0$. It is easy to see that $\hat{B}=(\hat{B}_t)_{t\geq 0}$ is a continuous local $\tilde{\mathbb{F}}$-martingale such that $\left\langle\hat{B}^i,\hat{B}^j\right\rangle_t=t\delta_{ij}$ for $i,j=1,2,\ldots,d$ and $t\geq 0$. By Theorem 5.8, $\hat{B}$ is an $\tilde{\mathbb{F}}$-Brownian motion such that $\mathrm{d}\hat{B}_t=\mathcal{R}^T(\rho_t,g(t,\tilde{X}_t))\mathrm{d}\tilde{B}_t$ for $t\geq 0$. Hence it follows that $\rho_t\mathcal{R}(\rho,g(t,\tilde{X}_t))\mathrm{d}\hat{B}_t=\rho_t\mathcal{R}(\rho,g(t,\tilde{X}_t))\cdot\mathcal{R}^T(\rho,g(t,\tilde{X}_t))\mathrm{d}\tilde{B}_t=\rho_t\mathrm{d}\tilde{B}_t$ for $t\geq 0$, which is equivalent to $g(t,\tilde{X}_t)\mathrm{d}\hat{B}_t=\rho_t\mathrm{d}\tilde{B}_t$. □

Corollary 9.2. *If a pair* $(\mathcal{P}_{\mathbb{F}},X)$ *satisfies the conditions of Theorem 9.1, then the distribution* PX^{-1} *is a probability measure on* $(C,\beta(C))$ *with a filtration* $\beta_t(C)=\sigma(\bigcup_{w\in C}\{w(s):0\leq s\leq t\})$ *such that the process* $(h(w(t))-h(w(0))-\int_0^t(\mathbb{L}_{fg}h)(s,w(s))\mathrm{d}s)_{t\geq 0}$ *is a continuous local* $(\beta_t(C))_{t\geq 0}$*-martingale on* $(C,\beta(C))$ *for every* $h\in C_0^2(\mathbb{R}^d))$*, where* $C=C(\mathbb{R}^+,\mathbb{R}^d)$. □

The properties of the distribution PX^{-1} described above can be equivalently expressed by saying that a measure $P_X=:PX^{-1}$ is a solution of the local martingale problem for the differential operator $\mathbb{L}_{fg}$. Such a problem is also called a local $\mathbb{L}_{fg}$-martingale problem. More generally, for a given differential operator $\mathbb{L}_{fg}$ and $(s,x)\in[0,\infty)\times\mathbb{R}^d$, a solution of the local martingale problem for $\mathbb{L}_{fg}$ (or for a pair (f,g)) starting from (s,x) is a probability measure $\mathbb{P}_{s,x}$ on $(C,\beta(C))$

satisfying $\mathbb{P}_{s,x}(\{x(t) = x \text{ for } 0 \le t \le s\}) = 1$ and such that the process $(h(x(t)) - h(x) - \int_s^t (\mathbb{L}_{fg}h)(\tau, x(\tau))\mathrm{d}\tau)_{t\ge s}$ is a continuous $(\mathbb{P}_{s,x}, (\beta_t(C))_{t\ge 0})$-local martingale for every $h \in C_0^2(\mathbb{R}^d)$. The local $\mathbb{L}_{fg}$-martingale problem is said to be well posed if for each (s,x), there is exactly one solution of the martingale problem starting from (s,x).

Corollary 9.3. *Given bounded measurable functions $f : \mathbb{R}^+\times\mathbb{R}^d \to \mathbb{R}^d$ and $g : \mathbb{R}^+\times\mathbb{R}^d \to \mathbb{R}^{d\times m}$, and $(s,x) \in \mathbb{R}^+\times\mathbb{R}^d$, the stochastic differential equation (9.2) possesses at least one weak solution starting from (s,x) if and only if the $\mathbb{L}_{fg}$-local martingale problem possesses at least one solution starting from (s,x). The stochastic differential equation (9.2) possesses exactly one in law weak solution starting from (s,x) if and only if the local $\mathbb{L}_{fg}$-martingale problem starting from (s,x) is well posed.* □

Sufficient conditions for the well-posedness of martingale problems are given by the following theorem.

Theorem 9.2. *Let $f : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^d$ and $g : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^{d\times m}$ be bounded and measurable and such that $\sigma = g\cdot g^*$ satisfies $\inf_{0\le t\le T} \inf_{\theta\in(\mathbb{R}^d\setminus\{0\})} \frac{\langle \theta, \sigma(t,x)\rangle}{|\theta|^2} > 0$ and $\lim_{y\to x} \sup_{0\le t\le T} ||\sigma(t,y) - \sigma(t,x)\| = 0$ for every $T > 0$ and $x \in \mathbb{R}^d$. For every $(s,x) \in \mathbb{R}^+ \times \mathbb{R}^d$, the local $\mathbb{L}_{fg}$-martingale problem starting from (s,x) is well posed.* □

We can now prove the following existence theorem.

Theorem 9.3. *Let $f : \mathbb{R}^+\times\mathbb{R}^d \to \mathbb{R}^d$ and $g : \mathbb{R}^+\times\mathbb{R}^d \to \mathbb{R}^{d\times m}$ be continuous and bounded. Then for every probability measure μ on $\beta(\mathbb{R}^d)$, there exists a weak solution $(\mathcal{P}_{\mathbb{F}}, X, P)$ of (9.2) with an initial distribution μ.*

Proof. By virtue of Theorem 9.1, for the existence of a weak solution $(\mathcal{P}_{\mathbb{F}}, X, P)$ of (9.2) with an initial distribution μ, it suffices to construct a d-dimensional process $\hat{X} = (\hat{X}_t)_{t\ge 0}$ on any filtered probability space $\hat{\mathcal{P}}_{\hat{\mathbb{F}}} = (\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{F}}, \hat{P})$ such that $\hat{X}$ is $\hat{\mathbb{F}}$-adapted, $P\hat{X}_0^{-1} = \mu$ and for every $h \in C_0^2$, the process $\varphi^h = (\varphi_t^h)_{t\ge 0}$ with $\varphi_t^h = h(\hat{X}_t) - h(\hat{X}_0) - \int_0^t (\mathbb{L}_{fg}h)(s, \hat{X}_s)\mathrm{d}s$ a.s. for $t \ge 0$ is a continuous local $\hat{\mathbb{F}}$-martingale. To do this, let us select a filtered probability space $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\ge 0}$ satisfying the usual conditions and such that there exists an m-dimensional $\mathbb{F}$-Brownian motion $B = (B_t)_{t\ge 0}$ defined on this space. Let $t_l^k = k/2^l$ for every $k, l = 0, 1, 2, \ldots$ and define f_l and g_l by setting $f_l(t,x) = f(t_l^k, x)$ and $g_l(t,x) = g(t_l^k, x)$ for $x \in \mathbb{R}^d$ and $k/2^l \le t < (k+1)/2^l$ for $k, l = 0, 1, 2, \ldots$. For every $l = 1, 2, \ldots$, the functions f_l and g_l are Borel-measurable and bounded. Select an $\mathcal{F}_0$-measurable $\mathbb{R}^d$-random variable ξ defined on $\mathcal{P}_{\mathbb{F}}$ such that $P\xi^{-1} = \mu$. Let us define on $\mathcal{P}_{\mathbb{F}}$ a sequence $(X^l)_{l=1}^\infty$ of d-dimensional stochastic processes $X^l = (X_t^l)_{t\ge 0}$ by setting $X_0^l = \xi$ a.s. and $X_t^l = X_{t_l^k}^l + f_l(t, X_{t_l^k}^l)(t - t_l^k) + g_l(t, X_{t_l^k}^l)(B_t - B_{t_l^k})$ a.s. for $l = 1, 2, \ldots$ and $t_l^k \le t \le t_l^{k+1}$ with $k = 0, 1, 2, \ldots$. It is clear that X^l is defined a.s. for $t \ge 0$, and for every $t \ge 0$ and $l = 1, 2, \ldots$ is $\mathcal{F}_t$-measurable. It is easy to see that X^l is continuous and that

$$X_t^l = X_s^l + \int_s^t f_l\left(\tau, X_\tau^l\right) \mathrm{d}\tau + \int_s^t g_l\left(\tau, X_\tau^l\right) \mathrm{d}B_\tau$$

a.s. for $0 \leq s < t \leq T$ and every $T > 0$. Hence in particular, it follows that for every $m = 1, 2, \ldots$, there is $C_m > 0$ such that

$$E[|X_t^l - X_s^l|^{2m}] \leq C_m E\left[\left(\int_s^t |f_l(\tau, X_\tau^l)|^2 \mathrm{d}\tau\right)^m\right] + C_m E\left[\left(\int_s^t \|g_l(\tau, X_\tau^l)\|^2 \mathrm{d}\tau\right)^m\right] \leq C_m M^2 |t-s|^m ,$$

where $M > 0$ is such that $\max(|f(t,x)|, \|g(t,x)\|) \leq M$ for $x \in \mathbb{R}^d$ and $t \geq 0$. Furthermore, $\sup_{l \geq 1} \sup_{0 \leq t \leq T} E[|X_t^l|^{2m}] \leq C_m$ for every $T > 0$. Therefore, by virtue of Theorem 2.4, there are an increasing subsequence $(l_i)_{i=1}^\infty$ of $(l)_{l=1}^\infty$, a probability space $\hat{\mathcal{P}} = (\hat{\Omega}, \hat{\mathcal{F}}, \hat{P})$, and d-dimensional continuous stochastic processes $\hat{X}$, $\hat{X}^{l_i}$, $i = 1, 2, \ldots$, defined on $\hat{\mathcal{P}}$ such that $P(X^{l_i})^{-1} = P(\hat{X}^{l_i})^{-1}$ for $i = 1, 2, \ldots$ and $\sup_{t \geq 0} |\hat{X}_t^{l_i} - \hat{X}_t| \to 0$ a.s. as $i \to \infty$. By Itô's formula, for every $h \in C_0^2(\mathbb{R}^d)$ and $i = 1, 2, \ldots$, one gets

$$E\left[h(X_t^{l_i}) - h(X_s^{l_i}) - \int_s^t \left(\mathbb{L}_{f_{l_i} g_{l_i}} h\right)(\tau, X_\tau^{l_i}) | \mathcal{F}_s\right] = E\left[\sum_{r=1}^d \sum_{j=1}^m \int_s^t h'_{x_r}(X_\tau^{l_i}) g_{l_i}^{r,j}(\tau, X_\tau^{l_i}) \mathrm{d}B_\tau | \mathcal{F}_s\right] = 0$$

a.s. for every $0 \leq s \leq t < \infty$. Then for every continuous and bounded function $F : \mathbb{R}^d \to \mathbb{R}$, we have

$$E\left(E\left[F(X_s^{l_i})\left(h(X_t^{l_i}) - h(X_s^{l_i}) - \int_s^t \left(\mathbb{L}_{f_{l_i} g_{l_i}} h\right)(\tau, X_\tau^{l_i})\right) | \mathcal{F}_s\right]\right) = 0$$

for every $0 \leq s \leq t < \infty$ and $i = 1, 2, \ldots$. Thus

$$E\left[F(X_s^{l_i})\left(h(X_t^{l_i}) - h(X_s^{l_i}) - \int_s^t \left(\mathbb{L}_{f_{l_i} g_{l_i}} h\right)(\tau, X_\tau^{l_i})\right)\right] = 0$$

for every $0 \leq s \leq t < \infty$ and $i = 1, 2, \ldots,$, which implies

$$\hat{E}\left[F(\hat{X}_s)\left(h(\hat{X}_t) - h(\hat{X}_s) - \int_s^t \left(\mathbb{L}_{fg} h\right)(\tau, \hat{X}_\tau)\right)\right] = \lim_{i \to \infty} \hat{E}\left[F(\hat{X}_s^{l_i})\left(h(\hat{X}_t^{l_i}) - h(\hat{X}_s) - \int_s^t \left(\mathbb{L}_{f_{l_i} g_{l_i}} h\right)(\tau, \hat{X}_\tau^{l_i})\right)\right] = 0$$

for every $0 \le s \le t < \infty$. Let $\hat{\mathcal{F}}_t = \bigcap_{\varepsilon>0} \sigma[\hat{X}_u : 0 \le u \le t+\varepsilon]$ for $t \ge 0$. Therefore,

$$\hat{E}\left(F(\hat{X}_s)\,\hat{E}\left[\left(h(\hat{X}_t) - h(\hat{X}_s) - \int_s^t \left(\mathbb{L}_{fg}h\right)(\tau, \hat{X}_\tau)\right) |\hat{\mathcal{F}}_s\right]\right) = 0$$

for every $0 \le s \le t < \infty$, which by the monotone class theorem can be extended for every bounded measurable function F. Taking in particular F such that

$$F(\hat{X}_s) = \hat{E}\left[\left(h(\hat{X}_t) - h(\hat{X}_s) - \int_s^t \left(\mathbb{L}_{fg}h\right)(\tau, \hat{X}_\tau)\right) |\hat{\mathcal{F}}_s\right],$$

we obtain

$$\hat{E}\left[\left(h(\hat{X}_t) - h(\hat{X}_s) - \int_s^t \left(\mathbb{L}_{fg}h\right)(\tau, \hat{X}_\tau)\right) |\hat{\mathcal{F}}_s\right] = 0$$

a.s. for $0 \le s \le t < \infty$ and $h \in C_0^2(\mathbb{R}^d)$. Then a process $\varphi^h = (\varphi_t^h)_{t\ge 0}$ defined by $\varphi_t^h = h(\hat{X}_t) - h(\hat{X}_0) - \int_0^t (\mathbb{L}_{fg}h)(\tau, \hat{X}_\tau)\mathrm{d}\tau$ a.s. for $t \ge 0$ is a continuous local $\hat{\mathbb{F}}$-martingale on $\hat{\mathcal{P}}_{\hat{\mathbb{F}}} = (\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{F}}, \hat{P})$ with $\hat{\mathbb{F}} = (\hat{\mathcal{F}}_t)_{t\ge 0}$. □

Remark 9.2. The boundedness of f and g in Theorem 9.3 can be weakened to the following linear growth condition: there exists a positive number K such that

$$|f(t,x)|^2 + \|g(t,x)\|^2 \le K^2(1 + |x|^2) \tag{9.5}$$

for every $t \ge 0$ and $x \in \mathbb{R}^d$. The condition (9.5) is necessary for the existence of global solutions of (9.2), i.e., of solutions of the form $X = (X_t)_{t\ge 0}$. □

The following uniqueness theorem follows immediately from Corollary 9.3 and Theorem 9.2.

Theorem 9.4. *Let $f : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^d$ and $g : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^{d\times m}$ be continuous and bounded. If g is such that the matrix function $\sigma : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$ defined by $\sigma(t,x) = g(t,x)\cdot g^*(t,x)$ is uniformly positive, then for every $(s,x) \in \mathbb{R}^+ \times \mathbb{R}^d$, the stochastic differential equation (9.2) possesses exactly one in law weak solution starting from (s,x).*

Proof. It is enough to observe that for fixed $(s,x) \in \mathbb{R}^+ \times \mathbb{R}^d$ and $T > 0$, the restriction of σ to $[0,T]\times B_x$, where B_x is a compact neighborhood of $x \in \mathbb{R}^d$, is uniformly continuous. Therefore, $\lim_{y\to x} \sup_{0\le s\le T} \|\sigma(s,y) - \sigma(s,x)\| = 0$. Now the result follows from Theorem 9.2 and Corollary 9.3. □

We can now define diffusion processes generated by weak solutions of stochastic differential equations. Let us note that a diffusion can be thought of as a continuous strong Markov process $X = (X_t)_{t\ge 0}$ on a filtered probability space $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$. It can be represented as a unique solution of the autonomous stochastic differential equation

$$X_t = X_0 + \int_0^t f(X_s)\mathrm{d}s + \int_0^t g(X_s)\mathrm{d}B_s\,, \tag{9.6}$$

where $f : \mathbb{R}^d \to \mathbb{R}^d$ and $g : \mathbb{R}^d \to \mathbb{R}^{d\times m}$ are given functions such that (9.6) has a unique in law weak solution. In what follows, by a diffusion starting with $(s,x) \in \mathbb{R}^+ \times \mathbb{R}^d$ we shall mean a unique in law weak solution $(\mathcal{P}_{\mathbb{F}}, X, B)$ for (9.6) with an initial condition $X_s = x$ a.s. Usually, it will be identified with a continuous process $X = (X(t))_{t\geq 0}$ instead of $(\mathcal{P}_{\mathbb{F}}, X, B)$. It is denoted simply by $X_{s,x}$ or $X^{fg}_{s,x}$. The following result can be obtained immediately from Theorem 9.4.

Corollary 9.4. *If $f : \mathbb{R}^d \to \mathbb{R}^d$ and $g : \mathbb{R}^d \to \mathbb{R}^{d\times m}$ are continuous and bounded, and g is such that the matrix function $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ defined by $\sigma(x) = g(x)\cdot g^*(x)$ is uniformly positive, then for every $(s,x) \in \mathbb{R}^+ \times \mathbb{R}^d$, there is a diffusion process $X^{fg}_{s,x} = (X^{fg}_{s,x}(t))_{t\geq 0}$ on a filtered probability space $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ supporting an m-dimensional $\mathbb{F}$-Brownian motion $B = (B_t)_{t\geq 0}$ such that $(\mathcal{P}_{\mathbb{F}}, X^{fg}_{s,x}, B)$ is a unique in law weak solution of (9.2) with the initial condition $X_s = x$ a.s.* □

In what follows, for a given $(s,x) \in \mathbb{R}^+ \times \mathbb{R}^d$, by $E^{s,x}$ we shall denote the mean value operator with respect to the probability law $Q^{s,x}$ of a diffusion $X^{s,x} = (X^{s,x}_t)_{t\geq 0}$. To define it, let $\mathcal{M}_{s,x}$ be a σ-algebra on Ω generated by random variables $X_t : \Omega \ni \omega \to X^{s,x}_t(\omega) \in \mathbb{R}^d$ with $t \geq s$ and $x \in \mathbb{R}^d$. Define a probability measure $Q^{s,x}$ on $\mathcal{M}_{s,x}$ such that $Q^{s,x}(\{X_{t_1} \in A_1, \ldots, X_{t_k} \in A_k\}) = P(\{X^{s,x}_{t_1} \in A_1, \ldots, X^{s,x}_{t_k} \in A_k\})$ for $A_i \in \beta(\mathbb{R}^d)$ with $i = 1, 2, \ldots, k$ and $k \geq 1$. If $s = 0$, we will write E^x and Q^x instead of $E^{0,x}$ and $Q^{0,x}$. We shall now prove that the above-defined diffusion has the following Markov properties.

Theorem 9.5. *Let $f : \mathbb{R}^d \to \mathbb{R}^d$ and $g : \mathbb{R}^d \to \mathbb{R}^{d\times m}$ be continuous and bounded, assume that g is such that the matrix function $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ defined by $\sigma(x) = g(x)\cdot g^*(x)$ is uniformly positive, and let $X = (X_t)_{t\geq 0}$ be a diffusion on $\mathcal{P}_{\mathbb{F}}$ starting with $(0,x)$ for $x \in \mathbb{R}^d$. For every bounded continuous function $F : \mathbb{R}^d \to \mathbb{R}$ and $t, h \geq 0$, one has*

$$E^x[F(X_{t+h})|\mathcal{F}_t](\omega) = E^{X_t(\omega)}[F(X_h)], \tag{9.7}$$

where $E^x[F(X_{t+h})|\mathcal{F}_t]$ denotes the conditional expectation with respect to Q^x.

Proof. For $r \geq t$, we have $X^{t,x}_r = X^{t,x}_t + \int_t^r f(X^{t,x}_\tau)\mathrm{d}\tau + \int_t^r g(X^{t,x}_\tau)\mathrm{d}B_\tau$ a.s. By the uniqueness in law of X, we have $E[F(X^{t,x}_r(\cdot)] = E[F(X^{t,X_t}_r(\cdot)]$. Denoting $X^{t,x}_r$ by $\varphi(x,t,r,\cdot)$ for $r \geq 1$, we get $E[F(X^{t,x}_r(\cdot)] = E[F(\varphi(x,t,r,\cdot))]$ for $r \geq t$. Note that $X^{t,x}_r$ is independent of $\mathcal{F}_t$. Then

$$E[F(\varphi(X_t,t,t+h,\cdot))|\mathcal{F}_t] = E[F(\varphi(x,0,h,\cdot))|\mathcal{F}_t]_{x=X_t}.$$

Let $\psi(x,\omega) = F\circ\varphi(x,t,t+h,\omega))$. It is clear that ψ is measurable. Then it can be approximated pointwise boundedly by a sequence $(l_k)_{k=1}^\infty$ of functions of the form $\sum_{j=1}^{p(k)} \lambda^k_j(x)\gamma^k_j(\omega)$. Therefore,

$$\begin{aligned}E[\psi(X_t,\cdot)|\mathcal{F}_t] &= E[\lim_{k\to\infty} l_k(X_t,\cdot)|\mathcal{F}_t] \\ &= \lim_{k\to\infty}\sum_{j=1}^{p(k)} \lambda_j^k(X_t)E[\gamma_j^k(\cdot)|\mathcal{F}_t] \\ &= \lim_{k\to\infty}\sum_{j=1}^{p(k)} E[\lambda_j^k(y)\gamma_j^k(\cdot)|\mathcal{F}_t]_{y=X_t} \\ &= E[\psi(y,\cdot)|\mathcal{F}_t]_{y=X_t} = E[\psi(y,\cdot)]_{y=X_t}\,.\end{aligned}$$

From the time-homogeneity of $X = (X_t)_{t\geq 0}$, it follows that

$$\begin{aligned}E[F(\varphi(X_t,t,t+h,\cdot))|\mathcal{F}_t] &= E[F(\varphi(y,t,t+h,\cdot))]_{y=X_t} \\ &= E[F(\varphi(y,0,h,\cdot))]_{y=X_t}\,.\end{aligned}$$

Then (9.7) is satisfied. □

Theorem 9.6. *Let* $f : \mathbb{R}^d \to \mathbb{R}^d$ *and* $g : \mathbb{R}^d \to \mathbb{R}^{d\times m}$ *be continuous and bounded and assume that* g *is such that the matrix function* $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ *defined by* $\sigma(x) = g(x)\cdot g^*(x)$ *is uniformly positive, and let* $X = (X_t)_{t\geq 0}$ *be a diffusion on* $\mathcal{P}_{\mathbb{F}}$ *starting with* $(0,x)$ *for* $x\in\mathbb{R}^d$. *For every bounded continuous function* $F:\mathbb{R}^d\to\mathbb{R}$ *and* $\mathbb{F}$*-stopping time* $\tau<\infty$, *a.s. one has*

$$E^x[F(X_{\tau+h})|\mathcal{F}_\tau](\omega) = E^{X_\tau(\omega)}[F(X_h)] \quad \text{for all } h\geq 0\,. \tag{9.8}$$

Proof. We have

$$X^{\tau,x}_{\tau+h} = x + \int_\tau^{\tau+h} f(X_s^{\tau,x})\mathrm{d}s + \int_\tau^{\tau+h} g(X_s^{\tau,x})\mathrm{d}B_s\,.$$

By the strong Markov property for Brownian motions, the process $\tilde{B}_v = B_{\tau+v}-B_\tau$ with $v\geq 0$ is again an $\mathbb{F}$-Brownian motion independent of $\mathcal{F}_\tau$. Therefore,

$$X^{\tau,x}_{\tau+h} = x + \int_0^{h} f(X_{\tau+v}^{\tau,x})\mathrm{d}v + \int_0^{h} g(X_{\tau+v}^{\tau,x})\mathrm{d}\tilde{B}_v\,.$$

Thus the process $(X^{\tau,x}_{\tau+h})_{h\geq 0}$ has the same distribution as $(Y_h)_{h\geq 0}$ defined by

$$Y_h = x + \int_0^h f(Y_v)\mathrm{d}v + \int_0^h g(Y_v)\mathrm{d}\tilde{B}_v\,.$$

Then $E[F(X^{\tau,x}_{\tau+h})] = E[F(Y_h)]$ for $h\geq 0$. Since $(Y_h)_{h\geq 0}$ is independent of $\mathcal{F}_\tau$, it follows that $(X^{\tau,x}_{\tau+h})_{h\geq 0}$ must be independent of $\mathcal{F}_\tau$, too. From the uniqueness in law of $(X_h^{0,x})_{h\geq 0}$, it follows that $(Y_h)_{h\geq 0}$ and $(X^{\tau,x}_{\tau+h})_{h\geq 0}$ have the same law as $(X_h^{0,x})_{h\geq 0}$. Let $\varphi(x,t,r,\omega) = X_r^{t,x}(\omega)$ for $r \geq t$ and $\omega \in \Omega$. Hence it follows that $E[F(\varphi(x,0,\tau+h,\cdot))|\mathcal{F}_\tau] = E[F(\varphi(x,0,h,\cdot))]_{x=X_\tau^{0,x}}$. Setting $X_t = X_t^{0,x}$, we get

$$\begin{aligned}\varphi(x,0,\tau+h,\omega)) = X_{\tau+h}(\omega) &= x + \int_0^{\tau+h} f(X_s)\mathrm{d}s + \int_0^{\tau+h} g(X_s)\mathrm{d}B_s \\ &= x + \int_0^{\tau} f(X_s)\mathrm{d}s + \int_0^{\tau} g(X_s)\mathrm{d}B_s \\ &\quad + \int_\tau^{\tau+h} f(X_s)\mathrm{d}s + \int_\tau^{\tau+h} g(X_s)\mathrm{d}B_s \\ &= \varphi(X_\tau,\tau,\tau+h,\omega)\,.\end{aligned}$$

Therefore, (9.8) can be written in the form $E[F(\varphi(X_\tau,\tau,\tau+h,\cdot))|\mathcal{F}_\tau] = E[F(\varphi(x,0,h,\cdot))]_{x=X_\tau}$. Putting $\psi(x,t,r,\omega) = F(\varphi(x,t,r,\omega))$, we can assume, similarly as in the proof of Theorem 9.5, that ψ has the form $\psi(x,t,r,\omega) = \sum_j \lambda_j(x)\gamma_j(t,r,\omega)$. Therefore,

$$\begin{aligned}E[\psi(X_\tau,\tau,\tau+h,\cdot)|\mathcal{F}_\tau] &= \sum_j E[\lambda_j(X_\tau)\gamma_j(\tau,\tau+h,\cdot)|\mathcal{F}_\tau] \\ &= \sum_j \lambda_j(X_\tau)E[\gamma_j(\tau,\tau+h,\cdot)|\mathcal{F}] \\ &= \sum_j E[\lambda_j(x)\gamma_j(\tau,\tau+h,\cdot)|\mathcal{F}_\tau]_{x=X_\tau} \\ &= E[\psi(x,\tau,\tau+h,\cdot)|\mathcal{F}_\tau]_{x=X_\tau} = E[F(X_{\tau+h}^{\tau,x})]_{x=X_\tau} \\ &= E[F(X_h^{0,x})]_{x=X_\tau} = E[F(\varphi(x,0,h,\cdot))]_{x=X_\tau}\,. \qquad \square\end{aligned}$$

Remark 9.3. By induction, we can extend (9.8) to k bounded Borel functions $f_1 : \mathbb{R}^d \to \mathbb{R}, \ldots, f_k : \mathbb{R}^d \to \mathbb{R}$ and get the relation

$$E^x[f_1(X_{\tau+h_1})f_2(X_{\tau+h_2})\ \ldots\ f_k(X_{\tau+h_k})] = E^{X_\tau}[f_1(X_{h_1})f_2(X_{h_2})\ \ldots\ f_k(X_{h_k})]$$

for the $\mathbb{F}$-stopping time τ and $h_1 \leq h_2 \leq \cdots \leq h_k$, which can be written in the general form $E^x[\theta_\tau\eta|\mathcal{F}_\tau] = E^{X_\tau}[\eta]$, where θ denotes the shift operator defined by $\theta_t\eta = f_1(X_{t_1+t})f_2(X_{t_2+t})\ \ldots\ f_k(X_{t_k+t})$ for $\eta = f_1(X_{t_1})f_2(X_{t_2})\ \ldots\ f_k(X_{t_k})$. $\square$

10 The Infinitesimal Generator of Diffusion

A diffusion process defined as the unique solution of the autonomous stochastic differential equation (9.6) is also known as Itô diffusion. There is another definition of diffusion processes. For a d-dimensional process $X = (X_t)_{t\geq 0}$ on a filtered probability space $\mathcal{P}_{\mathbb{F}}$, it must be a diffusion that is a continuous time-homogeneous $\mathbb{F}$-Markov process starting with $(0,x)$ such that the limit $\lim_{t\to 0}(1/t)\,[E^x[h(X_t)] - h(x)]$ exists for every h in a suitable subclass $\mathcal{D}_X$ of the space $C(\mathbb{R}^d, \mathbb{R})$. The existence of the above limit admits the definition of an operator on $\mathcal{D}_X$, called the infinitesimal generator of X. Such a diffusion is known as a Kolmogorov–Feller diffusion. Let us note that we can also associate the infinitesimal generator to an Itô diffusion $X = (X_t)_{t\geq 0}$ starting with $(s,x) \in \mathbb{R}^+ \times \mathbb{R}^d$. To define it, let us denote by $\mathcal{D}_X$ the set of all functions $h : \mathbb{R}^d \to \mathbb{R}$ such that the limit $\lim_{t\to 0}(1/t)\,[E^x[h(X_t)] - h(x)]$ exists for every $h \in \mathcal{D}_X$. The operator $\mathcal{A}_X$ defined on $\mathcal{D}_X$ by setting

$$(\mathcal{A}_X h)(x) = \lim_{t\to 0} \frac{E^x[h(X_t)] - f(x)}{t} \tag{10.1}$$

for every $h \in \mathcal{D}_X$ and $x \in \mathbb{R}^d$ is called the infinitesimal generator of the Itô diffusion X. The set $\mathcal{D}_X$ is called the domain of $\mathcal{A}_X$.

Remark 10.1. If X is the unique in law solution of the stochastic differential equation (9.6) starting from $(0,x)$, then $\mathcal{A}_X$ will be denoted by $\mathcal{A}_{fg}$. In such a case, the domain $\mathcal{D}_X$ is also denoted by $\mathcal{D}_{fg}$. □

Similarly as above, we shall consider an infinitesimal generator $\mathcal{A}_{fg}$ on the space $C_0^2(\mathbb{R}^d)$ defined above. It can be verified that if $h \in C_0^2(\mathbb{R}^d)$, then all partial derivatives of h up to order two are continuous and bounded. We have the following result.

Theorem 10.1. *Let $f : \mathbb{R}^d \to \mathbb{R}^d$ and $g : \mathbb{R}^d \to \mathbb{R}^{d\times m}$ be continuous and bounded and assume that g is such that the matrix function $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ defined by $\sigma(x) = g(x)\cdot g^*(x)$ is uniformly positive, and let $X = (X_t)_{t\geq 0}$ be an Itô diffusion on $\mathcal{P}_{\mathbb{F}}$ starting with $(0,x) \in \mathbb{R}^+ \times \mathbb{R}^d$ defined by the weak solution $(\mathcal{P}_{\mathbb{F}}, x, P)$ of (9.6). Then $C_0^2(\mathbb{R}^d) \subset \mathcal{D}_X$, and for every $h \in C_0^2(\mathbb{R}^d)$, one has $\mathcal{A}_{fg} h = \mathbb{L}_{fg} h$, i.e.,*

$$(\mathcal{A}_{fg} h)(x) = \sum_{i=1}^{d} f^i(x) h'_{x_i}(x) + \frac{1}{2}\sum_{i,j=1}^{d} \sigma^{ij}(x) h''_{x_i x_j}(x) \tag{10.2}$$

for $x \in \mathbb{R}^d$, where $(\sigma^{ij}(x))_{d\times d} = \sigma(x)$.

Proof. By Itô's formula, for every $h \in C_0^2(\mathbb{R}^d)$, we obtain

$$h(X_t) = h(X_0) + \int_0^t \left(\sum_{i=1}^d f^i(X_s) h'_{x_i}(X_s) + \frac{1}{2} \sum_{i,j=1}^d \sigma^{ij}(X_s) h''_{x_i x_j}(X_s) \right) \mathrm{d}s$$

$$+ \sum_{i=1}^d \sum_{j=1}^m \int_0^t \sigma^{ij}(X_s) h'_{x_i}(X_s) \mathrm{d}B^j . \tag{10.3}$$

Hence it follows that

$$E^x[h(X_t)] - h(x) = \int_0^t E^x \left[\sum_{i=1}^d f^i(X_s) h'_{x_i}(X_s) + \frac{1}{2} \sum_{i,j=1}^d \sigma^{ij}(X_s) h''_{x_i x_j}(X_s) \right] \mathrm{d}s$$

for $t \geq 0$, because $X_0 = x$ a.s. Therefore, by the continuity of functions f and g, the limit (10.1) exists for every x and is equal to the right-hand side of (10.2). In particular, it follows that $C_0^2(\mathbb{R}^d) \subset \mathcal{D}_X$. □

Corollary 10.1. *If the conditions of Theorem 10.1 are satisfied, then for every $h \in C_0^2(\mathbb{R}^d)$ and every $\mathbb{F}$-stopping time τ such that $E^x[\tau] < \infty$, the following Dynkin's formula is satisfied:*

$$E^x[h(X_\tau)] = h(x) + E^x \left[\int_0^\tau (\mathcal{A}_{fg} h)(X_s) \mathrm{d}s \right]$$

for $x \in \mathbb{R}^d$.

Proof. Let $G : \mathbb{R}^d \to \mathbb{R}$ be continuous and such that $|G(x)| \leq M$ for $M > 0$. For every positive integer k, one has $E^x[\int_0^{\tau \wedge k} G(X_s) \mathrm{d}B_s] = E^x[\int_0^k \mathbb{1}_{\{s<\tau\}} G(X_s) \mathrm{d}B_s] = 0$, because $G(X_s)$ and $\mathbb{1}_{\{s<\tau\}}$ are bounded and $\mathcal{F}_s$-measurable. Moreover,

$$E^x \left[\left(\int_0^{\tau \wedge k} G(X_s) \mathrm{d}B_s \right)^2 \right] = E^x \left[\int_0^{\tau \wedge k} |G(X_s)|^2 \mathrm{d}s \right] \leq M^2 E^x[\tau] < \infty .$$

Therefore, the family $\left(\int_0^{\tau \wedge k} G(X_s) \mathrm{d}B_s \right)_{k \geq 1}$ is uniformly integrable and

$$E^x \left[\left| \int_0^{\tau \wedge k} G(X_s) \mathrm{d}B_s - \int_0^\tau G(X_s) \mathrm{d}B_s \right|^2 \right] \to 0$$

as $k \to \infty$. Then $\lim_{k \to \infty} E^x[\int_0^{\tau \wedge k} G(X_s) \mathrm{d}B_s] = E^x[\int_0^\tau G(X_s) \mathrm{d}B_s] = 0$. Now the result follows immediately from (10.3). □

There are some problems for which it is much more suitable that the operators be defined in a more general way than the infinitesimal ones. Let $X = (X_t)_{t \geq 0}$ be an Itô diffusion defined by a weak solution $(\mathcal{P}_{\mathbb{F}}, X, B)$ of (9.6) with f and g satisfying the conditions of Corollary 9.4. The characteristic operator $\mathcal{L}_X$ of X is defined by

$$(\mathcal{L}_X h)(x) = \lim_{k \to \infty} \frac{E^x[h(X_{\tau_k})] - f(x)}{E^x[\tau_k]},$$

where the $\mathcal{U}_k$ are open sets decreasing to the point x, i.e., such that $\bigcap_{k=1}^{\infty} \mathcal{U}_k = \{x\}$, $\tau_k = \inf\{t > 0 : X_t \notin \mathcal{U}_k\}$, and h belongs to the set $\mathcal{C}_X$ of all functions $h : \mathbb{R}^d \to \mathbb{R}$ such that the above limit exists for all sequences $(\mathcal{U}_k)_{k=1}^{\infty}$. If $E^x[\tau_k] = 0$ for every open neighborhood $\mathcal{U}_k$ of x, then we define $(\mathcal{L}_X h)(x) = 0$. The set $\mathcal{C}_X$ is called the domain of $\mathcal{L}_X$.

Remark 10.2. It can be verified that $\mathcal{D}_X \subset \mathcal{C}_X$ and that for every $h \in \mathcal{D}_X$, one has $\mathcal{L}_X h = \mathcal{A}_X h$. If X is defined by (9.6), then $\mathcal{L}_X$ and $\mathcal{C}_X$ are denoted by $\mathcal{L}_{fg}$ and $\mathcal{C}_{fg}$, respectively. □

A point $x \in \mathbb{R}^d$ is called a trap for $X = (X_t)_{t \geq 0}$ if $Q^x(\{X_t = x \text{ for all } t \geq 0\}) = 1$, i.e., if $Q^x(\{\tau_x = \infty\}) = 1$, where $\tau_x = \inf\{t > 0 : |X_t| > |x|\}$.

Remark 10.3. It can be verified that if x is not a trap for X, then there exists an open set $\mathcal{U}$ containing x such that $E^x[\tau_{\mathcal{U}}] < \infty$. □

Let us consider characteristic operators on the space $C_0^2(\mathbb{R}^d)$. We have the following result.

Theorem 10.2. *Let f and g satisfy the conditions of Corollary 9.4 and let $X = (X_t)_{t \geq 0}$ be an Itô diffusion defined by the weak solution $(\mathcal{P}_{\mathbb{F}}, X, B)$ of (9.6). Then*

$$(\mathcal{L}_{fg} h)(x) = \sum_{i=1}^{d} f^i(x) h'_{x_i}(x) + \frac{1}{2} \sum_{i,j=1}^{d} \sigma^{ij}(x) h''_{x_i x_j}(x)$$

for every $h \in C_0^2(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$, where $\sigma = g \cdot g^*$ and $\sigma = (\sigma^{ij})_{d \times d}$. □

In applications of diffusion processes, very often the following question arises: when is the inequality $E^x[\tau_D] < \infty$ satisfied for a d-dimensional diffusion process $X = (X_x^{fg}(t))_{t \geq 0}$, a bounded domain $D \subset \mathbb{R}^n$, and $\tau_D = \inf\{t \geq 0 : X_x^{fg}(t) \notin D\}$? The answer is given by the following lemma.

Lemma 10.1. *Let $D \subset \mathbb{R}^d$ be a bounded domain and suppose $f : \mathbb{R}^d \to \mathbb{R}^d$ and $g : \mathbb{R}^d \to \mathbb{R}^{d \times m}$ are continuous, bounded, and such that the matrix function $\sigma : \mathbb{R}^n \to \mathbb{R}^{n \times n}$ defined by $\sigma(x) = g(x) \cdot g^*(x)$ for $x \in \mathbb{R}^d$ is uniformly positive on D, and let $X = (X_x^{fg}(t))_{t \geq 0}$ be a unique in law weak solution of $SDE(f, g)$ with initial condition $X_x^{fg}(0) = x$ a.s. Then $E^x[\tau_D] < \infty$.*

Proof. By the properties of the function g, we have $\min_{x \in \bar{D}} \sum_{i,j}^{d} \sigma_{ij}(x) \xi_i \xi_j \geq \lambda |\xi|^2 > 0$ for every $\xi \in \mathbb{R}^d$ and $\lambda > 0$. Hence it follows that $\min_{x \in \bar{D}} \sigma_{ii}(x) > 0$ for every $1 \leq i \leq d$. Fix $1 \leq i \leq d$ and let $a = \min_{x \in \bar{D}} \sigma_{ii}(x)$, $b =$

$\max_{x \in \bar{D}} |f(x)|$, $q = \min_{x \in \bar{D}} x_i$, and $\nu > (2b/a)$. Consider the function $h(x) = -\mu \cdot \exp(\nu x_i)$ for $x = (x_1, \ldots, x_d) \in D$, where the constant $\mu > 0$ will be appropriately selected later. It is clear that $h \in C^\infty(D, \mathbb{R})$ and $f_i(x) \geq -|f_i(x)| \geq -b$. Then

$$-(\mathbb{L}_{fg}h)(x) = \mu e^{\nu x_i} \left[\frac{1}{2} \nu^2 \sigma_{ii}(x) + \nu f_i(x) \right] \geq \frac{\mu}{2} \nu a \cdot e^{\nu q} \left(\nu - \frac{2b}{a} \right)$$

for $x \in D$. Choosing μ sufficiently large, we can guarantee that $(\mathbb{L}_{gh}h)(x) \leq -1$ for every $x \in D$. The function h and its derivatives are bounded on $\bar{D}$, so by the last inequality and Itô's formula, we get $E^x(t \wedge \tau_D) \leq h(x) - E^x[h(X_x^{fg}(t \wedge \tau_D))] \leq 2 \max_{z \in \bar{D}} |h(z)| < \infty$ for every $x \in D$ and $t \geq 0$ which in the limit $t \to \infty$ leads to the inequality $E^x[\tau_D] < \infty$. □

Remark 10.4. The above result is also true for continuous bounded functions $f : \mathbb{R}^d \to \mathbb{R}^d$, $g : \mathbb{R}^d \to \mathbb{R}^{d \times m}$, $\tilde{f} : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$, and $\tilde{g} : [0, T] \times \mathbb{R}^d \to \mathbb{R}^{d \times m}$ such that $\min_{x \in \bar{D}} \sigma_{ii}(x) > 0$ and $\min_{(t,x) \in [0,T] \times \bar{D}} \tilde{\sigma}_{ii}(t, x) > 0$, respectively, for some $1 \leq i \leq d$, where $\sigma(x) = g(x) \cdot g^*(x)$, $\tilde{\sigma}(t, x) = \tilde{g}(t, x) \cdot \tilde{g}^*(t, x)$, and $\tilde{\sigma} = (\tilde{\sigma}_{ij})_{n \times n}$. □

11 Diffusions Defined by Nonautonomous Stochastic Differential Equations

Let $f : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^d$ and $g : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^{d \times m}$ be given. In what follows, we shall say that f and g satisfy conditions $(\mathcal{C})$ if f and g are continuous and bounded, and g is such that the matrix function $\sigma = g \cdot g^*$ is uniformly positive. Immediately from Theorem 9.4, it follows that for such f and g and every $(s, x) \in \mathbb{R}^+ \times \mathbb{R}^d$, the stochastic differential equation

$$X_t = x + \int_s^t f(\tau, X_\tau) d\tau + \int_s^t g(\tau, X_\tau) dB_\tau \qquad (11.1)$$

possesses a unique in law weak solution $(\mathcal{P}_{\mathbb{F}}, X, B)$. Let us observe that the process $X = (X_t)_{t \geq 0}$ defined by (11.1) is not a strong Markov process, which makes it unable to be a diffusion process. However, with an extra argument, by extending its state space, we can conclude that the above weak solution X defines a $(d + 1)$-dimensional Itô diffusion $Y = (Y_t)_{t \geq 0}$ of the form $Y_t = (s + t, X_{s+t}^*)^*$. To get an appropriate stochastic differential equation for Y, let us define new functions $\mathbf{f} : \mathbb{R}^+ \times \mathbb{R}^d \ni (t, x) \to \mathbf{f}(t, x) \in \mathbb{R}^{d+1}$ and $\mathbf{g} : \mathbb{R}^+ \times \mathbb{R}^d \ni (t, x) \to \mathbf{g}(t, x) \in \mathbb{R}^{(d+1) \times m}$ by setting $\mathbf{f}(t, x) = (1, f(t, x))^*$ and $\mathbf{g}(t, x) = (\mathbf{0}, \mathbf{g}_1(t, x), \ldots, \mathbf{g}_d(t, x))^*$ with $\mathbf{0}, \mathbf{g}_i(t, x) \in \mathbb{R}^{1 \times m}$, where $\mathbf{0} = (0, \ldots 0)$ and $\mathbf{g}_i(t, x)$ denotes for every $i = 1, 2, \ldots, d$, the ith row of the matrix g. It is clear that $(\mathcal{P}_{\mathbb{F}}, Y, \tilde{B})$ is a weak solution of the autonomous stochastic differential equation

$$Y_t = (s,x) + \int_0^t \mathbf{f}(Y_\tau)\mathrm{d}\tau + \int_0^t \mathbf{g}(Y_\tau)\mathrm{d}\tilde{B}_\tau, \tag{11.2}$$

where $\tilde{B} = (\tilde{B}_t)_{t\geq 0}$ with $\tilde{B}_t = B_{s+t} - B_s$ for $t \geq 0$. Indeed, by a change of variables τ to $s+\tau$ in the formula (11.1) defining X_{s+t}, we get

$$X_{s+t} = x + \int_0^t f(s+\tau, X_{s+\tau})\mathrm{d}\tau + \int_0^t g(s+\tau, X_{s+\tau})\mathrm{d}\tilde{B}_\tau$$

a.s. for $t \geq 0$. Therefore,

$$\begin{aligned} Y_t = (s+t, X_{s+t}) &= \left(s+t, x + \int_0^t f(Y_\tau)\mathrm{d}\tau + \int_0^t g(Y_\tau)\mathrm{d}\tilde{B}_\tau\right) \\ &= (s,x) + \int_0^t (1, f(Y_\tau))\mathrm{d}\tau + \int_0^t (\mathbf{0}, \mathbf{g}_1(Y_\tau), \ldots, \mathbf{g}_d(Y_\tau))^*)\mathrm{d}\tilde{B}_\tau \\ &= (s,x) + \int_0^t \mathbf{f}(Y_\tau)\mathrm{d}\tau + \int_0^t \mathbf{g}(Y_\tau)\mathrm{d}\tilde{B}_\tau \end{aligned}$$

a.s. for $t \geq 0$. It can also be proved that uniqueness in law of $(\mathcal{P}_{\mathbb{F}}, X, B)$ implies that $(\mathcal{P}_{\mathbb{F}}, Y, \tilde{B})$ is a unique in law weak solution of (11.2). In what follows, for simplicity of notation we shall denote a vector $(s+t, X^*_{s+t})^*$ by $(s+t, X_{s+t})$. We have the following theorem.

Theorem 11.1. *If f and g satisfy conditions $(\mathcal{C})$, and $(\mathcal{P}_{\mathbb{F}}, X, B)$ is a unique in law weak solution of (11.1) with $X = (X_t)_{t\geq 0}$ then the process $Y = (Y_t)_{t\geq 0}$ defined by $Y_t = (s+t, X_{s+t})$ for $t \geq 0$ is an $(d+1)$-dimensional Itô diffusion defined for every fixed $(s,x) \in \mathbb{R}^+ \times \mathbb{R}^d$ by a weak solution of (11.2).*

Proof. It has been verified that $(\mathcal{P}_{\mathbb{F}}, Y, \tilde{B})$ is a weak solution of (11.2) with $\mathbf{f}$ and $\mathbf{g}$ as defined above. We have, therefore, only to verify its uniqueness in law. For simplicity, let us assume that $s = 0$. Assume that $(\mathcal{P}_{\tilde{\mathbb{F}}}, \tilde{X}, \tilde{B})$ is another weak solution of (11.1) and let $(\mathcal{P}_{\tilde{\mathbb{F}}}, \tilde{Y}, \hat{B})$ be a weak solution of (11.2) with $s = 0$ such that $\tilde{Y}_t = (t, \tilde{X}_t)$. By the uniqueness in law of the solution $(\mathcal{P}_{\mathbb{F}}, X, B)$, we have $PX^{-1} = P\tilde{X}^{-1}$ on $\beta(\mathbb{R}^d)$. This is equivalent to $P(X_{t_1}, \ldots, X_{t_r})^{-1} = P(\tilde{X}_{t_1}, \ldots, \tilde{X}_{t_r})^{-1}$ on $\beta(\mathbb{R}^{rd})$ for $0 \leq t_1 < \cdots < t_r < \infty$. Let $Q = \{A \times B : A \in \beta(\mathbb{R}^r_+), B \in \beta(\mathbb{R}^{rd}_+)\}$. It is clear that Q is a π-system such that $\beta(\mathbb{R}^r_+ \times \mathbb{R}^{rd}_+) = \sigma(Q)$. For every $0 \leq t_1 < \cdots < t_r < \infty$ and $A \times B \in Q$, one has $P((Y_{t_1}, \ldots, Y_{t_r})^{-1}(A \times B)) = \mathbb{1}_{(t_1,\ldots,t_r)}(A)P((X_{t_1}, \ldots, X_{t_r})^{-1}(B)) = \mathbb{1}_{(t_1,\ldots,t_r)}(A)\tilde{P}((\tilde{X}_{t_1}, \ldots, \tilde{X}_{t_r})^{-1}(B)) = \tilde{P}((\tilde{Y}_{t_1}, \ldots, \tilde{Y}_{t_r})^{-1}(A \times B))$. Hence by Dynkin's theorem, it follows that $P(Y_{t_1}, \ldots, Y_{t_r})^{-1} = P(\tilde{Y}_{t_1}, \ldots, \tilde{Y}_{t_r})^{-1}$ on $\sigma(Q)$ for every $0 \leq t_1 < \cdots < t_r < \infty$ that is equivalent to $PY^{-1} = P\tilde{Y}^{-1}$ on $\beta(C)$, where $C = C(\mathbb{R}^+, \mathbb{R}^{d+1})$. □

Remark 11.1. In a similar way as above, we can prove that if f and g are measurable, bounded, and such that for every $(s,x) \in \mathbb{R}^+ \times \mathbb{R}^d$, the stochastic differential equation (11.1) possesses a unique in law weak solution $(\mathcal{P}_{\mathbb{F}}, X, B)$ with $X = (X_t)_{t\geq 0}$, then the process $Y = (Y_t)_{t\geq 0}$ defined by $Y_t = (s+t, X_{s+t})$ for $t \geq 0$ is a $(d+1)$-dimensional Itô diffusion. □

The properties of d-dimensional Itô diffusions presented above are also true for $(d+1)$-dimensional diffusions defined by unique in law weak solutions of nonautonomous stochastic differential equations. In particular, the infinitesimal generator and the characteristic operator for $Y = (Y_t)_{t\geq 0}$ with $Y_0 = (s,x)$ and $Y_t = (s+t, Y_{s+t})$ for $t > 0$ can be defined by

$$(\mathcal{A}_{fg}\tilde{h})(s,x) = \tilde{h}'_t(s,x) + \sum_{i=1}^{d} f^i(s,x)\tilde{h}'_{x_i}(s,x) + \frac{1}{2}\sum_{i,j=1}^{d} \sigma^{ij}(s,x)\tilde{h}''_{x_i x_j}(s,x)$$

for every $\tilde{h} \in C_0^{1,2}(\mathbb{R}^{d+1})$ and $(s,x) \in \mathbb{R}^+ \times \mathbb{R}^d$ and

$$(\mathcal{L}_{fg}\tilde{h})(s,x) = \tilde{h}'_t(s,x) + \sum_{i=1}^{d} f^i(s,x)\tilde{h}'_{x_i}(s,x) + \frac{1}{2}\sum_{i,j=1}^{d} \sigma^{ij}(s,x)\tilde{h}''_{x_i x_j}(s,x)$$

for every $\tilde{h} \in C_0^{1,2}(\mathbb{R}^{d+1})$ and $(s,x) \in \mathbb{R}^+ \times \mathbb{R}^d$, where $\sigma = g \cdot g^*$ and $C_0^{1,2}(\mathbb{R}^{d+1}) = C_0^{1,2}(\mathbb{R}^+ \times \mathbb{R}^d, \mathbb{R})$.

12 The Feynman–Kac Formula

Given functions f and g defined above satisfying conditions $(\mathcal{C})$, we shall consider their restrictions to the set $[0,T] \times \mathbb{R}^d$ with $T > 0$. They will still be denoted by f and g, respectively. A weak solution of (11.1) corresponding to a pair (f,g) and $(s,x) \in [0,T] \times \mathbb{R}^d$ are denoted by $X_{s,x}^{fg}$. We still denote by $\mathcal{A}_{fg}$ the infinitesimal generator of a $(d+1)$-dimensional Itô diffusion $Y_{s,x}^{fg} = (Y_{s,x}^{fg}(t))_{t\geq 0}$ with $Y_{s,x}^{fg}(t) = (s+t, X_{s,x}^{fg}(s+t))$. We obtain the following Feynman–Kac theorem.

Theorem 12.1. *Assume that f and g satisfy conditions $(\mathcal{C})$, $T > 0$, and let $c \in C([0,T] \times \mathbb{R}^d)$ be bounded. Then for every $(s,x) \in [0,T) \times \mathbb{R}^d$, there is a unique in law solution $(\mathcal{P}_{\mathbb{F}}, X_{s,x}^{fg}, B)$ of (9.2) satisfying $X_{s,x}^{fg}(s) = x$ a.s. such that the function v defined by*

$$v(t,s,x) = E^{s,x}\left[\exp\left(-\int_0^t c(Y_{s,x}^{fg}(\tau))\mathrm{d}\tau\right)\tilde{h}(Y_{s,x}^{fg}(t))\right]$$

for $\tilde{h} \in C_0^{1,2}(\mathbb{R}^{d+1})$, $(s,x) \in [0,T) \times \mathbb{R}^d$ *and* $t \in [0, T-s]$ *satisfies*

$$\begin{cases} v_t'(t,s,x) = \left(\mathcal{A}_{fg} v(t,\cdot)\right)(s,x) - c(s,x)v(t,s,x) \\ \textit{for } \ (s,x) \in [0,T) \times \mathbb{R}^d \ \textit{and } t \in [0, T-s] \\ v(0,s,x) = \tilde{h}(s,x) \ \textit{for } (s,x) \in [0,T) \times \mathbb{R}^d \,. \end{cases}$$

Proof. The existence and in law uniqueness of $X_{s,x}^{fg}$ follow immediately from Theorem 9.4. Fix $s \in [0,T)$ and let $U_t^s = \tilde{h}(Y_{s,x}^{fg}(t))$ and $V_t^s = \exp(-\int_0^t c(Y_{s,x}^{fg}(\tau))\mathrm{d}\tau)$ for $t \in [0, T-s]$. Then

$$U_t^s = \tilde{h}(s,x) + \int_0^t (\mathcal{A}_{fg}\tilde{h})(Y_{s,x}^{fg}(\tau))\mathrm{d}\tau$$

$$+ \int_0^t \sum_{i=1}^n \sum_{j=1}^m (g \cdot g^*)_{i,j} (Y_{s,x}^{fg}(\tau)) h'_{x_i}(Y_{s,x}^{fg}(\tau))\mathrm{d}B_\tau \,,$$

$v(t,s,x)) = E^{s,x}[U_t^s V_t^s]$, and $\mathrm{d}V_t^s = -V_t^s \cdot c(Y_{s,x}^{fg}(t))\mathrm{d}t$. Hence it follows that $\mathrm{d}(U_t^s V_t^s) = U_t^s \mathrm{d}V_t^s + V_t^s \mathrm{d}U_t^s$, since $\mathrm{d}U_t^s \cdot \mathrm{d}V_t^s = 0$. Then $(U_t^s V_t^s)_{0 \le t \le T}$ is an Itô process, and by Itô's formula, we get

$$E^{s,x}[U_t^s V_t^s] = \tilde{h}(s,x) + E^{s,x}\left[\int_0^t V_\tau^s \cdot (\mathcal{A}_{fg}\tilde{h})(Y_{s,x}^{fg}(\tau))\mathrm{d}\tau\right]$$

$$-E^{s,x}\left[\int_0^t U_\tau^s \cdot c(Y_{s,x}^{fg}(\tau)) V_\tau^s \mathrm{d}\tau\right]$$

for $t \in [0, T-s]$. Hence it follows that $v(\cdot, s, x))$ is differentiable for fixed $(s,x) \in [0,T) \times \mathbb{R}^d$ and

$$\frac{1}{r}\left[E^{s,x}[v(t, Y_{s,x}^{fg}(r))] - v(t,s,x))\right]$$

$$= \frac{1}{r}\left[E^{s,x}\{E^{Y_{s,x}^{fg}(r)}[V_t^s \tilde{h}(Y_{s,x}^{fg}(t))] - E^{s,x}[V_t^s \tilde{h}(Y_{s,x}^{fg}(t))]\}\right]$$

$$= \frac{1}{r} E^{s,x}\Big\{E^{s,x}\Big[\tilde{h}(Y_{s,x}^{fg}(t+r))$$

$$\exp\left(-\int_0^t c(Y_{s,x}^{fg}(\tau+r))\mathrm{d}\tau\right) |\mathcal{F}_r\Big] - E^{s,x}[V_t^s \tilde{h}(Y_{s,x}^{fg}(t))|\mathcal{F}_r]\Big\}$$

$$= \frac{1}{r} E^{s,x} \Big[V^s_{t+r} \tilde{h}(Y^{fg}_{s,x}(t+r))$$
$$\exp\left(\int_0^r c(Y^{fg}_{s,x}(\tau)) \mathrm{d}\tau \right) - V^s_t \tilde{h}(Y^{fg}_{s,x}(t)) \Big]$$
$$= \frac{1}{r} E^{s,x} \left[V^s_{t+r} \tilde{h}(Y^{fg}_{s,x}(t+r)) - V^s_t \tilde{h}(Y^{fg}_{s,x}(t)) \right]$$
$$+ \frac{1}{r} E^{s,x} \left\{ V^s_{t+r} \tilde{h}(Y^{fg}_{s,x}(t+r)) \left[\exp\left(\int_0^r c(Y^{fg}_{s,x}(\tau)) \mathrm{d}\tau \right) - 1 \right] \right\}.$$

But

$$\lim_{r \to 0} \frac{1}{r} \left[V^s_{t+r} \tilde{h}(Y^{fg}_{s,x}(t+r)) - V^s_t \tilde{h}(Y^{fg}_{s,x}(t)) \right]$$
$$= \lim_{r \to 0} \frac{1}{r} \left[v(t+r, s, x)) - v(t, s, x)) \right] = v'_t(t, s, x)$$

and

$$\lim_{r \to 0} \frac{1}{r} \left[V^s_{t+r} \tilde{h}(Y^{fg}_{s,x}(t+r)) \left[\exp\left(\int_0^r c(Y^{fg}_{s,x}(\tau)) \mathrm{d}\tau \right) - 1 \right] \right]$$
$$= E^{s,x} \left\{ \tilde{h}(Y^{fg}_{s,x}(t)) V^s_t \cdot \lim_{r \to 0} \frac{1}{r} \left[\exp\left(\int_0^r c(Y^{fg}_{s,x}(\tau)) \mathrm{d}\tau \right) - 1 \right] \right\}$$
$$= E^{s,x} \left\{ \tilde{h}(Y^{fg}_{s,x}(t)) V^s_t \lim_{r \to 0} \frac{1}{r} \left[c(Y^{fg}_{s,x}(r) \exp\left(\int_0^r c(Y^{fg}_{s,x}(\tau)) \mathrm{d}\tau \right) \right] \right\}$$
$$= v(t, s, x) \cdot c(s, x).$$

Therefore, $\left(\mathcal{A}_{fg} v(t, \cdot)\right)(s, x) = v'_t(t, s, x) + v(t, s, x) \cdot c(s, x)$ and $v(0, s, x) = E^{s,x} \tilde{h}(Y^{fg}_{s,x}(0)) = \tilde{h}(s, x)$. ☐

Corollary 12.1. *Assume that f and g satisfy conditions ($\mathcal{C}$), $T > 0$, and let $c \in C([0, T] \times \mathbb{R}^d, \mathbb{R})$ be bounded. Then for every $x \in \mathbb{R}^d$, there is a unique in law solution $X^{fg}_{0,x}$ of (9.2) satisfying $X^{fg}_{0,x}(0) = x$ a.s. such that the function*

$$v(t, x) = E^x \left[\exp\left(- \int_0^t c(Y^{fg}_{0,x}(\tau)) \mathrm{d}\tau \right) (h \circ p_r)(Y^{fg}_{0,x}(t)) \right]$$

satisfies

$$\begin{cases} v'_t(t, x) = \left(\mathcal{A}_{fg} v(t, \cdot)\right)(t, x) - c(t, x) v(t, x) \ for \\ (t, x) \in (0, T] \times \mathbb{R}^d, v(0, x) = h(x) \ for \ x \in \mathbb{R}^d \end{cases}$$

for $h \in C^2_0(\mathbb{R}^d)$ and $(t, x) \in [0, T] \times \mathbb{R}^d$, where p_r is the orthogonal projection of $\mathbb{R}^{d+1}$ onto $\mathbb{R}^d$. ☐

We can also prove the following theorem.

Theorem 12.2. *Assume that* f *and* g *satisfy conditions* $(\mathcal{C})$*,* $T > 0$*, and let* $c \in C(\mathbb{R}^+ \times \mathbb{R}^n, \mathbb{R})$ *and* $v \in C^{1,1,2}([0,T] \times \mathbb{R}^+ \times \mathbb{R}^d, \mathbb{R})$ *be bounded such that*

$$\begin{cases} v'_t(t,s,x) = \left(\mathcal{A}_{fg} v(t,\cdot)\right)(s,x) - c(s,x)v(t,s,x) \\ for\ (s,x) \in [0,T) \times \mathbb{R}^n\ and\ t \in [0,T-s] \\ v(0,s,x) = \tilde{h}(s,x)\ for\ (s,x) \in [0,T) \times \mathbb{R}^d \end{cases} \tag{12.1}$$

for $\tilde{h} \in C_0^{1,2}(\mathbb{R}^d)$. *For every* $(s,x) \in [0,T) \times \mathbb{R}^d$*, there is a unique in law solution* $X^{fg}_{s,x}$ *of* (9.2) *satisfying* $X^{fg}_{s,x}(s) = x$ *such that*

$$v(t,s,x) = E^{s,x}\left[\exp\left(-\int_s^{s+t} c(\tau, X^{fg}_{s,x}(\tau))\mathrm{d}\tau\right)\tilde{h}(s+t, X^{fg}_{s,x}(s+t))\right]$$

with $(s,x) \in [0,T) \times \mathbb{R}^d$ *and* $t \in [0,T-s]$.

Proof. By virtue of Theorem 9.4, for every $(s,x) \in \mathbb{R}^+ \times \mathbb{R}^d$, there is a unique in law weak solution $(\mathcal{P}_{\mathbb{F}}, X^{fg}_{s,x}, B)$ of (9.2) such that $X^{fg}_{s,x}(s) = x$ a.s. Hence by Theorem 11.1, the process $Y^{fg}_{s,x} = (Y^{fg}_{s,x}(t))_{t\geq 0}$ defined by $Y^{fg}_{s,x}(t) = (s+t, X^{fg}_{s,x}(s+t))$ is a $(d+1)$-dimensional Itô diffusion on $\mathcal{P}_{\mathbb{F}}$. Assume that $v \in C^{1,1,2}([0,T] \times [0,T] \times \mathbb{R}^d, \mathbb{R})$ is bounded and satisfies conditions (12.1). Define $\hat{\mathcal{A}}_{fg}$ by setting $\left(\hat{\mathcal{A}}_{fg} v(t,\cdot)\right)(s,x) = -v'_t(t,s,x) + \left(\mathcal{A}_{fg} v(t,\cdot)\right)(s,x) - c(s,x)v(t,s,x)$ for $(s,x) \in [0,T) \times \mathbb{R}^d$ and $t \in [0,T-s]$. We have $\left(\hat{\mathcal{A}}_{fg} v(t,\cdot)\right)(s,x) = 0$ for $(s,x) \in [0,T) \times \mathbb{R}^d$ and $t \in [0,T-s]$. Fix $(u,s,x,z) \in \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R}$ and define $Z_{s,x}(t) = z + \int_0^t c(Y^{fg}_{s,x}(\tau))\mathrm{d}\tau$ and $H_{s,x}(t) = (u-t, Y^{f,g}_{s,x}(t), Z_{s,x}(t))$ for $t \in [0,T-s]$. Similarly as in the proof of Theorem 11.1, we can verify that $(H_{s,x}(t))_{0\leq t\leq T-s}$ is a $(d+3)$-dimensional Itô diffusion with infinitesimal generator $\mathcal{A}_H$ defined by

$$(\mathcal{A}_H \psi(t,\cdot))(s,x,z) = -\psi'_t(t,s,x,z) + (\mathcal{A}_{fg}\psi(t,\cdot)(s,x,z) + c(s,x)\psi'_z(t,s,x,z)$$

for $\psi \in C_0^{1,1,2,2}([0,T] \times \mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R})$. By Dynkin's formula, for $\psi(t,s,x,z) = \exp(-z)v(t,s,x)$, we get

$$E^{s,x,z}[\psi(H_{t\wedge\tau_{\mathcal{R}}})] = \psi(t,s,x,z) + E^{s,x,z}\left[\int_0^{t\wedge\tau_{\mathcal{R}}} (\mathcal{A}_H)\psi)(H_{s,x}(r)\mathrm{d}r\right]$$

for all $t \in [0,T-s)$ and $\mathcal{R} > 0$, where $\tau_{\mathcal{R}} = \inf\{t > 0 : |H_{s,x}(t)| \geq \mathcal{R}\}$. For such ψ, we have

$$(\mathcal{A}_H)\psi(t,\cdot)(s,x,z)$$
$$= \exp(-z)\Big[-v_t'(t,s,x) + \big(\mathcal{A}_{fg}v(t,\cdot)\big)(s,x) - c(s,x)v(t,s,x)\Big] = 0\,.$$

Therefore,

$$v(t,s,x) = \psi(t,s,x,0) = E^{s,x,0}[\psi(H_{t\wedge\tau_{\mathcal{R}}})]$$
$$= E^{s,x}\Big[\exp\Big(-\int_0^{t\wedge\tau_{\mathcal{R}}} c(Y_{s,x}^{fg}(r))\mathrm{d}r\Big)v(u-t\wedge\tau_{\mathcal{R}}, Y_{s,x}^{fg}(t\wedge\tau_{\mathcal{R}})\Big].$$

By continuity and the boundedness of v, it follows that

$$\lim_{\mathcal{R}\to\infty} E^{s,x}\Big[\exp\Big(-\int_0^{t\wedge\tau_{\mathcal{R}}} c(Y_{s,x}^{fg}(r))\mathrm{d}r\Big)v(u-t\wedge\tau_{\mathcal{R}}, Y_{s,x}^{fg}(t\wedge\tau_{\mathcal{R}})\Big]$$
$$= E^{s,x}\Big[\exp\Big(-\int_0^{t} c(Y_{s,x}^{fg}(r))\mathrm{d}r\Big)v(u-t, Y_{s,x}^{fg}(t)\Big].$$

Taking in particular $u = t$, we obtain

$$E^{s,x}\Big[\exp\Big(-\int_0^{t} c(Y_{s,x}^{fg}(r))\mathrm{d}r\Big)v(0, Y_{s,x}^{fg}(t)\Big]$$
$$= E^{s,x}\Big[\exp\Big(-\int_0^{t} c(Y_{s,x}^{fg}(r))\mathrm{d}r\Big)\tilde{h}(Y_{s,x}^{fg}(t)\Big] = v(t,s,x),$$

which can be written in the form

$$v(t,s,x) = E^{s,x}\left[\exp\left(-\int_s^{s+t} c(\tau, X_{s,x}^{fg}(\tau))\mathrm{d}\tau\right)\tilde{h}(s+t, X_{s,x}^{fg}(s+t))\right]$$

for $(s,x) \in [0,T)\times\mathbb{R}^d$ and $t \in [0, T-s]$. □

Corollary 12.2. *If the assumptions of Theorem 12.1 are satisfied,* $w \in C^{1,2}([0,T]\times\mathbb{R}^d,\mathbb{R})$ *and* $u \in C([0,T]\times\mathbb{R}^d)$ *are bounded such that the function* v *defined by* $v(s,x) = w(T-s,x)$ *for* $0 \le s < T$ *satisfies*

$$\begin{cases} v_s'(s,x) + (\mathcal{A}_{fg}v)(s,x) = -u(s,x) \;\; for \;\; (s,x)\in[0,T)\times\mathbb{R}^d, \\ v(0,x) = \tilde{h}(T,x) \;\; for \;\; x\in\mathbb{R}^d, \end{cases}$$

then for every $(s,x) \in [0,T)\times\mathbb{R}^d$*, there is a unique in law weak solution* $X_{s,x}^{fg}$ *of* (9.2) *satisfying* $X_{s,x}^{fg}(s) = x$ *such that* $w(T-s,x) = E^{s,x}\left[\tilde{h}(T, X_{s,x}^{fg}(T))\right] + E^{s,x}\left[\int_s^T u(\tau, X_{s,x}^{fg}(\tau))\mathrm{d}\tau\right]$. □

13 Harmonic Measure and the Mean Value Property of Diffusions

Let $H \subset \mathbb{R}^d$ be measurable and let τ_H be the first exit time from H for an Itô diffusion $X = (X_t)_{t \geq 0}$. Let α be another stopping time and g a bounded continuous function on $\mathbb{R}^d$. Put $\eta = g(X_{\tau_H})\mathbb{1}_{\{\tau_H < \infty\}}$ and $\tau_H^\alpha = \inf\{t > \alpha : X_t \notin H\}$. It can be verified that $\theta_\alpha \eta \mathbb{1}_{\{\alpha<\infty\}} = g(\tau_H^\alpha)\mathbb{1}_{\{\tau_H^\alpha<\infty\}}$, where θ_t is the shift operator defined above. Indeed, taking $\eta^k = \sum_j g(X_{t_j^k})\mathbb{1}_{[t_j^k, t_{j+1}^k)}(\tau_H)$, $t_j^k = j/2^k$ for $k = 1, 2, \ldots$ and $j = 0, 1, 2, \ldots$, we obtain $\theta_t \eta = \lim_{k\to\infty} \theta_t \eta^k = \lim_{k\to\infty} \sum_j g(X_{t_j^k+t})\mathbb{1}_{[t_j^k+t, t_{j+1}^k+t)}(\tau_H^t) = g(\tau_H^t)\mathbb{1}_{\{\tau_H^t<\infty\}}$, because, $\theta_t \mathbb{1}_{[t_j^k, t_{j+1}^k)}(\tau_H) = \mathbb{1}_{[t_j^k+t, t_{j+1}^k+t)}(\tau_H^t)$. In particular, if $\alpha = \tau_G$ for a measurable set $G \subset H$ and $\tau_H < \infty$, Q^x a.s., then we have $\tau_H^\alpha = \tau_H$ (see Fig. 1.1). Thus $\theta_{\tau_G} g(X_{\tau_H}) = g(X_{\tau_H})$.

Therefore, for a measurable bounded function f, we obtain

$$E^x[f(X_{\tau_H})] = E^x[E^{X_{\tau_G}}[f(X_{\tau_H})]] = \int_{\partial G} E^y[f(X_{\tau_H})] \cdot Q^x[X_{\tau_G} \in \mathrm{d}y] \quad (13.1)$$

for $x \in G$. In other words, the expected value of f at X_{τ_H} when starting at $x \in G$ can be obtained by integrating the expected value when starting at $y \in \partial G$ with respect to the "hitting distribution" μ_G^x of X on ∂G defined by $\mu_G^x(F) = Q^x(\{X_{\tau_G} \in F\})$ for $F \subset \partial G$ and $x \in G$. The measure μ_G^x is called the harmonic measure of X on ∂G. The formula (13.1) can be written in the following form:

$$E^x[f(X_{\tau_H})] = \int_{\partial G} E^y[f(X_{\tau_H})]\mathrm{d}\mu_G^x(y).$$

The above formula defines the mean property of an Itô diffusion X for $x \in G$ and Borel sets $G \subset H$.

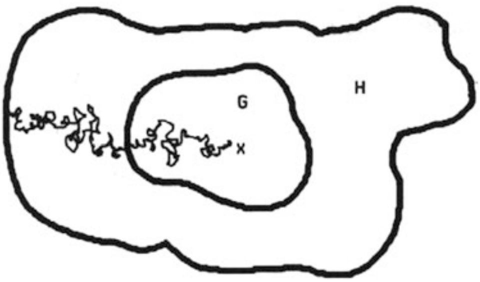

Fig. 1.1 The mean property of Itô diffusion

14 Notes and Remarks

The definitions and most results of this chapter are classical. They are selected from Friedman [31], Ikeda and Watanabe [43], Karatzas and Shreve [47], Øksendal [86], Protter [89], Jacod and Shiryaev [44], and Gihman and Skorohod [33]. In particular, the definition and the classical properties of stopping times are taken from Protter [89], whereas the properties of sequences of the first exit times of continuous processes contained in Sect. 5 of Chap. IV come from Kisielewicz [61]. Section 2 is entirely based on Ikeda and Watanabe [43]. The definitions and results dealing with stochastic processes contained in Sect. 3 are selected from Ikeda and Watanabe [43], Øksendal [86], Protter [89], and Billingsley [21]. In particular, Theorem 3.6 presented in this section is a slight author's generalization of Theorem 2.12.1 of Billingsley given in [21]. The properties of Poisson process and Brownian motion are taken from Protter [89], Friedman [31], and Lipcer and Shiryaev [72]. Stochastic integrals are mainly based on Friedman [31] and Øksendal [86]. The Itô formula is taken from Øksendal [86], and the martingale representation theorems from Karatzas and Shreve [47]. The theory of stochastic differential equations and diffusion processes is based on Ikeda and Watanabe [43], Karatzas and Shreve [47], Øksendal [86], and Stroock and Varadhan [94, 95]. In particular, the uniqueness of in law weak solutions of stochastic differential equations is based on Stroock and Varadhan [95]. Properties of diffusion processes described by nonautonomous stochastic differential equations are taken from Kisielewicz [60, 61]. The last two sections are entirely taken from Øksendal [86]. The proof of Remark 1.2 can be found in Jacod and Shiryaev [44]. Example 1.1 and Fig. 1.1 are also taken from Øksendal [86]. The proof of Theorem 1.5 can be found in Ikeda and Watanabe [43] and Gihman and Skorohod [33]. The proofs of Remark 3.1 and Theorem 3.3 are given in Jacod and Shiryaev [44], and Theorem 3.4 in Gihman and Skorohod [33] and Jacod and Shiryaev [44], where proofs of Theorems 4.4 and 4.5 can also be found. The proof of Remark 4.1 is taken from Øksendal [86]. Remarks 4.2 and 4.3 come from Proter [89]. The complete proof of Theorem 5.2 can be found in Protter [89], whereas Remark 5.7 is proved in Øksendal [86]. Theorems 5.6, 5.8, and 5.9 are proved in Friedman [31], Gihman and Skorohod [33], Jacod and Shiryaev [44], and Karatzas and Shreve [47]. Proofs of Remark 5.9 and Theorems 8.1 and 8.2 can be found in Øksendal [86], Kunita [68], Gihman and Skorohod [33], and Karatzas and Shreve [47]. Finally, proofs of Remark 9.3, Theorem 9.2, and Corollary 12.2 can be found in Ikeda and Watanabe [43], Strook and Varadhan [95], and Friedman [31].

Chapter 2
Set-Valued Stochastic Processes

This chapter is devoted to basic notions of the theory of set-valued mappings and set-valued stochastic processes. We begin with the notions and basic properties of the space of subsets of a given metric space. Selected properties of set-valued mappings, Aumann integrals, and set-valued stochastic processes are presented. The last two parts of this chapter discuss properties of a set-valued conditional expectation and selection properties of set-valued integrals depending on random parameters.

1 Spaces of Subsets of a Metric Space

Let (X, ρ) be a metric space and $(A_n)_{n=1}^{\infty}$ a sequence of subsets of X. The sets $\bigcap_{n=1}^{\infty} \bigcup_{k=0}^{\infty} A_{n+k}$ and $\bigcup_{n=1}^{\infty} \bigcap_{k=0}^{\infty} A_{n+k}$ are denoted by $\mathrm{Lim\,sup}\, A_n$ and $\mathrm{Lim\,inf}\, A_n$, respectively and said to be a limit superior and a limit inferior, respectively of a sequence $(A_n)_{n=1}^{\infty}$. Immediately from the above definitions, the following properties of $\mathrm{Lim\,sup}\, A_n$ and $\mathrm{Lim\,inf}\, A_n$ follow.

Lemma 1.1. *Let $(A_n)_{n=1}^{\infty}$ and $(B_n)_{n=1}^{\infty}$ be sequences of subsets of X and let $C \subset X$. Then*

(i) $\mathrm{Lim\,inf}\, A_n = (\mathrm{Lim\,sup}\, A_n^{\sim})^{\sim}$, *where* $D^{\sim} = X \setminus D$ *for* $D \subset X$,
(ii) $\mathrm{Lim\,inf}(A_n \cap B_n) = \mathrm{Lim\,inf}\, A_n \cap \mathrm{Lim\,inf}\, B_n$,
(iii) $\mathrm{Lim\,inf}(A_n \cap C) = (\mathrm{Lim\,inf}\, A_n) \cap C$,
(iv) $\bigcap_{n=1}^{\infty} A_n \subset \mathrm{Lim\,inf}\, A_n \subset \mathrm{Lim\,sup}\, A_n \subset \bigcup_{n=1}^{\infty} A_n$.

Corollary 1.1. *For every family $\{A_n^i : i, n = 1, 2, \ldots\}$ of subsets of X, one has* $\bigcap_{i=1}^{\infty}[\mathrm{Lim\,inf}\ A_n^i) = \mathrm{Lim\,inf}[\bigcap_{i=1}^{\infty} A_n^i]$. □

Apart from the limits $\mathrm{Lim\,sup}\, A_n$ and $\mathrm{Lim\,inf}\, A_n$, we can also define the Kuratowski limits $\mathrm{Li}\, A_n$ and $\mathrm{Ls}\, A_n$. The set $\mathrm{Li}\, A_n$ is defined by the property $x \in \mathrm{Li}\, A_n$ if and only if for every neighborhood $\mathcal{U}$ of x, there is an integer $N \geq 1$

M. Kisielewicz, *Stochastic Differential Inclusions and Applications*,
Springer Optimization and Its Applications 80, DOI 10.1007/978-1-4614-6756-4_2,

such that $\mathcal{U} \cap A_n \neq \emptyset$ for every $n \geq N$. It is said to be the Kuratowski limit inferior of a sequence $(A_n)_{n=1}^{\infty}$. Similarly, the Kuratowski limit superior $\mathrm{Ls}\, A_n$ of a sequence $(A_n)_{n=1}^{\infty}$ is defined by the property: $x \in \mathrm{Ls}\, A_n$ if and only if for every neighborhood $\mathcal{U}$ of x, there are infinitely many n with $\mathcal{U} \cap A_n \neq \emptyset$.

Corollary 1.2. *For every sequence $(A_n)_{n=1}^{\infty}$ of subsets of X, one has*

(i) $\mathrm{Li}\, A_n \subset \mathrm{Ls}\, A_n$,
(ii) $x \in \mathrm{Li}\, A_n$ *if and only if there exist an integer $N \geq 1$ and a sequence $(x_n)_{n=1}^{\infty}$ of X with $x_n \in A_n$ for $n \geq N$ such that $x = \lim_{n\to\infty} x_n$,*
(iii) $x \in \mathrm{Ls}\, A_n$ *if and only if there exist an increasing subsequence $(n_k)_{k=1}^{\infty}$ of $(n)_{n=1}^{\infty}$ and a sequence $(x_{n_k})_{k=1}^{\infty}$ of X such that $x_{n_k} \in A_{n_k}$ for $k = 1, 2, \ldots$ and $x = \lim_{k\to\infty} x_{n_k}$.* □

The following properties of the Kuratowski limits follow immediately from the above definitions.

Lemma 1.2. *Let $(A_n)_{n=1}^{\infty}$ and $(B_n)_{n=1}^{\infty}$ be sequences of subsets of X. Then*

(i) if $A_n \subset B_n$ for $n \geq 1$, then $\mathrm{Li}\, A_n \subset \mathrm{Li}\, B_n$ *and* $\mathrm{Ls}\, A_n \subset \mathrm{Ls}\, B_n$,
(ii) $\mathrm{Lim\,inf}\, A_n \subset \mathrm{Li}\, A_n$,
(iii) $\mathrm{Li}(A_n \cap B_n) \subset (\mathrm{Li}\, A_n) \cap (\mathrm{Li}\, B_n)$,
(iv) $\mathrm{Ls}(A_n \cap B_n) \subset (\mathrm{Ls}\, A_n) \cap (\mathrm{Ls}\, B_n)$,
(v) $\mathrm{Ls}\, A_n = \bigcap_{n=1}^{\infty} \overline{\bigcup_{k=0}^{\infty} A_{k+n}}$,
(vi) if $A_n = A$ for $n \geq 1$, then $\mathrm{Li}\, A_n = \bar{A} = \mathrm{Ls}\, A_n$.

Let $\mathrm{Cl}(X)$ denote the family of all nonempty closed subsets of X. For every $A, B \in \mathrm{Cl}(X)$, we define the Hausdorff distance $h(A, B)$ with respect to the metric ρ on X by setting $h(A, B) = \inf\{\varepsilon : A \subset V_\varepsilon(B) \text{ and } B \subset V_\varepsilon(A)\}$, where $V_\varepsilon(C)$ denotes the ε-neighborhood of $C \in \mathrm{Cl}(X)$, i.e., $V_\varepsilon(C) = \{x \in X : \mathrm{dist}(x, C) \leq \varepsilon\}$.

Lemma 1.3. *The function $h : \mathrm{Cl}(X) \times \mathrm{Cl}(X) \to [0, \infty]$ has the following properties:*

(i) $h(A, B) = 0$ *if and only if* $A = B$ *for* $A, B \in \mathrm{Cl}(X)$,
(ii) $h(A, B) = h(B, A)$ *for every* $A, B \in \mathrm{Cl}(X)$,
(iii) $h(A, B) \leq h(A, C) + h(C, B)$ *for every* $A, B, C \in \mathrm{Cl}(X)$.

Proof. To prove (i), let us observe that $h(A, B) = \max\{\bar{h}(A, B), \bar{h}(B, A)\}$, where $\bar{h}(C, D) = \sup_{x\in C} \mathrm{dist}(x, D)$ for $C, D \in \mathrm{Cl}(X)$. Hence it follows that $h(A, B) = 0$ implies that $A \subset B$ and $B \subset A$, because $A, B \in \mathrm{Cl}(X)$. Then $A = B$. Statement (ii) is evident. To prove (iii), if $A \subset V_\varepsilon(C)$ and $C \subset V_\eta(B)$, then $A \subset V_{\varepsilon+\eta}(B)$. Consequently, we get $\bar{h}(A, B) \leq \bar{h}(A, C) + \bar{h}(C, B)$. Thus $h(A, B) = \max\{\bar{h}(A, B), \bar{h}(B, A)\} \leq \max\{\bar{h}(A, C) + \bar{h}(C, B), \bar{h}(B, C) + \bar{h}(C, A)\} \leq \max\{h(A, C) + h(C, B), h(B, C) + h(C, A)\} = h(A, C) + h(C, B)$. □

Theorem 1.1. *Let* (X, ρ) *be a compact metric space. Then* $(\mathrm{Cl}(X), h)$ *is a compact metric space, too. In such a case, a sequence* $(A_n)_{n=1}^{\infty}$ *of* $\mathrm{Cl}(X)$ *converges to* $A \in \mathrm{Cl}(X)$ *in the* h*-metric topology if and only if* $\mathrm{Li}\, A_n = A = \mathrm{Ls}\, A_n$.

Proof. By virtue of Lemma 1.3, the mapping h is a metric on $\mathrm{Cl}(X)$. The proof of compactness of $(\mathrm{Cl}(X), h)$ can be found in [49]. If a sequence $(A_n)_{n=1}^{\infty}$ of $\mathrm{Cl}(X)$ converges to $A \in \mathrm{Cl}(X)$ in the h-metric topology, then by the definitions of the metric h and the Kuratowski limits $\mathrm{Li}\, A_n$ and $\mathrm{Ls}\, A_n$, we get $A \subset \mathrm{Li}\, A_n$ and $\mathrm{Ls}\, A_n \subset A$. Then $\mathrm{Li}\, A_n = A = \mathrm{Ls}\, A_n$. Conversely, let $A \subset X$ be such that $\mathrm{Li}\, A_n = A = \mathrm{Ls}\, A_n$. By the compactness of the metric space (X, ρ), we have $A \neq \emptyset$. Then $A \in \mathrm{Cl}(X)$. We have to show that for every $\varepsilon > 0$ and sufficiently large $n \geq 1$, one has $A_n \subset V_\varepsilon(A)$ and $A \subset V_\varepsilon(A_n)$. If the first inclusion were false, we would obtain a contradiction to $A = \mathrm{Ls}\, A_n$. If the second inclusion were false, we would obtain a contradiction to $\mathrm{Li}\, A_n = A$. □

Remark 1.1. The above results can be extended to the case of a locally compact separable metric space (X, ρ), because it possesses a one-point compactification, denoted by $X \cup \{\infty\}$. □

We can extend the definition of Hausdorff distance on the family $\mathcal{P}_b(X)$ of all nonempty bounded subsets of a metric space (X, ρ). Similarly as above, for every $A, B \in \mathcal{P}_b(X)$, we define $\bar{h}(A, B) = \inf\{\varepsilon > 0 : A \subset V_\varepsilon\}$, and then the Hausdorff pseudometric h on $\mathcal{P}_b(X)$ is defined by $h(A, B) = \max\{\bar{h}(A, B), \bar{h}(B, A)\}$ for every $A, B \in \mathcal{P}_b(X)$. It can be verified that $h(A, B) = 0$ if and only if $\bar{A} = \bar{B}$.

Corollary 1.3. *For every* $A, B \in \mathcal{P}_b(X)$*, one has* $\bar{h}(A, B) = \sup\{\mathrm{dist}(a, B) : a \in A\}$*, where* $\mathrm{dist}(a, B) = \inf\{\rho(a, b) : b \in B\}$.

Proof. For every $A, B \in \mathcal{P}_b(X)$, we have $A \subset V_\varepsilon(B)$ if for every $a \in A$, we have $\mathrm{dist}(a, B) \leq \varepsilon$. Then $A \subset V_\varepsilon(B)$ implies $\bar{h}(A, B) \leq \varepsilon$. Similarly, we can verify that $\bar{h}(A, B) \leq \varepsilon$ implies $A \subset V_\varepsilon(B)$. Hence it follows that $\inf\{\varepsilon > 0 : A \subset V_\varepsilon(B)\} = \inf\{\varepsilon > 0 : \varepsilon \geq \bar{h}(A, B)\} = \bar{h}(A, B)$. □

Lemma 1.4. *For every* $A, B \in \mathcal{P}_b(X)$*, one has* $h(\bar{A}, \bar{B}) \leq h(A, B)$.

Proof. For every $a \in \bar{A}$ and $\varepsilon > 0$, there is $a_\varepsilon \in A$ such that $\rho(a, a_\varepsilon) \leq \varepsilon$. Therefore, $\mathrm{dist}(a, \bar{B}) \leq \rho(a, a_\varepsilon) + \mathrm{dist}(a_\varepsilon, \bar{B}) \leq \varepsilon + \inf\{\rho(a_\varepsilon, b) : b \in \bar{B}\} \leq \varepsilon + \inf\{\rho(a_\varepsilon, b) : b \in B\} \leq \varepsilon + \bar{h}(A, B)$. Thus $\sup\{\mathrm{dist}(a, \bar{B}) : a \in \bar{A}\} \leq \varepsilon + \bar{h}(A, B)$, i.e., $\bar{h}(\bar{A}, \bar{B}) \leq \varepsilon + \bar{h}(A, B)$ for every $\varepsilon > 0$. Then $\bar{h}(\bar{A}, \bar{B}) \leq \bar{h}(A, B)$. Similarly, we get $\bar{h}(\bar{B}, \bar{A}) \leq \bar{h}(B, A)$. □

Remark 1.2. It can be verified that for every $A, B \in \mathcal{P}_b(X)$, one has $h(\bar{A}, \bar{B}) = h(A, B)$. □

If X is a linear normed space and $A, B \in \mathcal{P}_b(X)$, then we define $A + B = \{x \in X : x = a + b,\ a \in A,\ b \in B\}$. Similarly, for $A \in \mathcal{P}_b(X)$ and $\mu \in \mathbb{R}$, we define $\mu \cdot A = \{x \in X : x = \mu a,\ a \in A\}$. Immediately from the last definition, it follows that we can define a set $A + (-1)B$, which is often called the Minkowski difference of sets $A, B \in \mathcal{P}_b(X)$. In the general case, we have $A + (-1)A \neq \{0\}$.

For some nonempty compact convex sets $A, B \subset X$, a difference $A - B$, known as a Hukuhara difference, can be defined such that $A - A = \{0\}$. It is easy to verify that for all compact convex sets $A, B \in \mathcal{P}_b$ and $\lambda, \mu \in \mathbb{R}^+$, one has (i) $A + \{0\} = \{0\} + A = A$, (ii) $(A + B) + C = A + (B + C)$ (iii) $A + B = B + A$, (iv) $A + C = B + C$ implies $A = B$, (v) $1 \cdot A = A$, (vi) $\lambda \cdot (A + B) = \lambda \cdot A + \lambda \cdot B$, and (vii) $(\lambda + \mu) \cdot A = \lambda \cdot A + \mu \cdot A$.

Lemma 1.5. *Let* $(X, || \cdot ||)$ *be a linear normed space. For every* $A, B, C, D \in \mathcal{P}_b(X)$ *and* $\mu \in \mathbb{R}^+$, *one has (i)* $\bar{h}(\mu A, \mu B) = \mu \bar{h}(A, B)$ *and (ii)* $\bar{h}(A + B, C + D) \leq \bar{h}(A, C) + \bar{h}(B, D)$.

Proof. (i) If $A \subset V_\varepsilon(B)$, then $\mu A \subset V_{\mu\varepsilon}(\mu B)$. Hence it follows that $\inf\{\eta > 0 : \mu A \subset V_\eta(\mu B)\} = \mu \inf\{\eta > 0 : A \subset V_\eta(B)\} = \mu \bar{h}(A, B)$. (ii) If $A \subset V_\varepsilon(C)$ and $B \subset V_\eta(D)$, then $A + B \subset V_{\varepsilon+\eta}(C + D)$. Therefore, $\inf\{\varepsilon + \eta : A + B \subset V_{\varepsilon+\eta}(C + D)\} \leq \inf\{\varepsilon : A \subset V_\varepsilon(C)\} + \inf\{\eta : B \subset V_\eta(D)\} = \bar{h}(A, C) + \bar{h}(B, D)$. □

Corollary 1.4. *For every* $\mu \in [0, 1]$ *and* $A, B, C, D \in \mathcal{P}_b(X)$, *one has* $\bar{h}(\mu A + (1 - \mu)B, \mu C + (1 - \mu)D) \leq \mu \bar{h}(A, C) + (1 - \mu)\bar{h}(B, D)$. □

Corollary 1.5. *For every* $A, B, C, D \in \mathcal{P}_b(X)$, *one has* $\bar{h}(\overline{A + B}, \overline{C + D}) \leq \bar{h}(A, C) + \bar{h}(B, D)$. □

Corollary 1.6. *For every* $A, B, C, D \in \mathcal{P}_b(X)$, *one has* $h(\overline{A + B}, \overline{C + D}) \leq h(A, C) + h(B, D)$. □

2 Set-Valued Mappings

Let X and Y be nonempty sets and let $\mathcal{P}(Y)$ denote the family of all nonempty subsets of Y. By a set-valued mapping defined on X with values in $\mathcal{P}(Y)$ we mean a mapping $F : X \to \mathcal{P}(Y)$. It is clear that a set-valued mapping F can be defined as a relation contained in $X \times Y$ with the domain $\mathrm{Dom}(F) = X$. It is defined by its graph: $\mathrm{Graph}(F) = \{(x, y) \in X \times Y : y \in F(x)\}$. In applications, we need set-valued mappings having some special regularities, such as continuity and measurability. To define such set-valued mappings, we have to consider X and $\mathcal{P}(Y)$ as topological or measurable spaces. It can be verified that if $(Y, \mathcal{T})$ is a topological space, then we can define on $\mathcal{P}(Y)$ the upper topology $\mathcal{T}_u$ generated by the base $\mathcal{U} = \{[\cdot, G] : G \in \mathcal{T}\}$, where $[\cdot, G] = \{A \in \mathcal{P}(Y) : A \subset G\}$. Similarly, the lower topology $\mathcal{T}_l$ on $\mathcal{P}(Y)$ is generated by the subbase $\mathcal{L}$ defined by $\mathcal{L} = \{I_G : G \in \mathcal{T}\}$, where $I_G = \{U \in \mathcal{P}(Y) : U \cap G \neq \emptyset\}$. If (Y, d) is a separable metric space, then the Borel σ-algebra of the metric space $(\mathrm{Comp}(Y), h)$ is generated by sets $\{K \in \mathrm{Comp}(Y) : K \cap V \neq \emptyset\}$ for every open set $V \subset Y$, where Comp(Y) $\subset \mathcal{P}(Y)$ contains all compact subsets of Y, and h is the Hausdorff metric on $\mathrm{Comp}(Y)$. These observations lead to the following definitions.

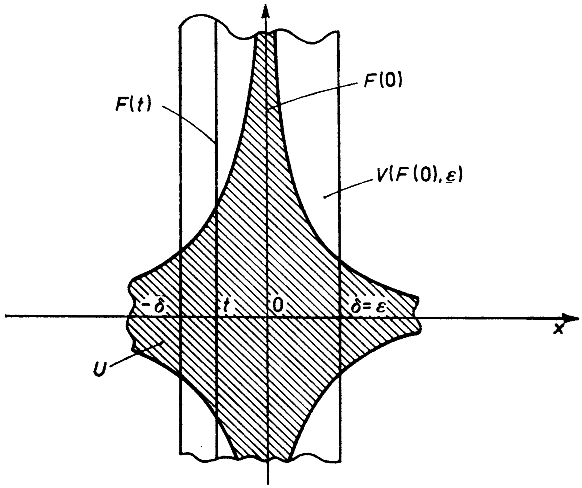

Fig. 2.1 A mapping that is H-u.s.c. but not u.s.c. at $t = 0$

If $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ are given topological spaces, then $F : X \to \mathcal{P}(Y)$ is said to be lower semicontinuous (l.s.c.) at $\bar{x} \in X$ if for every $U \in \mathcal{T}_Y$ with $F(\bar{x}) \cap U \neq \emptyset$, there is $V \in \mathcal{T}_X$ such that $\bar{x} \in V$ and $F(x) \cap U \neq \emptyset$ for every $x \in V$. We call $F : X \to \mathcal{P}(Y)$ upper semicontinuous (u.s.c.) at $\bar{x} \in X$ if for every $U \in \mathcal{T}_Y$ such that $F(\bar{x}) \subset U$, there is $V \in \mathcal{T}_X$ such that $\bar{x} \in V$ and $F(x) \subset U$ for every $x \in V$. If (X, ρ) and (Y, d) are given metric spaces, then a set-valued mapping $F : X \to \mathcal{P}(Y)$ is said to be H-l.s.c. at $\bar{x} \in X$ if for every $\varepsilon > 0$, there exists $\delta > 0$ such that $F(\bar{x}) \subset V(F(x), \varepsilon)$ for every $x \in B(\bar{x}, \delta)$, where $V(F(x), \varepsilon) = \{z \in X : \mathrm{dist}(z, F(x)) \leq \varepsilon\}$ and $B(\bar{x}, \delta)$ is an open ball of X centered at $\bar{x}$ with radius δ. It is clear that if F is H-l.s.c. at $\bar{x} \in X$, then it is also l.s.c. If $F(\bar{x}) \in \mathrm{Comp}(Y)$, then F is H-l.s.c. at $\bar{x} \in X$ if and only if it is l.s.c. at $\bar{x} \in X$. We say that F is l.s.c. (H-l.s.c.) on X if it is l.s.c. (H-l.s.c.) at every point $\bar{x} \in X$. In a similar manner, we can define H-u.s.c. set-valued mappings on X. There are some H-u.s.c. set-valued mappings that are not u.s.c. This is is illustrated in Fig. 2.1, where $F(t) = \{(y, z) \in \mathbb{R}^2 : y = t\}$ for $t \in \mathbb{R}$.

Let us observe that for a given l.s.c. set-valued mapping, we can change its values at finite points in such a way that it remains l.s.c. This follows from the following result.

Remark 2.1. If $F : X \to \mathcal{P}(Y)$ is l.s.c. on X and $(x_0, y_0) \in \mathrm{Graph}(F)$, then the set-valued mapping $G : X \to \mathcal{P}(Y)$ defined by taking $G(x) = F(x)$ for $x \in X \setminus \{x_0\}$ and $G(x) = \{y_0\}$ for $x = x_0$, is also l.s.c. on X.

Proof. It is clear that G is l.s.c. at every point $x \in X \setminus \{x_0\}$. By the lower semicontinuity of F at x_0 and the property of the point (x_0, y_0), for every neighborhood $\mathcal{U}$ of y_0 we have $F(x_0) \cap \mathcal{U} \neq \emptyset$, and there is a neighborhood V of x_0 such that $F(x) \cap \mathcal{U} \neq \emptyset$ for every $x \in V$. Therefore, for every $\mathcal{U} \in \mathcal{T}_Y$

such that $G(x_0) \cap \mathcal{U} \neq \emptyset$, there is $V \in \mathcal{T}_X$, a neighborhood of x_0, such that $G(x) \cap \mathcal{U} \neq \emptyset$ for every $x \in V$. Then G is l.s.c. at x_0. □

A set-valued mapping $F : X \to \mathcal{P}(Y)$ is said to be continuous (H-continuous) on X if it is l.s.c. (H-l.s.c.) and u.s.c. (H-u.s.c.) on X. It can be verified that a multifunction $F : X \to \mathrm{Comp}(Y)$ is continuous if and only if it is H-continuous. If $Y = \mathbb{R}^d$ and $F : X \to \mathrm{Comp}(Y)$ takes convex values, then F is continuous if and only if a function $X \ni x \to \sigma(p, F(x)) \in \mathbb{R}$ is continuous for every $p \in \mathbb{R}^d$, where $\sigma(\cdot, A)$ denotes the support function of $A \subset \mathbb{R}^d$. In optimal control theory, we have to deal with parameterized set-valued functions of the form $F(x) = \{f(x,u) : u \in U\}$, where $f : X \times U \to Y$ is a given function. We shall show that if $f(\cdot, u)$ is continuous, then the multifunction F is l.s.c. Some other properties of such multifunctions are given in Chap. 7.

Lemma 2.1. *Assume that X and Y are topological Hausdorff spaces and let $f : X \times U \to Y$, where $U \neq \emptyset$. If $f(\cdot, u)$ is continuous on X for every $u \in U$, then the set-valued mapping $F : X \to \mathcal{P}(Y)$ defined by $F(x) = f(x, U)$ is l.s.c. on X.*

Proof. Let $\bar{x} \in X$ be fixed and let $\mathcal{N}$ be an open set of Y. Suppose $\bar{u} \in U$ is such that $f(\bar{x}, \bar{u}) \in \mathcal{N}$. By the continuity of $f(\cdot, \bar{u})$ at $\bar{x}$, there is a neighborhood V of $\bar{x}$ such that $f(x, \bar{u}) \in \mathcal{N}$ for every $x \in V$. Therefore, for every $x \in V$, we get $F(x) \cap \mathcal{N} \neq \emptyset$. □

Let $(T, \mathcal{F})$ be a measurable space and (Y, d) a separable metric space. A set-valued mapping $F : T \to \mathcal{P}(Y)$ is said to be measurable (weakly measurable) if for every closed (open) set $E \subset Y$, we have $\{t \in T : F(t) \cap E \neq \emptyset\} \in \mathcal{F}$. It is clear that if F is measurable, then it is weakly measurable. The converse statement is not true in general.

Remark 2.2. Let $(T, \mathcal{F})$ be a measurable space and $(Y, ||\cdot||)$ a separable Banach space. For $F : T \to \mathcal{P}(Y)$, we denote by $\overline{\mathrm{co}}\, F$ the set-valued mapping $\overline{\mathrm{co}}\, F : T \to \mathcal{P}(Y)$ defined by $(\overline{\mathrm{co}}\, F)(t) = \overline{\mathrm{co}}\, F(t)$ for every $t \in T$, where $\overline{\mathrm{co}}\, F(t)$ denotes the closed convex hull of $F(t)$. It is clear that $\overline{\mathrm{co}}\, F$ is measurable whenever F is measurable. □

Remark 2.3. If $(T, \mathcal{F})$ is a measurable space, $Y = \mathbb{R}^d$, and $F : T \to \mathrm{Cl}(Y)$ is measurable, then the function $T \ni t \to \sigma(p, F(t)) \in \mathbb{R}$ is measurable for every $p \in \mathbb{R}^d$. If $F : T \to \mathrm{Cl}(\mathbb{R}^d)$ is convex-valued, then F is measurable if and only if $\sigma(p, F(\cdot))$ is measurable for every $p \in \mathbb{R}^d$. □

Remark 2.4. It can be proved that if X is a metric space, $Y = \mathbb{R}^d$, and $F : X \to \mathrm{Comp}(Y)$ is continuous, then $\sigma(p, F(\cdot))$ is continuous for every $p \in \mathbb{R}^d$. □

It is natural to expect that for a given multifunction $F : X \to \mathcal{P}(Y)$, there exists a function $f : X \to Y$ such that $f(x) \in F(x)$ for $x \in X$. The existence of such a function f, called a selector or a selection for F, follows immediately from Zermelo's axiom of choice. We recall first the Kuratowski–Zorn lemma, and then we will verify how from this principle, the axiom of choice can be deduced.

Lemma (Kuratowski–Zorn lemma). Let P be a nonempty partially ordered set with the property that every completely ordered subset of P has an upper bound in P. Then P contains at least one maximal element.

Lemma (Axiom of choice). Let $\mathcal{E}$ be a nonempty family of nonempty subsets of a set X. Then there exists a function $f : \mathcal{E} \to X$ such that $f(E) \in E$ for each E in $\mathcal{E}$.

Proof. Consider the class P of all functions $p : \mathcal{D}(p) \to X$ such that the domain $\mathcal{D}(p)$ of p belongs to $\mathcal{E}$ and $p(E) \in E$ for each E in $\mathcal{D}(p)$. This is a nonempty class, because $\mathcal{E}$ contains a nonempty set E, and if $x \in E$, the function with domain $\{E\}$ and range $\{x\}$ is a member of P. We order P by the inclusion relation in $\mathcal{E} \times X$. It can be verified that P satisfies the conditions of the Kuratowski–Zorn lemma. Therefore, we infer that there exists a function $f : \mathcal{E} \to X$ such that $f(E) \in E$ for each $E \in \mathcal{E}$. $\square$

Corollary 2.1. *For nonempty sets X and Y, every set-valued mapping $F : X \to \mathcal{P}(Y)$ possesses at least one selector.*

Proof. Let $\mathcal{E} = \{F(x)\}_{x \in X}$. The family $\mathcal{E}$ satisfies the conditions of Zermelo's axiom of choice. Therefore, there exists a function $g : \mathcal{E} \to Y$ such that $g(F(x)) \in F(x)$ for every $x \in X$. Thus the function $f : X \to Y$ defined by $f(x) = g(F(x))$ for $x \in X$ is a selector for F. $\square$

In applications of the theory of set-valued mappings, the existence of special selectors for given multifunctions plays a crucial role. The most difficult part is to deduce the existence of selectors with prescribed properties. In what follows, we shall present some results dealing with the existence of continuous, measurable, and Lipschitz continuous selectors. The fundamental problem deals with the existence of continuous selections. The following example shows that continuous set-valued mappings need not have, in general, continuous selections.

Example 2.1. Let F be the set-valued mapping defined on the interval $(-1, 1)$ by setting

$$F(x) = \begin{cases} \{(v_1, v_2) : v_1 = \cos\theta,\ v_2 = t\sin\theta \text{ and } \frac{1}{t} \le \theta \le \frac{1}{t} + 2\pi - |t|\} & \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{for} \quad t \in (-1, 2) \setminus \{0\}, & \\ \{(v_1, v_2) : -1 \le v_1 \le 1,\ \ v_2 = 0\} \qquad \text{for} \quad t = 0\,. & \end{cases}$$

For $t \neq 0$ and $t \in (-1, 1)$, $F(t)$ is a subset of an ellipse in $\mathbb{R}^2$ (see Fig. 2.2), whose minor axis shrinks to zero as $t \to 0$, so that the ellipse collapses to a segment $F(0)$.

The subset of the ellipse given by $F(t)$ is obtained by removing a section, from the angle $(1/t) - |t|$ to the angle $(1/t)$. As t gets smaller, the arc length of this hole decreases, while the initial angle increases like $(1/t)$, i.e., it spins around the origin with increasing angular velocity. However, F is continuous at the origin, while no selection $f : (-1, 0) \to \mathbb{R}^2$ or $g : (0, 1) \to \mathbb{R}^2$, for example

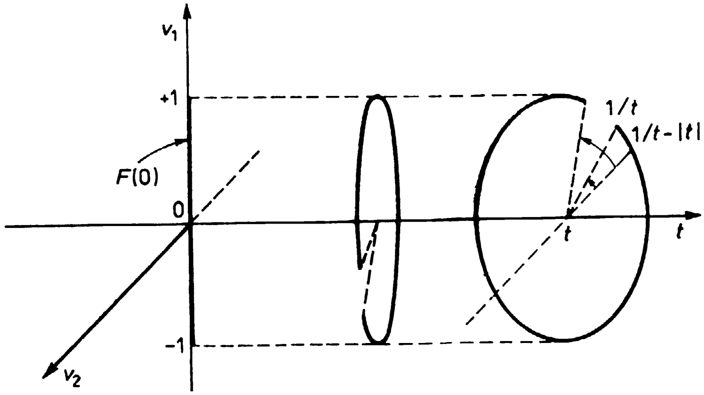

Fig. 2.2 The mapping F

$f(t) = (\cos(1/t), t\sin(1/t))$, can be continuously extended to the whole interval $(-1, 1)$. In fact, the hole in the ellipse would force this selection to rotate around the origin with an angle $\rho(t)$ between $(1/t)$ and $(1/t) + 2\pi - |t|$, and $\lim_{t\to 0} f(t)$ cannot exist.

We shall show that in some special cases, l.s.c. multifunctions possess continuous selections. This follows from the famous Michael continuous selection theorem. We precede it by the following lemmas.

Lemma 2.2. *Let* (X, ρ) *and* $(Y, ||\cdot||)$ *be a metric and a Banach space, respectively, and let* $\Phi : X \to \mathcal{P}(Y)$ *be a convex-valued and l.s.c. multifunction. For every* $\varepsilon > 0$*, there is a continuous function* $\varphi : X \to Y$ *such that* $\mathrm{dist}(\varphi(x), \Phi(x)) \le \varepsilon$ *for* $x \in X$.

Proof. Let $x \in X$ be fixed and select $y_x \in \Phi(x)$ and $\delta_x > 0$ such that $(y_x + \varepsilon K_0) \cap \Phi(x') \neq \emptyset$ for every $x' \in B_x$, where $B_x = B(x, \delta_x)$ denotes the open ball of X centered at x with radius $\delta_x > 0$, and K_0 is the unit open ball of Y centered at $0 \in Y$. Since X is paracompact, there exists a locally finite refinement $\{U_z\}_{z\in\Lambda}$ of $\{B_z\}_{z\in X}$. Let $\{p_x\}_{x\in\Lambda}$ be a partition of unity subordinated to it and define a function $\varphi : X \to Y$ by setting $\varphi(x) = \sum_{z\in\Lambda} p_z(x) y_z$ for $x \in X$. It is clear that φ is a continuous function on X. Furthermore, we have $x \in U_z \subset B_z$ whenever $p_z(x) > 0$. Hence it follows that $y_z \in \Phi(x) + \varepsilon K_0$. Since this set is convex, every convex combination of such y_z, in particular $\varphi(x)$, belongs to it, too. Therefore, $\mathrm{dist}(\varphi(x), \Phi(x)) \le \varepsilon$ for $x \in X$. □

Lemma 2.3. *Let* (X, d) *and* (Y, ρ) *be metric spaces, let* $G : X \to \mathcal{P}(Y)$ *be l.s.c., and let* $g : X \to Y$ *be continuous on* X*. If a real-valued function* $X \ni x \to \varepsilon(x) \in \mathbb{R}^+$ *is l.s.c. on* X*, then the set-valued mapping* $\Phi : X \to \mathcal{P}(Y)$ *defined by* $\Phi(x) = B(g(x), \varepsilon(x)) \cap G(x)$ *is l.s.c. at every* $x \in X$ *such that* $\Phi(x) \neq \emptyset$.

Proof. Let $\bar{x} \in X$ be such that $\Phi(\bar{x}) \neq \emptyset$. Select $\bar{y} \in \Phi(\bar{x})$ and let $\eta > 0$. Assume $\varepsilon(\bar{x}) > \rho(\bar{y}, g(\bar{x}))$ and let $\sigma > 0$ be such that $\rho(\bar{y}, g(\bar{x})) = \varepsilon(\bar{x}) - \sigma$. There exists $\sigma_1 > 0$ such that to every $x \in X$ with $d(x, \bar{x}) < \sigma_1$ we can associate $y_x \in G(x)$ such that $\rho(y_x, \bar{y}) < \min(\eta, (1/3)\sigma)$, $\sigma_2 > 0$ such that $d(x, \bar{x}) < \sigma_2$ implies $\varepsilon(x) > \varepsilon(\bar{x}) - (1/3)\sigma$, and $\sigma_3 > 0$ such that $d(x, \bar{x}) < \sigma_3$ implies $\rho(g(\bar{x}), g(x)) < (1/3)\sigma$. Thus

$$\begin{aligned}\rho(y_x, g(x)) &\leq \rho(y_x, \bar{y}) + \rho(\bar{y}, g(\bar{x})) + \rho(g(\bar{x}), g(x)) \\ &< (1/3)\sigma + \varepsilon(\bar{x}) - \sigma + (1/3)\sigma = \varepsilon(\bar{x}) - (1/3)\sigma < \varepsilon(x),\end{aligned}$$

whenever $d(x, \bar{x}) < \min\{\sigma_1, \sigma_2, \sigma_3\}$. Then $y_x \in \Phi(x)$ and $\rho(y_x, y) < \eta$. □

Now we can prove Michael's continuous selection theorem.

Theorem 2.1 (Michael). *Let (X, ρ) and $(Y, |\cdot|)$ be a metric and a Banach space, respectively, and let $F : X \to \mathcal{P}(Y)$ be l.s.c. with closed convex values. Then there exists a continuous function $f : X \to Y$ such that $f(x) \in F(x)$ for $x \in X$.*

Proof. By virtue of Lemma 2.2, for $\varepsilon_1 = 1/2$ and $\Phi = F$, there exists a continuous function $f_1 : X \to Y$ such that $\text{dist}(f_1(x), F(x)) \leq \varepsilon_1$ for $x \in X$. Let $\Phi_1(x) = (f_1(x) + \varepsilon_1 K_0) \cap F(x)$ for $x \in X$. We have $\Phi_1(x) \neq \emptyset$ for $x \in X$. By Lemma 2.3, the multifunction Φ_1 is l.s.c. Then by Lemma 2.2, for $\varepsilon_2 = (1/2)^2$, there exists a continuous function $f_2 : X \to Y$ such that $\text{dist}(f_2(x), \Phi_1(x)) \leq \varepsilon_2$ for $x \in X$. Thus $\text{dist}(f_2(x), F(x)) \leq \varepsilon_2$ and $\text{dist}(f_2(x), (f_1(x) + \varepsilon_1 K_0)) \leq \varepsilon_2$, i.e., $f_2(x) - f_1(x) \in (\varepsilon_1 + \varepsilon_2)K_0$ for $x \in X$. Continuing the above procedure, we can deduce that for every $\varepsilon_n = (1/2)^n$ with $n = 0, 1, 2, \ldots$, there exists a continuous function $f_n : X \to Y$ such that $\text{dist}(f_n(x), F(x)) \leq \varepsilon_n$ and $f_n(x) - f_{n-1}(x) \in (\varepsilon_{n-1} + \varepsilon_n)K_0$ for $x \in X$. Hence in particular, it follows that $\sup_{x \in X} \|f_n(x) - f_{n-1}(x)\| \leq \varepsilon_{n-1} + \varepsilon_n$ for $n \geq 1$, which implies that $(f_n)_{n=1}^{\infty}$ is a Cauchy sequence in the Banach space $C(X, Y)$ of all continuous bounded functions $g : X \to Y$ with the supremum norm. Thus there exists a continuous function $f : X \to Y$ such that $\sup_{x \in X} \|f_n(x) - f(x)\| \to 0$ as $n \to \infty$. Hence it follows that $f(x) \in F(x)$ for $x \in X$, because $F(x)$ is a closed subset of Y and $\text{dist}(f_n(x), F(x)) \leq \varepsilon_n$ for $x \in X$ and $n = 1, 2, \ldots$. □

Remark 2.5. There are closed convex-valued u.s.c. multifunctions that do not possess continuous selections. A simple example is the set-valued mapping F defined by the formula

$$F(x) = \begin{cases} \{-1\} & \text{for} \quad x < 0, \\ [-1, 1] & \text{for} \quad x = 0, \\ \{+1\} & \text{for} \quad x > 0, \end{cases}$$

with the graph presented in Fig. 2.3. □

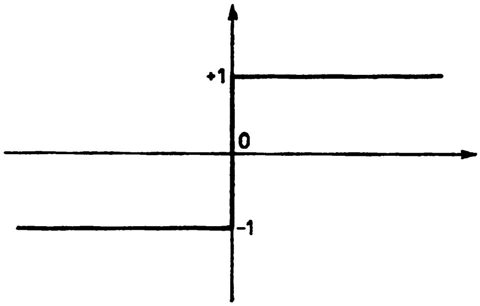

Fig. 2.3 The mapping F

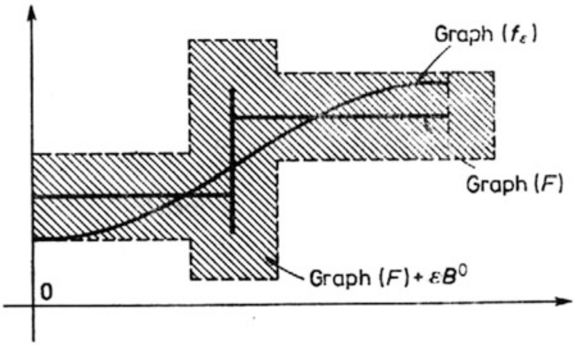

Fig. 2.4 Approximation continuous selection of u.s.c mutifunction

It can be proved that the above set-valued mapping possesses an approximation continuous selection of u.s.c mutifunction

Immediately from Michael's continuous selection theorem we obtain the existence of continuous approximation selections for some special multifunctions. The proof of such a theorem is based on the following lemma.

Lemma 2.4. *Let* (X, ρ), $(Y, |\cdot|)$ *and* $(Z, \|\cdot\|)$ *be Polish and Banach spaces, respectively. If* $\lambda : X \times Y \to Z$ *and* $v : X \to Z$ *are continuous and* $F : X \to \mathcal{P}(Y)$ *is l.s.c. such that* $v(x) \in \lambda(\{x\} \times F(x))$ *for* $x \in X$, *then for every l.s.c. function* $\varepsilon : X \to (0, \infty)$, *the set-valued mapping* $\Phi : X \to \mathcal{P}(Y)$ *defined by* $\Phi(x) = F(x) \cap \{u \in Y : \|\lambda(x, u) - v(x)\| < \varepsilon(x)\}$ *for* $x \in X$ *is l.s.c. on* X.

Proof. Let $\bar{x} \in X$. For every open set $\mathcal{U} \subset Y$ such that $\mathcal{U} \cap \Phi(\bar{x}) \neq \emptyset$, there are $\bar{u} \in \Phi(\bar{x})$ and $\eta > 0$ such that $(\bar{x}, \bar{u}) \in \text{Graph}(\Phi)$ and $(\bar{u} + \eta K_0) \subset \mathcal{U}$, where K_0 is the unit ball of Y. There is $\sigma > 0$ such that $\|\lambda(\bar{x}, \bar{u}) - v(\bar{x}))\| = \varepsilon(\bar{x}) - \sigma$. Let $\delta > 0$ be such that $\|\lambda(x, u) - \lambda(\bar{x}, \bar{u})\| < (1/3)\sigma$ for every $(x, u) \in X \times Y$ satisfying $\max\{\rho(x, \bar{x}), |u - \bar{u}|\} < \delta$. By the lower semicontinuity of F, there is $\sigma_1 > 0$ such that for every $x \in X$ satisfying $\rho(x, \bar{x}) < \sigma_1$, there exists $y_x \in F(x)$ such that $|y_x - \bar{u}| < \min\{\eta, (1/3)\sigma, \delta\}$. By the continuity of v, there is $\sigma_2 > 0$ such that $\|v(x) - v(\bar{x})\| < (1/3)\sigma$ for $x \in X$ satisfying $\rho(x, \bar{x}) < \sigma_2$. Furthermore, by the lower semicontinuity of ε, there is $\sigma_3 > 0$ such that $\varepsilon(x) > \varepsilon(\bar{x}) - (1/3)\sigma$

for every $x \in X$ satisfying $\rho(x,\bar{x}) < \sigma_3$. Then for every $x \in X$ satisfying $\rho(x,\bar{x}) < \min\{\delta,\sigma_1,\sigma_2,\sigma_3\}$, we get

$$\begin{aligned}\|\lambda(x,y_x) - v(x)\| &\le \|\lambda(x,y_x) - \lambda(\bar{x},\bar{u})\| \\ &\quad + \|\lambda(\bar{x},\bar{u}) - v(\bar{x})\| + \|v(\bar{x}) - v(x)\| \\ &< (1/3)\sigma + \varepsilon(\bar{x}) - \sigma + (1/3)\sigma < \varepsilon(x)\,.\end{aligned}$$

Thus $y_x \in \Phi(x)$ and $\|y_x - \bar{u}\| < \eta$. For $\bar{u} \in \Phi(\bar{x})$ and $\eta > 0$ chosen above, we can choose $\bar{\varepsilon} = \min\{\delta,\sigma_1,\sigma_2,\sigma_3\}$ such that $(\bar{u} + \eta K_0) \cap \Phi(x) \neq \emptyset$ for every $x \in B(\bar{x},\bar{\varepsilon})$. Therefore, for every open set $\mathcal{U} \subset Y$ such that $\mathcal{U} \cap \Phi(\bar{x}) \neq \emptyset$, there is $\bar{\varepsilon} > 0$ such that $(\bar{u} + \eta K_0) \cap \Phi(x) \neq \emptyset$ and $(\bar{u} + \eta K_0) \cap \Phi(x) \subset \mathcal{U} \cap \Phi(x)$ for every $x \in B(\bar{x},\bar{\varepsilon})$. □

Theorem 2.2. *Let (X,ρ), $(Y,|\cdot|)$ and $(Z,\|\cdot\|)$ be Polish and Banach spaces, respectively. Assume that $\lambda : X \times Y \to Z$ and $v : X \to Z$ are continuous and $F : X \to \mathcal{P}(Y)$ is l.s.c. with closed convex values. If $\lambda(x,\cdot)$ is affine and $v(x) \in \lambda(x,F(x))$ for $x \in X$, then for every $\varepsilon > 0$, there exists a continuous function $f_\varepsilon : X \to Y$ such that $f_\varepsilon(x) \in F(x)$ and $\|\lambda(x,f_\varepsilon(x)) - v(x)\| \le \varepsilon$ for $x \in X$.*

Proof. By virtue of Lemma 2.4, for every $\varepsilon > 0$, the set-valued mapping $\Phi_\varepsilon : X \to \mathcal{P}(Y)$ defined by $\Phi_\varepsilon(x) = F(x) \cap \{u \in Y : \|\lambda(x,u) - v(x)\| < \varepsilon\}$ for $x \in X$ is l.s.c. on X. Therefore, $\mathrm{cl}(\Phi_\varepsilon)$ is also l.s.c. on X. By the convexity of $F(x)$ and the property of $\lambda(x,\cdot)$ for fixed $x \in X$, it follows that $\Phi_\varepsilon(x)$ and $\mathrm{cl}(\Phi_\varepsilon)(x)$ are convex for $x \in X$. Therefore, by Michael's theorem, for every $\varepsilon > 0$, there exists a continuous selector f_ε for $\mathrm{cl}(\Phi_\varepsilon)$. It is clear that f_ε is a selector of F and satisfies $\|\lambda(x,f_\varepsilon(x)) - v(x)\| < \varepsilon$ for $x \in X$. □

Now we consider the problem of the existence of more regular selections of multifunctions. Such selections are connected with special properties of the "Steiner point map" $s : \mathrm{Conv}(\mathbb{R}^d) \to \mathbb{R}^d$ defined by

$$s(A) = \begin{cases} (d/2)\,[\sigma(1,A) + \sigma(-1,A)] & \text{for } d = 1, \\ d\int_{\Gamma_1} y\,\sigma(y,A)dr & \text{for } d \ge 1, \end{cases} \tag{2.1}$$

for $A \in \mathrm{Conv}(\mathbb{R}^d)$, where Γ_1 is the boundary of an open unit ball of $\mathbb{R}^d$ and dr denotes a differential of the surface measure r on Γ_1 proportional to the Lebesgue measure such that $r(\Gamma_1) = 1$. As usual, $\sigma(\cdot,A)$ denotes the support functions of $A \in \mathrm{Conv}(\mathbb{R}^d)$, and $\mathrm{Conv}(\mathbb{R}^d)$ is the family of all nonempty convex compact subsets of $\mathbb{R}^d$.

Immediately from the above definition, it follows that (i) $s(\{x\}) = x$ for every $x \in \mathbb{R}^d$. Furthermore, (ii) $s(A+B) = s(A) + s(B)$ and (iii) $s(\lambda A) = \lambda s(A)$ for $A,B \in \mathrm{Conv}(\mathbb{R}^d)$ and $\lambda \in \mathbb{R}$. Indeed, for every $A,B \in \mathrm{Conv}(\mathbb{R}^d)$, one obtains

$$\begin{aligned} s(A+B) &= d\int_{\Gamma_1} \sigma(y, A+B)y\,\mathrm{d}r \\ &= d\int_{\Gamma_1} \sigma(y,A)\,\mathrm{d}r + d\int_{\Gamma_1} \sigma(y,B)\,\mathrm{d}r \\ &= s(A)+s(B). \end{aligned}$$

Quite similarly, we also get $s(\lambda A) = \lambda s(A)$ for $\lambda \in \mathbb{R}$ and $A \in \mathrm{Conv}(\mathbb{R}^d)$. Then conditions (ii) and (iii) are also satisfied.

We shall show that for every $A \in \mathrm{Conv}(\mathbb{R}^d)$, one has $s(A) \in A$. To prove this, let us recall some properties of the group $\mathcal{O}(\mathbb{R}^d)$ of all orthogonal linear transformations on $\mathbb{R}^d$. It can be verified that $s(l[A]) = l[s(A)]$ for every $l \in \mathcal{O}(\mathbb{R}^d)$ and $A \in \mathrm{Conv}(\mathbb{R}^d)$. It is also known that the surface measure $r(\cdot)$ on Γ_1 is invariant under the action of elements in $\mathcal{O}(\mathbb{R}^d)$.

Lemma 2.5. *For every $A \in \mathrm{Conv}(\mathbb{R}^d)$, one has $s(A) \in A$.*

Proof. Suppose there is $A \in \mathrm{Conv}(\mathbb{R}^d)$ such that $s(A) \notin A$. Define $C = A - s(A)$. Then $0 \notin C$, and by (i)–(iii), we get $s(C) = 0$. Let $0 \neq \hat{c}$ be such that $\langle c - \hat{c}, \hat{x}\rangle > 0$ for every $c \in C$, where $\hat{x} = \hat{c}\|\hat{c}\|^{-1}$, and $\langle\cdot,\cdot\rangle$ denotes the inner product in $\mathbb{R}^d$. But $\langle c, \hat{x}\rangle = \langle \hat{c} + (c-\hat{c}), \hat{x}\rangle = \langle \hat{c}, \hat{x}\rangle + \langle c - \hat{c}, \hat{x}\rangle$ and $\langle \hat{c}, \hat{x}\rangle = \|\hat{c}\|$. Then for every $c \in C$, one has $\|\hat{c}\| \leq \langle c, \hat{x}\rangle$.

Let $l : \mathbb{R}^d \to \mathbb{R}^d$ be the linear transformation defined by $l(\hat{x}) = \hat{x}$ and $l(x) = -x$ for $x \in \mathbb{R}^d$ orthogonal to $\hat{x}$. It can be verified that l belongs to the group $\mathcal{O}(\mathbb{R}^d)$ of orthogonal linear transformations on $\mathbb{R}^d$ and $l^2 = I$, the identity map. So $l = l^*$. Let $D = C + l(C)$. Then $l(D) = D$, and so $s(D) = 0$. In addition, for every $d \in D$, we have $\langle d, \hat{x}\rangle \geq 2\|\hat{c}\| > 0$, and so $0 \notin D$. Now let

$$\begin{aligned} &\Gamma_1^0 = \{y \in \Gamma_1 : \langle y, \hat{x}\rangle = 0\}, \Gamma_1^+ = \{y \in \Gamma_1 : \langle y, \hat{x}\rangle > 0\} \text{ and} \\ &\Gamma_1^- = \{y \in \Gamma_1 : \langle y, \hat{x}\rangle < 0\}. \end{aligned}$$

Then $\Gamma_1 = \Gamma_1^0 \cup \Gamma_1^+ \cup \Gamma_1^-$, and these three sets Γ_1^0, Γ_1^+, Γ_1^- are disjoint. Also, $r(\Gamma_1^0) = 0$. So we have

$$\begin{aligned} s(D) &= d\int_{\Gamma_1^+} \sigma(y,D)\,\mathrm{d}r + d\int_{\Gamma_1^-} \sigma(y,D)\,\mathrm{d}r \\ &= d\int_{\Gamma_1} [\sigma(y,D) - \sigma(-y,D)]\,\mathrm{d}r. \end{aligned}$$

Let $y \in \Gamma_1^+$ and $e \in D$ be such that $\sigma(-y, D) = \langle -y, e\rangle$. Then

$$\begin{aligned} \sigma(y,D) - \sigma(-y,D) &= \sigma(y, l(D)) - \sigma(-y,D) \\ &= \sigma(l(y), D) - \sigma(-y,D) \\ &\geq \langle l(y), e\rangle + \langle y, e\rangle = \langle (l+I)(y), e\rangle\,. \end{aligned}$$

But $(l + I)(y) = 2\langle y, \hat{x}\rangle \hat{x}$. Then $\sigma(y, D) - \sigma(-y, D) \geq 2\langle y, \hat{x}\rangle \cdot \langle \hat{x}, e\rangle > 0$, since $y \in \Gamma_1^+$ and $\langle \hat{x}, e\rangle > 0$. Therefore,

$$\langle s(D), \hat{x}\rangle = d \int_{\Gamma_1} [\sigma(y, D) - \sigma(-y, D)] \cdot \langle y, \hat{x}\rangle \, \mathrm{d}r > 0,$$

which contradicts $s(D) = 0$. Then $s(A) \in A$ for $A \in \mathrm{Conv}(\mathbb{R}^d)$. □

Corollary 2.2. *There is* $K > 0$ *such that for every* $A, B \in \mathrm{Conv}(\mathbb{R}^d)$*, one has* $|s(A) - s(B)| \leq K \cdot h(A, B)$.

Proof. Let us observe that for $A, B \in \mathrm{Conv}(\mathbb{R}^d)$, we have $h(A, B) = \max\{|\sigma(x, A) - \sigma(x, B)| : \|x\| = 1\}$. Then $|s(A) - s(B)| \leq d \int_{\Gamma_1} y\, |\sigma(y, A) - \sigma(y, B)| \, \mathrm{d}r \leq K \cdot h(A, B)$ for every $K \geq d$. □

Remark 2.6. In the above inequality we can compute the optimal Lipschitz constant $K(d) > 0$. It is equal to $d!!/(d-1)!!$ if d is odd, and $K(d) = d!!/[\pi(d-1)!!]$ if d is even. □

Theorem 2.3. *If* (X, ρ) *is a metric space and* $F : X \to \mathrm{Conv}(\mathbb{R}^d)$ *is Lipschitz continuous, then* F *admits a Lipschitz continuous selection.*

Proof. Let $h(F(x_1), F(x_2)) \leq L\rho(x_1, x_2)$ for some $L > 0$ and every $x_1, x_2 \in X$. Put $f(x) = s(F(x))$ for $x \in X$. By Corollary 2.2, we get $|f(x_1) - f(x_2)| = |s(F(x_1)) - s(F(x_2))| \leq K(d) \cdot h(F(x_1), F(x_2)) \leq K(d) \cdot L\, \rho(x_1, x_2)$, where $K(d)$ is as in Remark 2.6. By Lemma 2.5, for every $x \in X$, we have $f(x) \in F(x)$. □

Remark 2.7. Theorem 2.3 cannot be extended to multifunctions with values in an infinite-dimensional Banach space $(Y, ||\cdot||)$. It can be proved that if a Lipschitz continuous multifunction $F : X \to \mathrm{Conv}(Y)$ admits a Lipschitz continuous selection, then Y is finite-dimensional. □

Remark 2.8. It can be proved that if $F : X \to \mathcal{P}(\mathbb{R}^d)$ with $X \in \mathrm{Conv}(\mathbb{R}^m)$ is convex-valued such that $F^-(\{y\}) = \{x \in X : y \in F(x)\}$ is an open set in X for every $y \in \mathbb{R}^d$, then F admits an C^∞- selection. □

We shall now show that some measurable multifunctions admit measurable selections. We begin with the following lemma.

Lemma 2.6. *Let* (X, ρ) *be a separable metric space and* $(T, \mathcal{F})$ *a measurable space. Then a multifunction* $F : T \to \mathcal{P}(X)$ *is weakly measurable if and only if the function* $T \ni t \to \mathrm{dist}(x, F(t)) \in \mathbb{R}^+$ *is measurable for each* $x \in X$.

Proof. Let us observe that F is weakly measurable if and only if $F^-(B(x, \varepsilon)) \in \mathcal{F}$ for every $x \in X$ and $\varepsilon > 0$. On the other hand, a function $T \ni t \to \mathrm{dist}(x, F(t)) \in \mathbb{R}^+$ is measurable for fixed $x \in X$ if and only if $\{t \in T : \mathrm{dist}(x, F(t)) < \varepsilon\} \in \mathcal{F}$ for every $\varepsilon > 0$. But $F^-(B(x, \varepsilon)) = \{t \in T : F(t) \cap B(x, \varepsilon) \neq \emptyset\} = \{t \in T : \mathrm{dist}(x, F(t)) < \varepsilon\}$. This completes the proof. □

Theorem 2.4 (Kuratowski and Ryll-Nardzewski). *Let (X, ρ) be a Polish space and $(T, \mathcal{F})$ a measurable space. If $F : T \to \mathrm{Cl}(X)$ is measurable, then F admits a measurable selector.*

Proof. Let $\{x_1, x_2, \ldots\}$ be a countable dense subset in X and let $B_n(i) = \{x \in X : \rho(x, x_i) \le 1/n\}$ for $i, n \ge 1$. Without any loss of generality, we may assume that $\mathrm{diam}(X) < 1$, where $\mathrm{diam}(X) = \sup\{\rho(x, y) : x, y \in X\}$. We will construct a sequence $(f_n)_{n=1}^{\infty}$ of measurable functions $f_n : T \to X$ such that

$$\text{(i)} \quad \mathrm{dist}(f_n(t), F(t)) \le \varepsilon_n \quad \text{and} \quad \text{(ii)} \quad \rho(f_n(t), f_{n-1}(t)) \le \varepsilon_{n-1}$$

for $n \ge 0$ and $t \in T$, where $\varepsilon_n = (1/2)^n$ for $n = 0, 1, 2, \ldots$. Let $f_0(t) = x_1$ for $t \in T$. Then $\mathrm{dist}(f_0(t), F(t)) < 1$. Suppose $f_0, \ldots, f_{n-1}$ have been constructed and let $A_k^n = \{t \in T : \mathrm{dist}(f_k(t), F(t)) < \varepsilon_n\}$ and $C_k^n = \{t \in T : \rho(x_k, f_{n-1}(t)) < \varepsilon_{n-1}\}$. Put $D_k^n = A_k^n \cap C_k^n$. We claim that $T = \bigcup_{k \ge 1} D_k^n$ for $n \ge 1$. Fix $t \in T$. By the inductive hypothesis, we can find $z \in F(t)$ such that $\rho(f_{n-1}(t), z) < \varepsilon_{n-1}$. On the other hand, there is $k \ge 1$ such that $\rho(x_k, z) < \varepsilon_n$ and $\rho(x_k, z) + \rho(z, f_{n-1}(t)) < \varepsilon_n + \varepsilon_{n-1} < 2\varepsilon_{n-2} = \varepsilon_{n-1}$. Therefore, $t \in D_k^n$ and $T \subset \bigcup_{k \ge 1} D_k^n$. By virtue of Lemma 2.6 and the continuity of the function $\mathrm{dist}(\cdot, F(t))$ for fixed $t \in T$, we obtain that $A_k^n \in \mathcal{F}$. The inductive hypothesis gives that $C_k^n \in \mathcal{F}$. Then $D_k^n \in \mathcal{F}$. Now define $f_n : T \to X$ by setting $f_n(t) = x_k$ for $t \in D_k^n \setminus \bigcup_{i=1}^{k-1} D_i^n$. Clearly, f_n is measurable. Moreover, by (ii), we see that $(f_n(t))_{n=1}^{\infty}$ is a Cauchy sequence in X for every fixed $t \in T$. Then there exists a function $f : T \to X$ such that $f_n(t) \to f(t)$ for every $t \in T$ as $n \to \infty$. We also have $\mathrm{dist}(f(t), F(t)) = 0$ for every $t \in T$. Hence it follows that f is measurable such that $f(t) \in F(t)$ for every $t \in T$. □

In what follows, we shall consider "complete" measurable spaces defined in the following way. For a given measurable space $(T, \mathcal{F})$ and every probability measure μ on $\mathcal{F}$, we denote by $\mathcal{F}_\mu$ the μ-completion of $\mathcal{F}$ and define $\tilde{\mathcal{F}} = \bigcap_\mu \mathcal{F}_\mu$. The space $(T, \mathcal{F})$ is said to be complete if $\mathcal{F} = \tilde{\mathcal{F}}$.

Remark 2.9. It can be proved that for a given complete measure space $(T, \mathcal{F}, \mu)$, a multifunction $F : T \to \mathcal{P}(\mathbb{R}^n)$ such that $\mathrm{Graph}(F) \in \mathcal{F} \otimes \beta(\mathbb{R}^m)$ is measurable and admits a measurable selection. □

A consequence of the above measurable selection theorem is the following implicit function theorem.

Theorem 2.5. *Assume that (X, ρ) is a Polish space, $(T, \mathcal{F})$ a measurable space, and (Y, d) a metric space. Suppose $f : T \times X \to Y$ is a function measurable in $t \in T$ and continuous in $x \in X$, and let $\Gamma : T \to \mathrm{Comp}(X)$ be a measurable multifunction and $g : T \to Y$ a measurable function such that $g(t) \in f(t, \Gamma(t))$ for $t \in T$. Then there exists a measurable function $\gamma : T \to X$ such that $\gamma(t) \in \Gamma(t)$ and $g(t) = f(t, \gamma(t))$ for $t \in T$.*

Proof. Let us observe that the set-valued function $F : T \to \mathcal{P}(X)$ defined by $F(t) = \{x \in X : f(t, x) \in \mathcal{U}\}$ for $t \in T$ is measurable for every open set $\mathcal{U} \subset Y$.

Indeed, let B be a closed subset of X and let A be a countable dense subset of B. We have

$$\begin{aligned} F^{-}(B) &= \{t \in T : F(t) \cap B \neq \emptyset\} \\ &= \{t \in T : f(t,x) \in \mathcal{U} \ \text{ for some } \ x \in B\} \\ &= \{t \in T : f(t,a) \in \mathcal{U} \ \text{ for some } \ a \in A\} \\ &= \bigcup_{a \in A} \{t \in T : f(t,a) \in \mathcal{U}\}. \end{aligned}$$

Therefore, $F^{-}(B) \in \mathcal{F}$, because we have $\{t \in T : f(t,a) \in \mathcal{U}\} \in \mathcal{F}$ for every fixed $a \in A$. Define multifunctions $H(t) = \Gamma(t) \cap \{x \in X : d(f(t,x), g(t)) = 0\}$ for $t \in T$ and $F_n(t) = \{x \in X : d(f(t,x), g(t)) < 1/n\}$ for $t \in T$ and $n \geq 1$. For every $n = 1, 2, \ldots$, a multifunction F_n is measurable and also weakly measurable. Hence it follows that its closure $\bar{F}_n$ is weakly measurable, because $F_n^{-}(B) = \bar{F}_n^{-}(B)$ for every open set $B \subset X$. Clearly, $\{x \in X : d(f(t,x), g(t)) = 0\} = \bigcap_{n=1}^{\infty} \bar{F}_n(t)$ for $t \in T$, because $\bar{F}_n(t) \subset \{x \in X : d(f(t,x), g(t)) \leq 1/n\}$ for $t \in T$ and $n \geq 1$. Hence it follows that the multifunction H defined above can be also defined by $H(t) = \Gamma(t) \cap \left[\bigcap_{n=1}^{\infty} \bar{F}_n(t)\right]$ for $t \in T$, which implies that H is measurable. Therefore, by Theorem 2.4, there is a measurable selector γ for H that in particular is a selector for Γ satisfying $d(f(t, \gamma(t)), g(t)) = 0$ for $t \in T$. $\square$

Corollary 2.3. *If (X, ρ) is a Polish space, $(T, \mathcal{F})$ a measurable space, and $\Gamma : T \to \mathrm{Comp}(X)$ and $g : T \to X$ are measurable, then there exists a measurable selector γ for Γ such that $\mathrm{dist}(g(t), \Gamma(t)) = \rho(g(t), \gamma(t))$ for $t \in T$.* $\square$

The following important result follows immediately from the Kuratowski and Ryll-Nardzewski measurable selection theorem.

Theorem 2.6. *Let (X, ρ) be a Polish space, $(T, \mathcal{F})$ a measurable space, and let $F : T \to \mathrm{Cl}(X)$. The following conditions are equivalent:*

(i) F is measurable;
(ii) there exists a sequence $(f_n)_{n=1}^{\infty}$ of measurable selectors of F such that $F(t) = \mathrm{cl}\{f_1(t), f_2(t), \ldots\}$ for every $t \in T$.

Proof. Let F be measurable and $(x_n)_{n=1}^{\infty}$ a dense sequence of X. For every $n, k \geq 1$, we define

$$F_{n,k}(t) = \begin{cases} F(t) \cap B(x_n, \varepsilon_k) & \text{if } t \in F^{-}(B(x_n, \varepsilon_k)), \\ F(t) & \text{otherwise}, \end{cases}$$

where $\varepsilon_k = (1/2)^k$ and $F^{-}(B(x_n, \varepsilon_k)) = \{t \in T : F(t) \cap B(x_n, \varepsilon_k) \neq \emptyset\}$. Note that $F^{-}(B(x_n, \varepsilon_k)) \in \mathcal{F}$ and that the set-valued function $T \ni t \to F(t) \cap B(x_n, \varepsilon_k) \subset X$ is measurable. So $F_{n,k}$ is measurable, which implies that $\mathrm{cl}\,[F_{n,k}]$

is also measurable. Therefore, by Theorem 2.4, there exist measurable functions $f_{n,k} : T \to X$ such that $f_{n,k}(t) \in \text{cl}\,[F_{n,k}]\,(t)$ for every $t \in T$. We shall show that $F(t) = \text{cl}\{f_{n,k}(t) : n, k \geq 1\}$ for $t \in T$. Indeed, fix $t \in T$ and let $x \in F(t)$ and $\varepsilon > 0$. Let $k \geq 1$ and $n \geq 1$ be such that $\varepsilon_{k-1} \leq \varepsilon$ and $x \in B(x_n, \varepsilon_k)$. Then $t \in F^-(B(x_n, \varepsilon_k))$ and $f_{n,k}(t) \in \overline{B(x_n, \varepsilon_k)}$. So $\rho(f_{n,k}(t), x) \leq \rho(f_{n,k}(t), x_n) + \rho(x_n, x) \leq \varepsilon$, which proves that $F(t) = \text{cl}\{f_{n,k}(t) : n, k \geq 1\}$. Then (i) $\Rightarrow$ (ii). Assume that (ii) is satisfied. Then for every open set $\mathcal{U} \subset X$, we have

$$F^-(\mathcal{U}) = \{t \in T : F(t) \cap \mathcal{U} \neq \emptyset\} = \bigcup_{n \geq 1} \{t \in T : f_n(t) \in \mathcal{U}\} \in \mathcal{F}.$$

Then F is weakly measurable and therefore measurable. Thus (ii)$\Rightarrow$(i). □

Remark 2.10. It can be proved that if $(T, \mathcal{F})$ is a complete measurable space, $(G, \mathcal{G})$ is a measurable space, X is a Suslin space, $g : T \times G \to X$ is jointly measurable, $\Gamma : T \to \mathcal{P}(G)$ is a multifunction such that $\text{Graph}(\Gamma) \in \mathcal{F} \otimes \mathcal{G}$, and $h : T \to X$ is a measurable map such that $h(t) \in g(t, \Gamma(t))$ for $t \in T$, then there exists a measurable selector $\gamma : T \to G$ of Γ such that $h(t) = g(t, \gamma(t))$ for $t \in T$. □

We shall consider now the existence of Carathéodory-type selections of measurable multifunctions depending on two variables. More precisely, let $(T, \mathcal{F})$ be a measurable space, (X, ρ) a Polish space, and $(Y, \|\cdot\|)$ a separable Banach space. Consider the set-valued mapping $F : T \times X \to \text{Cl}(Y)$, which is assumed to be measurable, i.e., for every closed set $A \subset Y$, we have $F^-(A) = \{(t, x) \in T \times X : F(t, x) \cap A \neq \emptyset\} \in \mathcal{F} \otimes \beta(X)$. We are interested in the existence of a function $f : T \times X \to Y$, a selector of F, such that $f(\cdot, x)$ is measurable for fixed $x \in X$, and $f(t, \cdot)$ is continuous for fixed $t \in T$. Such selectors of F are said to be of Carathéodory type or simply to be Carathéodory selectors for F.

Theorem 2.7. *Let $(T, \mathcal{F})$ be a complete measurable space, (X, ρ) a Polish space, $(Y, \|\cdot\|)$ a separable Banach space, and $F : T \times X \to \text{Cl}(Y)$ a convex-valued measurable set-valued mapping. If furthermore, $F(t, \cdot)$ is l.s.c. for fixed $t \in T$, then F admits a Carathéodory selection.*

Proof. Let $(y_n)_{n=1}^{\infty}$ be a dense sequence of Y. For $t \in T$, $n \geq 1$, and $\varepsilon > 0$, define $G_n^{\varepsilon}(t) = \{x \in X : y_n \in (F(t, x) + \varepsilon B)\}$, where B is an open unit ball in Y. By the lower semicontinuity of $F(t, \cdot)$, a set $G_n^{\varepsilon}(t)$ is open for every $t \in T$, $\varepsilon > 0$, and $n \geq 1$. Also, the family $\{G_n^{\varepsilon}(t) : n \geq 1\}$ is an open covering of X. Moreover,

$$\text{Graph}(G_n^{\varepsilon}) = \{(t, x) \in T \times X : \text{dist}(y_n, F(t, x)) < \varepsilon\} \in \mathcal{F} \otimes \beta(X),$$

because of the measurability of F. Let $\varepsilon_m = (1/2)^m$ and

$$G_{n,m}^{\varepsilon}(t) = \{x \in G_n^{\varepsilon}(t) : \text{dist}(x, X \setminus G_n^{\varepsilon}) \geq \varepsilon_m\} \text{ and } \mathcal{U}_n^{\varepsilon}(t) = G_n^{\varepsilon}(t) \setminus \bigcup_{1 \leq k < n} G_{n,k}^{\varepsilon}(t)$$

for $n, m \geq 1$. It can be verified that the family $\{\mathcal{U}_n^{\varepsilon}(t) : n \geq 1\}$ is a locally finite covering of X and every multifunction $\mathcal{U}_n^{\varepsilon} : T \to \mathcal{P}(X)$ has a measurable graph. Hence it follows that the set-valued mapping $T \ni t \to X \setminus \mathcal{U}_n^{\varepsilon}(t) \subset X$ is measurable with closed values. Let

$$p_n^{\varepsilon}(t,x) = \frac{\mathrm{dist}(x, X \setminus \mathcal{U}_n^{\varepsilon}(t))}{\sum_{n\geq 1} \mathrm{dist}(x, X \setminus \mathcal{U}_n^{\varepsilon}(t))}.$$

By virtue of Lemma 2.6, the function $p_n^{\varepsilon}(\cdot, x)$ is measurable for every $n \geq 1$ and fixed $x \in X$. By the above definition, $p_n^{\varepsilon}(t, \cdot)$ is continuous for fixed $t \in T$. Then p_n^{ε} is a Carathéodory function for every $\varepsilon > 0$ and $n \geq 1$. Furthermore, $\sum_{n\geq 1} p_n^{\varepsilon}(t,x) = 1$. Let $f^{\varepsilon}(t,x) = \sum_{n\geq 1} p_n^{\varepsilon}(t,x) \cdot y_n$. It is clear that f^{ε} is a Carathéodory function. By the convexity of $F(t,x)$, for every $(t,x) \in T \times X$ we get $f^{\varepsilon}(t,x) \in F(t,x) + \varepsilon B$ for $(t,x) \in T \times X$ and every $\varepsilon > 0$.

Let $\varepsilon_n = (1/2)^n$ for $n = 1, 2, \ldots$. We define now a sequence $(f_n)_{n=1}^{\infty}$ of Carathéodory functions $f_n : T \times X \to Y$ such that $f_n(t,x) \in F(t,x) + \varepsilon_n B$ and $\| f_n(t,x) - f_{n-1} \| < \varepsilon_{n-1}$ for $(t,x) \in T \times X$ and $n \geq 2$. We start with $f_1 = f^{\varepsilon_1}$ and then we put $F_2(t,x) = F(t,x) \cap \{f_1(t,x) + \varepsilon_1 B\}$ for $(t,x) \in T \times X$. By virtue of Lemma 2.3, a multifunction $F_2(t,\cdot)$ is l.s.c. for fixed $t \in T$. It is easy to see that F_2 is measurable. Consequently, its closure $\mathrm{cl}\,[F_2]$ is measurable and $\mathrm{cl}\,[F_2]\,(t,\cdot)$ is l.s.c. for fixed $t \in T$. From this and the first part of the proof, it follows that for $\varepsilon = \varepsilon_2$, there exists a Carathéodory function f_2 such that $f_2(t,x) \in \mathrm{cl}\,[F_2]\,(t,x) + \varepsilon_2 B$ for $(t,x) \in T \times X$. It is clear that $f_2(t,x) \in F(t,x) + \varepsilon_2 B$ and $\| f_2(t,x) - f_1(t,x) \| < \varepsilon_1$ for $(t,x) \in T \times X$. By the inductive procedure, we can define a sequence $(f_n)_{n=1}^{\infty}$ of Carathéodory functions $f_n : T \times X \to Y$ such that $f_n(t,x) \in F(t,x) + \varepsilon_n B$ and $\| f_n(t,x) - f_{n-1}(t,x) \| < \varepsilon_{n-1}$ for $(t,x) \in T \times X$. Hence it follows that there exists a Carathéodory function $f : T \times X \to Y$ such that $f_n(t,x) \to f(t,x)$ as $n \to \infty$ for $(t,x) \in T \times X$. By the closedness of $F(t,x)$, this implies that $f(t,x) \in F(t,x)$ for $(t,x) \in T \times X$. □

Remark 2.11. It can be proved that if T is a locally compact metric space furnished with a Radon measure μ, X is a Polish space, Y is a separable reflexive Banach space, and $F : T \times X \to \mathrm{Cl}(Y)$ is as in Theorem 2.7, then there exists a sequence $(f_m)_{m=1}^{\infty}$ of Carathéodory selectors $f_m : T \times X \to Y$ of F such that $F(t,x) = \mathrm{cl}\{f_m(t,x) : m \geq 1\}$ for every $(t,x) \in T \times X$. □

There are quite a number of set-valued fixed-point theorems. We present below one of them that generalizes the classical Banach fixed-point theorem.

Theorem 2.8 (Covitz–Nadler). *Let (X, ρ) be a complete metric space and let $F : X \to \mathrm{Cl}(X)$ be such that $h(F(x), F(\bar{x})) \leq K\rho(x, \bar{x})$ for every $x, \bar{x} \in X$ with $K \in (0,1)$. Then there exists $x \in X$ such that $x \in F(x)$.*

Proof. Let $L \in (K, 1)$ and $\lambda = K^{-1}L$. For some $x \in X$, we have $B(x, \lambda \cdot \mathrm{dist}(x, F(x))) \cap F(x) \neq \emptyset$, because $\lambda > 1$. Then we can select $x_1 \in F(x)$ such that $\rho(x, x_1) \leq \lambda \cdot \mathrm{dist}(x, F(x))$. For such $x_1 \in X$, we can select $x_2 \in F(x_1)$ such that $\rho(x_1, x_2) \leq \lambda \cdot \mathrm{dist}(x_1, F(x_1))$. Continuing this procedure, we can find a sequence $(x_n)_{n=1}^{\infty}$ of X such that $\rho(x_n, x_{n+1}) \leq \lambda \cdot \text{dist}(x_n, F(x_n))$ for $n \geq 1$. Hence it follows that $\rho(x_n, x_{n+1}) \leq \lambda \cdot \text{dist}(x_n, F(x_n)) \leq \lambda \cdot h(F(x_{n-1}, F(x_n)) \leq L\rho(x_{n-1}, x_n) \leq L^n \text{dist}(x, F(x))$. Now, similarly as in the proof of the Banach fixed-point theorem, we can verify that the above defined sequence $(x_n)_{n=1}^{\infty}$ has a limit, say x,, belonging to X. Since F is H-continuous and $\text{dist}(x, F(x)) \leq \rho(x, x_n) + \text{dist}(x_n, F(x_{n+1})) + h(F((x_{n+1}), F(x))$ for $n \geq 1$, it follows that $x \in F(x)$. □

3 The Aumann Integral

Let $(T, \mathcal{F}, \mu)$ be a σ-finite measure space that is not necessarily complete. For $p \geq 1$, by $\mathbb{L}^p(T, \mathbb{R}^d)$ we denote the Banach space $\mathbb{L}^p(T, \mathcal{F}, \mu, \mathbb{R}^d)$ with the norm $\|\cdot\|$ defined in the usual way, i.e., by $\|f\|^p = \int_T |f(t)|^p d\mu$ for $f \in \mathbb{L}^p(T, \mathbb{R}^d)$. In what follows, we shall consider $\mathbb{L}^p(T, \mathbb{R}^d)$ with $p = 1$ and $p = 2$. Instead of $\mathbb{L}^1(T, \mathbb{R}^d)$, we shall write $\mathbb{L}(T, \mathbb{R}^d)$. Let us recall that if $\mu(T) < \infty$, then a set $K \subset \mathbb{L}^p(T, \mathbb{R}^d)$ is relatively sequentially weakly compact if K is bounded and uniformly integrable, i.e., if $\lim_{\mu(E)\to 0} \int_E f(t) d\mu = 0$ uniformly for $f \in K$. By the reflexivity of $\mathbb{L}^2(T, \mathbb{R}^d)$, a set $K \subset \mathbb{L}^2(T, \mathbb{R}^d)$ is relatively sequentially weakly compact if and only if it is bounded. By the Eberlein–Šmulian theorem, it follows that for a bounded set $K \subset \mathbb{L}^2(T, \mathbb{R}^d)$, its closure $\text{cl}_w K$ with respect to the weak topology of $\mathbb{L}^2(T, \mathbb{R}^d)$ is weakly compact. In particular, if K is also closed and convex, then it is weakly compact, because in such a case, we have $K = \text{cl}_w K$.

Given a measurable set-valued mapping $F : T \to \text{Cl}(\mathbb{R}^d)$, we define subtrajectory integrals $S(F)$ of F as the subset of the space $\mathbb{L}^p(T, \mathbb{R}^d)$ defined by $S(F) = \{f \in \mathbb{L}^p(T, \mathbb{R}^d) : f(t) \in F(t) \ a.e.\}$. It can be verified that $S(F)$ is a closed subset of $\mathbb{L}^p(T, \mathbb{R}^d)$. In what follows we shall consider only the cases $p = 1$ and $p = 2$. Immediately from properties of multifunction F it will be easily seen if $S(F)$ is a subset of $\mathbb{L}(T, \mathbb{R}^d)$ or $\mathbb{L}^2(T, \mathbb{R}^d)$, respectively. In what follows, we shall denote by $\mathcal{M}(T, \mathbb{R}^d)$ the space of all measurable set-valued mappings $F : T \to \text{Cl}(\mathbb{R}^d)$ and by $\mathcal{A}(T, \mathbb{R}^d)$ the subspace of $\mathcal{M}(T, \mathbb{R}^d)$ containing all $F \in \mathcal{M}(T, \mathbb{R}^d)$ such that $S(F) \neq \emptyset$. It can be proved that every $F \in \mathcal{M}(T, \mathbb{R}^d)$ belongs to $\mathcal{A}(T, \mathbb{R}^d)$ if and only if there exists $k \in \mathbb{L}^p(T, \mathbb{R}^+)$ such that $\text{dist}(0, F(t)) \leq k(t)$ for a.e. $t \in T$. We have the following simple results.

Lemma 3.1. *If $F \in \mathcal{A}(T, \mathbb{R}^d)$, then there exists a sequence $(f_n)_{n=1}^{\infty}$ of functions $f_n \in S(F)$ such that $F(t) = \text{cl}\{f_1(t), f_2(t), \ldots\}$ for $t \in T$.*

Proof. By virtue of Theorem 2.6, there exists a sequence $(g_n)_{n=1}^{\infty}$ of measurable functions $g_n : T \to \mathbb{R}^d$ such that $F(t) = \text{cl}\{g_1(t), g_2(t), \ldots\}$ for $t \in T$. Taking a countable measurable partition $\{A_1, A_2, \ldots\}$ of T with $\mu(A_k) < \infty$ and a function $f \in \mathbb{L}^p(T, \mathbb{R}^d)$ such that $f(t) \in F(t)$ for $t \in T$, we define $B_{j,m,k} = \{t \in T : m - 1 \leq |g_j(t)| < m\} \cap A_k$ and $f_{j,m,k} = \mathbb{1}_{B_{j,m,k}} g_j + \mathbb{1}_{T \setminus B_{j,m,k}} f$ for $j, m, k \geq 1$. It is easy to see that $f_{j,m,k} \in S(F)$ and $F(t) = \overline{\{f_{j,m,k}(t) : j, m, k \geq 1\}}$ for $t \in T$. □

Corollary 3.1. *If* $F, G \in \mathcal{A}(T, \mathbb{R}^d)$*, then* $S(F) = S(G)$ *if and only if* $F(t) = G(t)$ *for a.e.* $t \in T$. □

Lemma 3.2. *Let* $F \in \mathcal{A}(T, \mathbb{R}^d)$ *and let* $(f_n)_{n=1}^{\infty}$ *be a sequence of* $S(F)$ *such that* $F(t) = \mathrm{cl}\{f_1(t), f_2(t), \ldots\}$ *for* $t \in T$*. Then for every* $f \in S(F)$ *and* $\varepsilon > 0$*, there exists a finite measurable partition* $\{A_1, \ldots, A_m\}$ *of* T *such that* $\| f - \sum_{i=1}^{m} \mathbb{1}_{A_i} f_i \| < \varepsilon$.

Proof. Assume $f(t) \in F(t)$ for every $t \in T$ and let $\rho \in \mathbb{L}^p(T, \mathbb{R})$ be strictly positive such that $\int_T \rho \mathrm{d}\mu < \varepsilon/3$. Then there exists a countable measurable partition $\{B_1, B_2, \ldots\}$ of T such that $|f(t) - f_i(t)| < \rho(t)$ for $t \in B_i$ and $i \geq 1$. Take an integer m such that $\sum_{i=m+1}^{\infty} \int_{B_i} |f(t)| \mathrm{d}\mu < \varepsilon/6$ and $\sum_{i=m+1}^{\infty} \int_{B_i} |f_i(t)| \mathrm{d}\mu < \varepsilon/6$ and define a finite measurable partition $\{A_1, \ldots, A_m\}$ as follows: $A_1 = B_1 \cup (\bigcup_{i=m+1}^{\infty} B_i)$ and $A_j = B_j$ for $2 \leq j \leq m$. Then we have

$$\left\| f - \sum_{i=1}^{m} \mathbb{1}_{A_i} f_i \right\| = \sum_{i=1}^{m} \int_{B_i} |f(t) - f_i(t)| \mathrm{d}\mu + \sum_{i=m+1}^{\infty} \int_{B_i} |f(t) - f_i(t)| \mathrm{d}\mu$$

$$\leq \int_T \rho \mathrm{d}\mu + \sum_{i=m+1}^{\infty} \int_{B_i} (|f(t)| + |f_i(t)|) \mathrm{d}\mu < \varepsilon. \qquad \square$$

Lemma 3.3. *Let* $(T, \mathcal{F}, \mu)$ *be a measure space with a* σ*-finite measure* μ*. If* $F \in \mathcal{A}(T, \mathbb{R}^d)$*, then* $\overline{\mathrm{co}}\, S(F) = S(\overline{\mathrm{co}}\, F)$.

Proof. We have $\overline{\mathrm{co}}\, S(F) \subset S(\overline{\mathrm{co}}\, F)$. Assume that there exists $f \in S(\overline{\mathrm{co}}\, F)$ such that $f \notin \overline{\mathrm{co}}\, S(F)$. By the strong separation theorem, we can find $h \in \mathbb{L}^{\infty}(T, \mathbb{R}^d)$ such that $\sup\{(h, g) : g \in S(F)\} < (h, f)$, where $(\cdot, \cdot)$ denotes the duality bracket. Hence it follows that $\int_T \sigma(h(t), \overline{\mathrm{co}}\, F(t)) \mathrm{d}\mu < \int_T \langle h(t), f(t) \rangle \, \mathrm{d}\mu$. On the other hand, $f(t) \in \overline{\mathrm{co}}\, F(t)$ a.e. Then $\int_T \langle h(t), f(t) \rangle \, \mathrm{d}\mu \leq \int_T \sigma(h(t), \overline{\mathrm{co}}\, F(t)) \mathrm{d}\mu$, a contradiction. Therefore, $\overline{\mathrm{co}}\, S(F) = S(\overline{\mathrm{co}} F)$. □

A multifunction $F : T \to \mathcal{P}(\mathbb{R}^n)$ is said to be p-integrably bounded if there is $k \in \mathbb{L}^p(T, \mathbb{R}^+)$ such that $\|F(t)\| =: h(\{0\}, F(t)) \leq k(t)$ for a.e. $t \in T$. In particular, for $p = 1$, we say simply integrably bounded instead of 1-integrably bounded. Similarly, if $p = 2$, then instead of 2-integrably bounded, we say square integrably bounded. It is clear that F is p-integrably bounded if and only if the function $T \ni t \to \|F(t)\| \in \mathbb{R}^+$ belongs to $\mathbb{L}^p(T, \mathbb{R}^+)$. For every p-integrably bounded multifunction $F \in \mathcal{M}(T, \mathbb{R}^n)$, we have $S(F) \neq \emptyset$.

Remark 3.1. Immediately from the definition of subtrajectory integrals, it follows that for every measurable and p-integrably bounded multifunction $F : T \to \mathrm{Conv}(\mathbb{R}^d)$, its subtrajectory integral $S(F)$ is a nonempty convex weakly sequentially compact subset of $\mathbb{L}^p(T, \mathbb{R}^d)$. In particular, it is a weakly compact convex subset of this space for $p > 1$. □

Lemma 3.4. *If* $F, G \in \mathcal{A}(T, \mathbb{R}^d$ *then* $S(\overline{F + G}) = \overline{S(F) + S(G)}$.

Proof. Immediately from Theorem 2.6, it follows that $H = \overline{F+G}$ is measurable. It is clear that $S(H)$ is closed, and therefore, $\overline{S(F)+S(G)} \subset S(H)$. On the other hand, we may find sequences $(f_n)_{n=1}^{\infty} \subset S(F)$ and $(g_m)_{m=1}^{\infty} \subset S(G)$ such that $F(t) = \text{cl}\{f_n(t) : n \geq 1\}$ and $G(t) = \text{cl}\{g_n(t) : n \geq 1\}$ a.e. Evidently, $H(t) = \overline{\{f_n(t) + g_m(t) : n, m \geq 1\}}$, which, by Lemma 3.2, implies that for given $h \in S(H)$ and $\varepsilon > 0$, we can select a finite $\mathcal{F}$-measurable partition $(A_k)_{k=1}^{N}$ of T and positive integers $n_1, \ldots, n_N$ and $m_1, \ldots, m_N$ such that $\|h - \sum_{k=1}^{N} \mathbb{1}_{A_k}(f_{n_k} + g_{m_k})\| < \varepsilon$. Hence it follows that $h \in \overline{S(F)+S(G)}$. Then $S(H) \subset \overline{S(F)+S(G)}$. □

Let $(T, \mathcal{F}, \mu)$ be a measure space, $\tilde{\mathbb{R}} = [-\infty, +\infty]$ and let $\phi : T \times X \to \tilde{\mathbb{R}}$ be an $\mathcal{F} \otimes \beta(\mathbb{R}^d)$-measurable function. The functional $\mathcal{T}_\phi$ defined on the space $\mathbb{L}^0(T, \mathbb{R}^d)$ of measurable functions $f : T \to \mathbb{R}^d$ by setting $\mathcal{T}_\phi(f) = \int_T \phi(t, f(t))\mathrm{d}\mu$ if the integral exists, permitting $+\infty$ or $-\infty$, is called the integral functional.

Lemma 3.5. *Let $F \in \mathcal{M}(T, \mathbb{R}^d)$ and let $\phi : T \times \mathbb{R}^d \to \tilde{\mathbb{R}}$ be $\mathcal{F} \otimes \beta(\mathbb{R}^d)$-measurable. Assume either that (i) $\phi(t,x)$ is u.s.c. in x for every fixed $t \in T$ or that (ii) $(T, \mathcal{F}, \mu)$ is complete and $\phi(t,x)$ is l.s.c. in x for every fixed $t \in T$. Then the function $T \ni t \to \inf\{\phi(t,x) : x \in F(t)\} \subset \tilde{\mathbb{R}}$ is measurable.*

Proof. Let $\xi(t) = \inf\{\phi(t,x) : x \in F(t)\}$ and assume that (i) is satisfied. By Theorem 2.6, there exists a sequence $(f_n)_{n=1}^{\infty}$ of measurable selectors of F such that $F(t) = \text{cl}(\{f_1(t), f_2(t), \ldots\})$ for $t \in T$. Then we have $\xi(t) = \inf_{n \geq 1} \phi(t, f_n(t))$ for $t \in T$, which implies that ξ is measurable. Let (ii) be satisfied and let $H : T \to \mathcal{P}(\mathbb{R}^d \times \mathbb{R})$ be defined by $H(t) = \{(x, \alpha) \in \mathbb{R}^d \times \mathbb{R} : x \in F(t),\ \phi(t,x) \leq \alpha\}$ for $t \in T$. Then $H(t)$ is closed in $\mathbb{R}^d \times \mathbb{R}$ for every $t \in T$, and $\text{Graph}(H) = [\text{Graph}(F) \cap \mathbb{R}] \cap \{(t, x, \alpha) : \Phi(t,x) - \alpha \leq 0\}$ belongs to $\mathcal{F} \otimes \beta(\mathbb{R}^d) \otimes \beta(\mathbb{R}) = \mathcal{F} \otimes \beta(\mathbb{R}^d \otimes \mathbb{R})$. Therefore, by virtue of Remark 2.9 and Theorem 2.6, there exists a sequence $(g_n, \xi_n)_{n=1}^{\infty}$ of measurable functions $g_n : T \to \mathbb{R}^d$ and $\xi_n : T \to \mathbb{R}$ such that $H(t) = \text{cl}(\{(g_1, \xi_1)(t), (g_2, \xi_2)(t), \ldots\})$ for $t \in \text{Dom}(H)$. Hence we have $\xi(t) = \inf_{n \geq 1} \xi_n(t)$ for $t \in \text{Dom}(H)$ and $\xi(t) = \infty$ for $t \in T \setminus \text{Dom}(H)$. This shows that ξ is measurable. □

Theorem 3.1. *Let $F \in \mathcal{A}(T, \mathbb{R}^d)$ and let $\phi : T \times X \to \tilde{\mathbb{R}}$ be $\mathcal{F} \otimes \beta(\mathbb{R}^d)$-measurable. Assume either that (i) $\phi(t,x)$ is u.s.c. in x for every fixed $t \in T$, or that (ii) $(T, \mathcal{F}, \mu)$ is complete and $\phi(t,x)$ is l.s.c. in x for every fixed $t \in T$. If the integral functional $\mathcal{T}_\phi$ is defined for all $f \in S(F)$ and $\mathcal{T}_\phi(f_0) < \infty$ for some $f_0 \in S(F)$, then $\inf\{\mathcal{T}_\phi(f) : f \in S(F)\} = \int_T \inf\{\phi(t,x) : x \in F(t)\}\mathrm{d}\mu$.*

Proof. Let $\xi(t) = \inf\{\phi(t,x) : x \in F(t)\}$. By virtue of Lemma 3.4, ξ is measurable and $\xi(t) \leq \phi(t, f(t))$ a.e. for every $f \in S(F)$. Taking $f = f_0$, we can see that the integral of ξ exists and $\int_T \xi \mathrm{d}\mu \leq \inf\{\mathcal{T}_\phi(f) : f \in S(F)\}$. If $\mathcal{T}_\phi(f_0) = -\infty$, then the proof is complete. Thus assume $\mathcal{T}_\phi(f_0)$ to be finite, so that the function $T \ni t \to \phi(t, f_0(t)) \in \tilde{\mathbb{R}}$ is in $\mathbb{L}(T, \mathbb{R})$. Let $\beta > \int_T \xi \mathrm{d}\mu$ be given. We shall show that $\mathcal{T}_\phi(f) < \beta$ for some $f \in S(F)$. Take a sequence

$(A_n)_{n=1}^\infty$ of measurable sets $A_n \in \mathcal{F}$ such that $\mu(A_n) < \infty$ and such that $A_n \uparrow T$ and a strictly positive function $\rho \in \mathbb{L}(T, \mathbb{R})$. For $n \geq 1$, define $B_n = A_n \cap \{t \in T : \phi(t, f_0(t)) \geq -n\}$ and

$$\xi_n(t) = \begin{cases} \xi(t) + \rho(t)/n \quad textif \; t \in B_n \text{ and } \xi(t) \geq -n, \\ -n + \rho(t)/n \quad \text{if } \; t \in B_n \text{ and } \xi(t) < -n, \\ \phi(t, f_0(t)) + \rho(t)/n \quad \text{if } \; t \in T \setminus B_n \; . \end{cases}$$

It is easy to see that $\xi_n \in \mathbb{L}(T, \mathbb{R})$ for $n \geq 1$ and $\xi_n(t) \downarrow \xi(t)$ a.e., so that $\int_T \xi_{n_0} \mathrm{d}\mu < \beta$ for some n_0. Setting $\zeta = \xi_{n_0}$, we have $\int_T \zeta \mathrm{d}\mu < \beta$ and $\xi(t) < \zeta(t)$ a.e. We claim now that there exists a measurable function $g : T \to \mathbb{R}^d$ satisfying $g(t) \in F(t)$ a.e. and $\phi(t, g(t)) \leq \zeta(t)$ a.e. For case (i), take a sequence $(g_i)_{i=1}^\infty$ of measurable functions such that $F(t) = \mathrm{cl}(\{g_1(t), g_2(t), \ldots\}$ for all $t \in T$. Since $\inf_{i \geq 1} \phi(t, g_i(t)) = \xi(t)$ a.e., there exists a measurable function g satisfying the conditions desired above. For case (ii), define $F_1(t) = F(t) \cap \{x \in \mathbb{R}^d : \phi(t, x) \leq \zeta(t)\}$ for $t \in T$. Since $F_1(t)$ is closed for every $t \in T$ and $\mathrm{Graph}(F_1) \in \mathcal{F} \otimes \beta(\mathbb{R}^d)$ it follows by Remark 2.9 that F_1 has a measurable selection on $\mathrm{Dom}(F_1) \in \mathcal{F}$. Thus the desired g is obtained from the condition $\mu(T \setminus \mathrm{Dom}(F_1)) = 0$. Using the function g defined above, we define $C_n = A_n \cap \{t \in T : |g(t)| \leq n\}$ and $f_n = \mathbb{1}_{C_n} g + \mathbb{1}_{T \setminus C_n} f_0$ for $n \geq 1$ such that $f_n \in S(F)$ for $n \geq 1$ and

$$\begin{aligned} \mathcal{T}_\phi(f_n) &= \int_{C_n} \phi(t, g(t)) \mathrm{d}\mu + \int_{T \setminus C_n} \phi(t, f_0(t)) \mathrm{d}\mu \\ &\leq \int_T \zeta \mathrm{d}\mu + \int_{T \setminus C_n} [\phi(t, f_0(t)) - \zeta] \mathrm{d}\mu. \end{aligned}$$

Since $\int_T \zeta \mathrm{d}\mu < \beta$ and $C_n \uparrow T$, we have $\mathcal{T}_\phi(f_n) < \beta$. □

Corollary 3.2. *If $F \in \mathcal{A}(T, \mathbb{R}^d)$ if $\phi : T \times X \to \tilde{\mathbb{R}}$ is $\mathcal{F} \otimes \beta(\mathbb{R}^d)$-measurable and satisfies (i) or (ii) of Theorem 3.1, and if $\mathcal{T}_\phi$ is defined for all $f \in S(F)$ and $\mathcal{T}_\phi(f_0) > -\infty$ for some $f_0 \in S(F)$, then* $\sup\{\mathcal{T}_\phi(f) : f \in S(F)\} = \int_T \sup\{\phi(t, x) : x \in F(t)\} \mathrm{d}\mu$. □

Corollary 3.3. *For every $F \in \mathcal{A}(T, \mathbb{R}^d)$, one has* $\sup\{\|f\|^p : f \in S(F)\} = \int_T \sup\{|x|^p : x \in F(t)\} \mathrm{d}\mu = \int_T \|F(t)\|^p \mathrm{d}\mu$. *Then F is p-integrably bounded if and only if $S(F)$ is a bounded subset of $\mathbb{L}^p(T, \mathbb{R}^d)$.* □

Let $M \subset \mathbb{L}^0(T, \mathbb{R}^d)$ be a set of measurable functions $f : T \to \mathbb{R}^d$. We call M decomposable with respect to $\mathcal{F}$ if $f_1, f_2 \in M$ and $A \in \mathcal{F}$ imply $\mathbb{1}_A f_1 + \mathbb{1}_{T \setminus A} f_2 \in M$. It is clear that if M is decomposable, then $\sum_{i=1}^m \mathbb{1}_{A_i} f_i \in M$ for each finite $\mathcal{F}$-measurable partition $\{A_1, \ldots, A_m\}$ of T and $\{f_1, \ldots, f_m\} \subset M$. The following theorem is a characterization of decomposable subsets of the space $\mathbb{L}^p(T, \mathbb{R}^d)$.

Theorem 3.2. *Let M be a nonempty closed subset of $\mathbb{L}^p(T,\mathbb{R}^d)$ with $p \geq 1$. Then there exists an $F \in \mathcal{A}(T,\mathbb{R}^d)$ such that $M = S(F)$ if and only if M is decomposable.*

Proof. Let us observe that $S(F)$ is decomposable for every $F \in \mathcal{A}(T,\mathbb{R}^d)$. If $M \subset \mathbb{L}^p(T,\mathbb{R}^d)$ is such that there exists $F \in \mathcal{A}(T,\mathbb{R}^d)$ such that $M = S(F)$, then it is decomposable. To prove the converse, assume that M is a nonempty closed decomposable subset of $\mathbb{L}^p(T,\mathbb{R}^d)$. Let us observe that a multifunction G defined by $G(t) = \mathbb{R}^d$ for every $t \in T$ belongs to $\mathcal{A}(T,\mathbb{R}^d)$. Therefore, by virtue of Lemma 3.1, there exists a sequence $(f_i)_{i=1}^{\infty}$ of $\mathbb{L}^p(T,\mathbb{R}^n)$ such that $\mathbb{R}^d = \text{cl}(f_i(t) : i \geq 1\}$ for every $t \in T$. Let $\alpha_i = \inf\{\|f_i - g\| : g \in M\}$ for $i \geq 1$ and choose a sequence $\{g_{ij} : j \geq 1\} \subset M$ such that $\|f_i - g_{ij}| \to \alpha_i$ as $j \to \infty$. Define $F \in \mathcal{A}(T,\mathbb{R}^d)$ by $F(t) = \text{cl}\{g_{ij}(t) : i,j \geq 1\}$. We shall prove that $M = S(F)$. By Lemma 3.2, for each $f \in S(F)$ and $\varepsilon > 0$, we can select a finite measurable partition $\{A_1,\dots,A_m\}$ of T and $\{h_1,\dots,h_m\} \subset \{g_{ij}(t) : i,j \geq 1\}$ such that $\|f - \sum_{k=1}^m \mathbb{1}_{A_k} h_k\| < \varepsilon$. Since $\sum_{k=1}^m \mathbb{1}_{A_k} h_k \in M$, this implies that $f \in M$. Then $S(F) \subset M$. Now suppose that $S(F) \neq M$. Then there exist an $f \in M$, an $A \in \mathcal{F}$ with $\mu(A) > 0$, and a $\delta > 0$ such that $\inf_{i,j\geq 1} |f(t) - g_{ij}(t)| \geq \delta$ for $t \in A$. Take an integer i, fixed in the rest of the proof, such that the set $B = A \cap \{t \in T : |f(t) - f_i(t)| < \delta/3\}$ has positive measure, and let $g'_j = \mathbb{1}_B f + \mathbb{1}_{T\setminus B} g_{ij}$, for $j \geq 1$. Since $g'_j \in M$ for $j \geq 1$ and $|f_i(t) - g_{ij}(t)| \geq |f(t) - g_{ij}(t)| - |f(t) - f_i(t)| > 2\delta/3$ it follows that

$$
\begin{aligned}
\|f_i - g_{ij}\|^p - \alpha_i &\geq \|f_i - g_{ij}\|^p - \|f_i - g'_j\|^p \\
&= \int_B \left(|f_i(t) - g_{ij}(t)|^p - |f_i(t) - f(t)|^p\right) \mathrm{d}\mu \\
&\geq [(2\delta/3)^p - (\delta/3)^p] \cdot \mu(B) > 0
\end{aligned}
$$

for $j \geq 1$. If j tends to infinity, we get $\lim_{j\to\infty} \|f_i - g_{ij}\| > \alpha_i$, a contradiction. Thus $M = S(F)$. □

Remark 3.2. The above result is also true for nonempty closed subsets of $\mathbb{L}^p(T,X)$, where X is a separable Banach space. □

Remark 3.3. Similarly as in the proof of Michael's continuous selection theorem, it can be proved that if (X,ρ) is a separable metric space and $(T,\mathcal{F},\mu)$ is a measure space, then every l.s.c. multifunction $F : X \to \text{Cl}(\mathbb{L}^p(T,\mathbb{R}^d))$ with decomposable values admits a continuous selection $f : X \to \mathbb{L}^p(T,\mathbb{R}^d)$.

Proof (Sketch of proof). The proof follows from the following construction procedure. For every $\varepsilon > 0$, we define continuous mappings $f_\varepsilon : X \to \mathbb{L}^p(T,\mathbb{R}^d)$ and $\varphi_\varepsilon : X \to \mathbb{L}^p(T,\mathbb{R}^+)$ such that $F_\varepsilon(x) = \{u \in F(x) : |u(t) - f_\varepsilon(t)| < \varphi_\varepsilon(t)\text{ a.e.}\}$ is nonempty and $\|\varphi_\varepsilon\|_p < \varepsilon$. Now, by the inductive procedure, we can define sequences $(f_n)_{n\geq 0}$, $(\varphi_n)_{n\geq 0}$, and $(F_n)_{n\geq 0}$ such that $\|\varphi_n(x)\| < 1/2^n$, $|f_n(x)(t) - f_{n-1}(x)(t)| \leq \varphi_n(x)(t) + \varphi_{n-1}(x)(t)$ a.e., and $F_n(x) \neq \emptyset$ for $x \in X$.

Hence the existence of a continuous selector f for F follows similarly as in the proof of Michael's theorem. □

Given $C \subset \mathbb{L}^p(T,\mathbb{R}^d)$, by $\mathrm{dec}\{C\}$ we denote the decomposable hull of C, i.e., the smallest decomposable set of $\mathbb{L}^p(T,\mathbb{R}^d)$ containing C. The closed decomposable hull $\overline{\mathrm{dec}}\{C\}$ of C is defined by $\overline{\mathrm{dec}}\{C\} = \mathrm{cl}_{\mathbb{L}}[\mathrm{dec}\{C\}]$. It is easy to see that

$$\mathrm{dec}\{C\} = \left\{ \sum_{i=1}^{m} \mathbb{1}_{A_i} f_i \ : (A_i)_{i=1}^m \in \Pi(T,\mathcal{F}) \text{ and } (f_i)_{i=1}^m \subset C \right\},$$

where $\Pi(T,\mathcal{F})$ denotes the family of all finite $\mathcal{F}$-measurable partitions of T. Immediately from the above definition, it follows that the decomposable hull of the unit ball $\mathcal{B}$ of $\mathbb{L}^p(T,\mathbb{R}^d)$ is equal to the whole space, i.e., $\mathrm{dec}\{\mathcal{B}\} = \mathbb{L}^p(T,\mathbb{R}^d)$. We have the following results dealing with decomposable hulls.

Lemma 3.6. *Let (X,ρ) be a metric space. If $\Gamma : X \to \mathcal{P}(\mathbb{L}^p(T,\mathbb{R}^d))$ is l.s.c., then the multifunction $X \ni x \to \mathrm{dec}\{\Gamma(x)\} \subset \mathbb{L}^p(T,\mathbb{R}^d)$ is also l.s.c.*

Proof. By virtue of ([49], Theorem II.2.8), one has to verify that $\mathrm{dec}(\Gamma)_-(C) := \{x \in X : \mathrm{dec}\{\Gamma(x)\} \subset C\}$ is a closed subset of X for every closed set $C \subset \mathbb{L}^p(T,\mathbb{R}^d)$. Let C be a closed subset of $\mathbb{L}^p(T,\mathbb{R}^d)$ and $(x_n)_{n=1}^\infty$ a sequence of $\mathrm{dec}(\Gamma)_-(C)$ converging to $x \in X$. For every $u \in \mathrm{dec}\{\Gamma(x)\} \subset \overline{\mathrm{dec}}\{\Gamma(x)\}$ and $\varepsilon > 0$, there exist a measurable partition $(A_k^\varepsilon)_{k=1}^{N_\varepsilon}$ of T and a family $(v_k^\varepsilon)_{i=1}^{N_\varepsilon} \subset \mathbb{L}^p(T,\mathbb{R}^d)$ such that $\|u - \sum_{i=1}^{N_\varepsilon} \mathbb{1}_{A_k^\varepsilon} v_k^\varepsilon\| < \varepsilon$ and $v_k^\varepsilon \in \Gamma(x)$ for every $k = 1,\dots,N_\varepsilon$. But Γ is l.s.c. at $x \in X$. Therefore, by virtue of ([49], Theorem II.2.9), for every $k = 1,\dots,N_\varepsilon$ and $\varepsilon > 0$ there exists a sequence $(v_k^{n,\varepsilon})_{n=1}^\infty$ converging to v_k^ε such that $v_k^{n,\varepsilon} \in \Gamma(x_n)$ for every $n \geq 1$, $k = 1,\dots,N_\varepsilon$ and $\varepsilon > 0$. Hence it follows that $\| \sum_{k=1}^{N_\varepsilon} \mathbb{1}_{A_k^\varepsilon} v_k^{n,\varepsilon} - \sum_{k=1}^{N_\varepsilon} \mathbb{1}_{A_k^\varepsilon} v_k^\varepsilon \| \to 0$ as $n \to \infty$ for every $\varepsilon > 0$. Therefore, $\lim_{n\to\infty} \|u - \sum_{k=1}^{N_\varepsilon} \mathbb{1}_{A_k^\varepsilon} v_k^{n,\varepsilon}\| \leq \varepsilon$ for every $\varepsilon > 0$. But $\sum_{i=1}^{N_\varepsilon} \mathbb{1}_{A_k^\varepsilon} v_k^{n,\varepsilon} \in \mathrm{dec}\{\Gamma(x_n)\} \subset C$ for every $n \geq 1$ and $\varepsilon > 0$. Then $u \in C + \varepsilon B$, where B denotes the closed unit ball of $\mathbb{L}^p(T,\mathbb{R}^d)$. Therefore, for every $u \in \mathrm{dec}\{\Gamma(x)\}$, one has $u \in \overline{C} = C$. Thus $\mathrm{dec}\{\Gamma(x)\} \subset C$, which implies that $x \in \mathrm{dec}(\Gamma)_-(C)$. Therefore, $\mathrm{dec}(\Gamma)_-(C)$ is a closed subset of X for every closed set $C \subset \mathbb{L}^p(T,\mathbb{R}^d)$. □

Remark 3.4. Immediately from Lemma 3.6, it follows that by the assumption of Lemma 3.6, the multifunction $X \ni x \to \overline{\mathrm{dec}}\{\Gamma(x)\} \subset \mathbb{L}^p(T,\mathbb{R}^d)$ is l.s.c.

Proof. By virtue of ([49], Theorem II.2.9) one has to verify that for every $x \in X$, every sequence $(x_n)_{n=1}^\infty$ of X converging to x, and $u \in \overline{\mathrm{dec}}\{\Gamma(x)\}$, there exists a sequence $(y_n)_{n=1}^\infty$ of $\mathbb{L}^p(T,\mathbb{R}^d)$ converging to u such that $y_n \in \overline{\mathrm{dec}}\{\Gamma(x_n)\}$ for every $n \geq 1$. Let $x \in X$ be fixed, let $(x_n)_{n=1}^\infty$ be a sequence of X converging to x, and let $u \in \overline{\mathrm{dec}}\{\Gamma(x)\}$. For every $\varepsilon > 0$, one has $\mathrm{dec}\{\Gamma(x)\} \cap B(u,\varepsilon) \neq \emptyset$. By virtue of ([49], Proposition II.2.4) and Lemma 3.6, a multifunction $\Phi(x) = \mathrm{dec}\{\Gamma(x)\} \cap B(u,\varepsilon)$ is l.s.c. Then there exists a sequence $(y_n)_{n=1}^\infty$ of $\mathbb{L}^p(T,\mathbb{R}^d)$

converging to u such that $y_n \in \text{dec}\{\Gamma(x_n)\} \cap B(u, \varepsilon)$, which implies that $y_n \in \overline{\text{dec}}\{\Gamma(x_n)\}$. □

Theorem 3.3. *The decomposable hull of a convex set* $K \subset \mathbb{L}^p(T, \mathbb{R}^d)$ *is itself convex, and its closure is convex and sequentially weakly closed. If* $(\Omega, \mathcal{F}, \mu)$ *is a* σ*-fini te nonatomic space and* K *is a nonempty subset of* $\mathbb{L}^p(\Omega, \mathcal{F}, \mu, \mathbb{R}^d)$, *then* $\overline{\text{dec}}_w\{K\} = \overline{\text{co}}[\text{dec}\{K\}]$, *where* $\overline{\text{dec}}_w\{K\}$ *denotes the closure of* dec$\{K\}$ *with respect to a weak topology of* $\mathbb{L}^p(T, \mathbb{R}^d)$.

Proof. Let K be a convex subset of $\mathbb{L}^p(T, \mathbb{R}^d)$ and $u, v \in \text{dec}\{K\}$. There are partitions $(A_n)_{n=1}^N, (B_m)_{m=1}^M \in \Pi(T, \mathcal{F})$, and $(u_n)_{n=1}^N, (v_m)_{m=1}^M \subset K$ such that $u = \sum_{n=1}^N \mathbb{1}_{A_n} u_n$ and $v = \sum_{m=1}^M \mathbb{1}_{B_m} v_m$. Let $(D_k)_{k=1}^K \in \Pi(T, \mathcal{F})$ be such that $u = \sum_{k=1}^K \mathbb{1}_{D_k} \bar{u}_k$ and $v = \sum_{k=1}^K \mathbb{1}_{D_k} \bar{v}_k$, where $\bar{u}_k = u_{n_k}$ and $\bar{v}_k = v_{m_k}$ for $n_k \in \{1, \ldots, N\}$ and $m_k \in \{1, \ldots, M\}$ for every $k = 1, \ldots, K$. For every $\lambda \in [0, 1]$ and $1 \le k \le K$, one has $\lambda \bar{u}_k + (1-\lambda)\bar{v}_k \in K$. Therefore, $\lambda u + (1-\lambda) v = \sum_{k=1}^K \mathbb{1}_{D_k}[\lambda \bar{u}_k + (1-\lambda)\bar{v}_k] \in \text{dec}\{K\}$. Thus $\text{dec}\{K\}$ is a convex subset of $\mathbb{L}^p(\Omega, \mathcal{F}, \mathbb{R}^r)$. Hence the convexity of $\overline{\text{dec}}_w\{K\}$ follows. Now, immediately from Mazur's theorem ([4], Theorem 9.11), it follows that $\overline{\text{dec}}\{K\}$ is sequentially weakly closed. Finally, immediately from ([41], Theorem 2.3.17), the equality $\overline{\text{dec}}_w\{K\} = \overline{\text{co}}[\text{dec}\{K\}]$ follows. □

Remark 3.5. If $K \subset \mathbb{L}^2(T, \mathbb{R}^d)$ is convex and square integrably bounded, then $\overline{\text{dec}}\{K\}$ is convex and weakly compact.

Proof. If $K \subset \mathbb{L}^2(T, \mathbb{R}^d)$ is square integrably bounded, then $\overline{\text{dec}}\{K\}$ is square integrably bounded, too. Therefore, $\overline{\text{dec}}\{K\}$ is relatively weakly compact, which by virtue of Theorem 3.3, implies that it is convex and weakly compact. □

Remark 3.6. If $F : T \to \mathbb{R}^d$ is measurable and p-integrably bounded, then the interior $\text{Int}[S(F)]$ of $S(F)$ is the empty set and $S(F) = \overline{\text{dec}}\{f_n : n \ge 1\}$, where $f_n \in S(F)$ for $n \ge 1$ are such that $F(t) = \text{cl}\{f_n(t) : n \ge 1\}$ for $t \in T$.

Proof. Suppose $\text{Int}[S(F)] \ne \emptyset$. For every $f \in \text{Int}[S(F)])$, there exists an open ball $\mathcal{B}(f)$ containing f such that $\mathcal{B}(f) \subset \text{Int}[S(F)] \subset S(F)$. Hence it follows that $\text{dec}\{\mathcal{B}(f)\} \subset \text{dec}\{S(F)\}$. But $S(F)$ is a decomposable subset of $\mathbb{L}^p(T, \mathbb{R}^d)$. Therefore, $\text{dec}\{\mathcal{B}(f)\} \subset S(F)$, which is a contradiction, because $S(F)$ is bounded and $\text{dec}\{\mathcal{B}(f)\} = \mathbb{L}^p(T, \mathbb{R}^d)$. Then $\text{Int}[S(F)] = \emptyset$. Let us observe that by the properties of $S(F)$, we have $\overline{\text{dec}}\{f_n : n \ge 1\} \subset S(F)$. On the other hand, by virtue of Lemma 3.2, for every $f \in S(F)$ and $\varepsilon > 0$ there exist a partition $(A_k)_{k=1}^N \in \Pi(T, \mathcal{F})$ and a family $(f_{n_k})_{k=1}^N \subset \{f_n : n \ge 1\}$ such that $\|f - \sum_{k=1}^N \mathbb{1}_{A_k} f_{n_k}\| \le \varepsilon$, which implies that $f \in \overline{\text{dec}}\{f_n : n \ge 1\}$. Thus $S(F) = \overline{\text{dec}}\{f_n : n \ge 1\}$. □

Lemma 3.7. *Assume that* $(T, \mathcal{F}, \mu)$ *and* (X, ρ) *are measure and metric spaces, respectively. Let* $F : T \times X \to \text{Cl}(\mathbb{R}^d)$ *be such that* $F(\cdot, x)$ *is measurable for fixed* $x \in X$ *and there exist* $m, k \in \mathbb{L}^2(T, \mathbb{R}^+)$ *such that* $\|F(t, x)\| \le m(t)$ *and* $h(F(t, x), F(t, \bar{x})) \le k(t)\rho(x, \bar{x})$ *for* μ*-a.e.* $t \in T$ *and* $x, \bar{x} \in X$.

Then $H(S(F(\cdot,x)), S(F(\cdot,\bar{x}))) \le K\rho(x,\bar{x})$ *for every* $x,\bar{x} \in X$, *where* $K = (\int_T k^2(t)\mathrm{d}\mu)^{1/2}$ *and* H *is the Hausdorff metric on* $\mathrm{Cl}(\mathbb{L}^2(T,\mathbb{R}^d))$.

Proof. Assume $x,\bar{x} \in X$ and select arbitrarily $f^x \in S(F(\cdot,x))$. By virtue of Theorem 3.1, one has

$$\begin{aligned}\mathrm{dist}(f^x, S(F(\cdot,\bar{x})) &= \inf\left\{\left(\int_T |f_t^x - f_t|^2\mathrm{d}\mu\right)^{1/2} : f \in S(F(\cdot,\bar{x}))\right\}\\ &= \left(\int_T \mathrm{dist}^2(f_t^x, F(t,\bar{x}))\mathrm{d}\mu\right)^{1/2}\\ &\le \left(\int_T k^2(t)\rho^2(x,\bar{x})\mathrm{d}\mu\right)^{1/2} \le K\rho(x,\bar{x}),\end{aligned}$$

where $K = (\int_0^T k^2(t)\mathrm{d}t)^{1/2}$. Then $\bar{H}(S(F(\cdot,x)), S((F(\cdot,\bar{x}))) \le K\rho(x,\bar{x})$. In a similar way, we obtain $\bar{H}(S(F(\cdot,\bar{x})), S(F(\cdot,x))) \le K\rho(x,\bar{x})$. □

Remark 3.7. Similarly as above, one can prove that if $(T,\mathcal{F},\mu)$ and (X,ρ) are as above and $F : T \times X \to \mathrm{Cl}(\mathbb{R}^d)$ is measurable and uniformly square integrably bounded such that $F(t,\cdot)$ is l.s.c. for a.e. fixed $t \in T$, then a set-valued mapping $X \ni x \to S(F(\cdot,x)) \in \mathrm{Cl}(\mathbb{L}^2(T,\mathbb{R}^d))$ is l.s.c.

Proof. Let us observe first that for given metric spaces X and Y, a multifunction $\Phi : X \to \mathcal{P}(Y)$ is l.s.c. at $\bar{x} \in X$ if it is H-l.s.c., i.e., if for every $\varepsilon > 0$, there exists a $\delta > 0$ such that for every $x \in X$ satisfying $\rho(x,\bar{x}) < \delta$, one has $\bar{h}(\Phi(\bar{x}), \Phi(x)) \le \varepsilon$. Indeed, suppose the above condition is satisfied and Φ is not l.s.c. at $\bar{x}$. There exists an open set $U \subset Y$ with $\Phi(\bar{x}) \cap U \ne \emptyset$ such that in every neighborhood V of $\bar{x}$, there exists $\tilde{x} \in V$ such that $\Phi(\tilde{x}) \cap U = \emptyset$. Therefore, we can select a sequence $(x_n)_{n=1}^\infty$ of X converging to $\tilde{x}$ such that $\Phi(x_n) \cap U = \emptyset$ for every $n = 1,2,\ldots$. On the other hand, for every $\varepsilon > 0$, there exists $N_\varepsilon \ge 1$ such that for every $n \ge N_\varepsilon$, we have $\Phi(\bar{x}) \subset V^0[\Phi(x_n),\varepsilon]$. Hence in particular, it follows that $\Phi(\bar{x}) \cap U \subset V^0[\Phi(x_n),\varepsilon]$ for $n \ge N_\varepsilon$. Let $y \in \Phi(\bar{x}) \cap U$, $n_k = N_{1/k}$ for every $k = 1,2,\ldots$ and select $y_k \in \Phi(x_{n_k})$ such that $d(y_k, y) < 1/k$. For k sufficiently large, we have $y_k \in U$ and therefore $\Phi(x_{n_k}) \cap U \ne \emptyset$, a contradiction.

Let us observe now that if $\Phi(\bar{x})$ is a compact subset of Y, then Φ is l.s.c. at $\bar{x} \in X$ if and only if for every $\varepsilon > 0$, there exists a $\delta > 0$ such that for every $x \in X$ satisfying $\rho(x,\bar{x}) < \delta$, one has $\bar{h}(\Phi(\bar{x}), \Phi(x)) \le \varepsilon$. Indeed, for $i = 1,\ldots,m$, let y_i be such that $\{B^0(y_i,(1/2)\varepsilon) : i = 1,\ldots,m\}$ covers $\Phi(\bar{x})$ and for $i = 1,\ldots,m$, let $\delta_i > 0$ be such that $\rho(x,\bar{x}) < \delta_i$ implies $\Phi(x) \cap B^0(y_i,(1/2)\varepsilon) \ne \emptyset$. Let $\delta = \min\{\delta_i : i = 1,\ldots,m\}$. Then $\rho(x,\bar{x}) < \delta$ implies that $y_i \in V^0(\Phi(x),(1/2)\varepsilon)$ for $i = 1,\ldots,m$, i.e., $B^0(y_i,(1/2)\varepsilon) \subset V^0(\Phi(x),(1/2)\varepsilon)$ for all $i = 1,\ldots,m$. Therefore, $\Phi(\bar{x}) \subset \bigcap_{i=1}^m B^0(y_i,(1/2)\varepsilon) \subset V^0(\Phi(x),(1/2)\varepsilon)$ for $x \in B^0(\bar{x},\delta)$, which is equivalent to $\bar{h}[\Phi(\bar{x}), \Phi(x)] \le \varepsilon$ for $x \in B^0(\bar{x},\delta)$.

Let $m \in \mathbb{L}^2(T, \mathbb{R}^+)$ be such that $\|F(t,x)\| \leq m(t)$ for every $x \in X$ and a.e. $t \in T$. Therefore, $F(t,x)$ is a compact subset of $\mathbb{R}^d$ for every $x \in X$ and a.e. $t \in T$. Similarly as in the proof of Lemma 3.7, we can verify that for every $\bar{x}, x \in X$, one has

$$\bar{H}[S(F(\cdot,\bar{x})), S(F(\cdot,x))] \leq \left(\int_T \bar{h}^2[F(t,\bar{x}), F(t,x)]\mathrm{d}t\right)^{\frac{1}{2}}.$$

Thus for every $\bar{x} \in X$ and every sequence $(x_n)_{n=1}^{\infty}$ of X converging to $\bar{x}$, we obtain $\int_T \bar{h}^2[F(t,\bar{x}), F(t,x_n)]\mathrm{d}t \to 0$ as $n \to \infty$, which implies that $\bar{H}[S(F(\cdot,\bar{x})), S(F(\cdot,x_n))] \to 0$ as $n \to \infty$. Then the set-valued mapping $X \ni x \to S(F(\cdot,x)) \in \mathrm{Cl}(\mathbb{L}^2(T,\mathbb{R}^d))$ is l.s.c. at $\bar{x}$. $\square$

Lemma 3.8. *Assume that T is an interval of the real line and let $F : T \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : T \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ be measurable uniformly p-integrably bounded and such that $F(t,\cdot)$ and $G(t,\cdot)$ are l.s.c. for fixed $t \in T$. There are continuous functions $u : \mathbb{R}^d \to \mathbb{L}^p(T,\mathbb{R}^d)$ and $v : \mathbb{R}^d \to \mathbb{L}^p(T,\mathbb{R}^{d\times m})$ such that*

(i) $u(x) \in S(F(\cdot,x))$ *and* $v(x) \in S(G(\cdot,x))$ *for* $x \in \mathbb{R}^d$;
(ii) *mappings* $f : T \times \mathbb{R}^d \ni (t,x) \to u(x)(t) \in \mathbb{R}^d$ *and* $g : T \times \mathbb{R}^d \ni (t,x) \to v(x)(t) \in \mathbb{R}^{d\times m}$ *are* $\beta_T \otimes \beta(\mathbb{R}^d)$*-measurable such that* $f(t,x) \in F(t,x)$ *and* $g(t,x) \in G(t,x)$ *for a.e.* $t \in T$ *and* $x \in \mathbb{R}^d$.

Proof. The existence of continuous functions u and v satisfying (i) follows immediately from Remarks 3.3 and 3.7. Let $\mathcal{I}$ be the identity mapping on T and define $(\mathcal{I} \times u) : T \times \mathbb{R}^d \to T \times \mathbb{L}^p(T,\mathbb{R}^d)$ by setting $(\mathcal{I} \times u)(t,x) = (t, u(x))$ for $(t,x) \in T \times \mathbb{R}^d$. The function $\mathcal{I} \times u$ is continuous on $T \times \mathbb{R}^d$ and therefore $(\beta_T \otimes \beta(\mathbb{R}^d), \beta_T \otimes \beta(\mathbb{L}^p))$-measurable, where β_T, $\beta(\mathbb{R}^d)$ and $\beta(\mathbb{L}^p)$ denote the Borel σ-fields on T, $\mathbb{R}^d$ and $\mathbb{L}^p(T,\mathbb{R}^d)$, respectively. Let $\rho : T \times \mathbb{L}^p(T,\mathbb{R}^d) \to \mathbb{R}^d$ be defined by $\rho(t,z) = z(t)$ for $(t,z) \in T \times \mathbb{L}^p(T,\mathbb{R}^d)$. The mapping ρ is $(\beta_T \otimes \beta(\mathbb{L}^p), \beta(\mathbb{R}^d))$-measurable because ρ is such that $\rho(t,\cdot)$ is continuous and $\rho(\cdot,z)$ is measurable for fixed $t \in T$ and $z \in \mathbb{L}^p(T,\mathbb{R}^d)$, respectively. Hence it follows that a mapping $f : T \times \mathbb{R}^d \ni (t,x) \to u(x)(t) \in \mathbb{R}^d$ is measurable on $T \times \mathbb{R}^d$, i.e., is $(\beta_T \otimes \beta(\mathbb{R}^d), \beta(\mathbb{R}^d))$-measurable because $f(t,x) = [\rho \circ (\mathcal{I} \times u)](t,x) = \rho(t, u(x))$ for $(t,x) \in T \times \mathbb{R}^d$. Measurability of a mapping g can be verified in a similar way. It is clear that $f(t,x) \in F(t,x)$ and $g(t,x) \in G(t,x)$ for a.e. $t \in T$ and $x \in \mathbb{R}^d$. $\square$

Similarly as above, let T be an interval of the real line. Denote by J the linear mapping defined on $\mathbb{L}^p(T,\mathbb{R}^d)$ by setting $J(f) = \int_T f(t)\mathrm{d}t$ for $f \in \mathbb{L}^p(T,\mathbb{R}^d)$. For a nonempty set $K \subset \mathbb{L}^p(T,\mathbb{R}^d)$, by $J(K)$ we denote its image by the mapping J, i.e., a set of the form $\{\int_T f(t)\mathrm{d}t : f \in K\}$.

Lemma 3.9. *If $K \subset \mathbb{L}^p(T,\mathbb{R}^d)$ is nonempty decomposable, then $J(K)$ is a nonempty convex subset of $\mathbb{R}^d$.*

Proof. Let $z_1, z_2 \in J(K)$ and $\lambda \in [0,1]$. There exist $f_1, f_2 \in K$ such that $z_1 = \int_T f_1(t)\mathrm{d}t$ and $z_2 = \int_T f_2(t)\mathrm{d}t$. Let $\mathcal{L}_T$ be the family of all Lebesgue measurable subsets of T and put $\mu(E) = (\int_E f_1(t)\mathrm{d}t, \int_E f_2(t)\mathrm{d}t)$ for $E \in \mathcal{L}_T$. By Lyapunov's theorem, $\mu(\mathcal{L}_T)$ is a convex compact subset of $\mathbb{R}^{2d}$. Since $(0,0)$ and (z_1, z_2) belong to $\mu(\mathcal{L}_T)$, then we have also $(\lambda z_1, \lambda z_2) \in \mu(\mathcal{L}_T)$. Therefore, there exists $H \in \mathcal{L}_T$ such that $(\lambda z_1, \lambda z_2) = \mu(H)$, which by the definition of the measure μ implies that $\lambda z_1 = \int_T \mathbb{1}_H f_1(t)\mathrm{d}t$ and $\lambda z_2 = \int_T \mathbb{1}_H f_2(t)\mathrm{d}t$. Let $f = \mathbb{1}_H f_1 + \mathbb{1}_{T\setminus H} f_2$. By the decomposability of K, we have $f \in K$. Therefore, $\int_T f(t)\mathrm{d}t \in J(K)$. But $\int_T f(t)\mathrm{d}t = \int_T (\mathbb{1}_H f_1 + \mathbb{1}_{T\setminus H} f_2)(t)\mathrm{d}t = \int_T \mathbb{1}_H (f_1 - f_2)(t)\mathrm{d}t + \int_T f_2(t)\mathrm{d}t = \lambda z_1 - \lambda z_2 + z_2 = \lambda z_1 + (1-\lambda) z_2$. Then $\lambda z_1 + (1-\lambda) z_2 \in J(K)$. □

For $F \in \mathcal{A}(T, \mathbb{R}^d)$, the set $J(S(F))$ is denoted by $\int_T F(t)\mathrm{d}t$ and is said to be the Aumann integral of F on the interval T.

Corollary 3.4. *For every $F \in \mathcal{A}(T, \mathbb{R}^d)$, the Aumann integral $\int_T F(t)\mathrm{d}t$ is a nonempty convex subset of $\mathbb{R}^d$. If furthermore, F is p-integrably bounded, then $\int_T F(t)\mathrm{d}t$ is a bounded subset of $\mathbb{R}^d$.* □

Denote by $V(\sigma^r)$ the set of $r+1$ vertices of the $(r+1)$-dimensional simplex $\sigma^r = \{(\xi_0, \ldots, \xi_r) \in \mathbb{R}^{r+1} : 0 \le \xi_i \le 1, \sum_{i=0}^r \xi_i = 1\}$. It is clear that if $u_i \in \mathbb{L}^\infty(T, \mathbb{R}^1)$ for $i = 0, 1, \ldots, r$, then $(u_0, \ldots, u_r) \in \mathbb{L}^\infty(T, \mathbb{R}^{r+1})$, where $\mathbb{L}^\infty(T, \mathbb{R}^1)$ consists of all μ-essentially bounded measurable scalar functions defined on T.

Lemma 3.10. *Let $Y(t)$ be an $n \times (r+1)$-matrix-valued function with components in $\mathbb{L}^\infty(T, \mathbb{R}^1)$, $\Psi = \{u \in \mathbb{L}^\infty(T, \mathbb{R}^{r+1}) : u(t) \in \sigma^r$ for $t \in T\}$, and $\Psi_0 = \{u \in \mathbb{L}^\infty(T, \mathbb{R}^{r+1}) : u(t) \in V(\sigma^r)$ for $t \in T\}$. Then $\{\int_T Y(t) \cdot u(t)\mathrm{d}t : u \in \Psi\} = \{\int_T Y(t) \cdot u(t)\mathrm{d}t : u \in \Psi_0\}$, and both of these sets are compact and convex.*

Proof. Let $J(u) = \int_T Y(t) \cdot u(t)\mathrm{d}t$ for $u \in \mathbb{L}^\infty(T, \mathbb{R}^{r+1})$. Clearly, Ψ is convex and bounded in the $\mathbb{L}^\infty(T, \mathbb{R}^{r+1})$-norm topology. Hence if we can show that Ψ is weakly*-closed, it will imply that Ψ is weakly*-compact. Suppose u^0 is a weak*-limit of a sequence of Ψ that does not belong to Ψ. Then there is a set $E \subset T$ of positive measure such that $u^0(t) \in \sigma^r$ for $t \in E$ and $u^0 \in \Psi$. One may readily establish the existence of an $\varepsilon > 0$ and $\eta \in \mathbb{R}^{r+1}$ such that the inner product satisfies $\langle \eta, \xi \rangle \ge C$ if $\xi \in \sigma^r$ and $\langle \eta, u^0(t) \rangle < C - \varepsilon$ for t in a subset E_1 of E having a positive measure $\mu(E_1)$. Define a function $w(t) = (w_0(t), \ldots, w_r(t))$ by setting

$$w_i(t) = \begin{cases} \eta_i / \mu(E_1) & \text{for } t \in E_1, \\ 0 \quad \text{for} & t \notin E_1, \end{cases}$$

for $i = 1, \ldots, r$. It is clear that $w \in \mathbb{L}^\infty(T, \mathbb{R}^{r+1})$. From the properties of $\eta \in \mathbb{R}^{r+1}$, it follows that w separates u^0 and Ψ, contradicting u^0 being a weak*-limit of a sequence of Ψ. Thus Ψ is closed, convex, and weak*-compact. It is easily seen that J is weak*-continuous, because the weak topology was defined so that the linear functionals that were continuous on a given normed space X with its norm topology are still continuous when X has its weak topology. In particular, $J = (J_1, \ldots, J_n)$ is a continuous linear mapping from X^* taken with its norm topology to $\mathbb{R}^d$ such that components J_i of J are representable as elements of X. Then J is continuous as a mapping of X^* with the weak*-topology to $\mathbb{R}^d$. Therefore, $J\Psi = \{Ju : u \in \Psi\}$ is a compac,t convex subset of $\mathbb{R}^d$. Clearly, $J\Psi_0 \subset J\Psi$. Similarly as in the proof of Lyapunov's theorem, we can also show that $J\Psi \subset J\Psi_0$. □

Lemma 3.11. *Let* $F : T \to \mathrm{Cl}(\mathbb{R}^d)$ *be measurable and integrably bounded. Then* $\int_T F(t)\mathrm{d}t = \int_T \mathrm{co}\, F(t)\mathrm{d}t$ *, and both sets are nonempty and convex in* $\mathbb{R}^d$.

Proof. The nonemptiness and convexity of $\int_T F(t)\mathrm{d}t$ follow from Corollary 3.4. By the definition of the Aumann integral, it follows that $\int_T F(t)\mathrm{d}t \subset \int_T \mathrm{co}\, F(t)\mathrm{d}t$. Suppose $y \in \int_T \mathrm{co}\, F(t)\mathrm{d}t$, and let $f \in S(\mathrm{co}\, F)$ be such that $y = \int_T f(t)\mathrm{d}t$. By Carathéodory's theorem, for every $t \in T$, the point $f(t) \in \mathrm{co}\, F(t)$ may be expressed as a convex combination $f(t) = \sum_{i=0}^d \xi_i(t) f^i(t)$ with $f^i(t) \in F(t)$, $0 \le \xi_i(t) \le 1$, and $\sum_{i=0}^d \xi_i(t) = 1$. Let σ^d denote the simplex in the space $\mathbb{R}^{d+1}$, i.e., $\sigma^d = \{(\xi_0, \ldots, \xi_d) \in \mathbb{R}^{n+1} : 0 \le \xi_i \le 1, \sum_{i=0}^d \xi_i = 1\}$. Denote by $\xi(t)$ the vector $(\xi_0(t), \ldots, \xi_d(t)) \in \sigma^d$. Let us observe that the functions ξ_i and f^i can be chosen to be measurable. Indeed, let $g(t, \xi, \beta^0, \ldots, \beta^d) = \sum_{i=0}^{\mathrm{d}} \xi_i(t)\beta^i$ for $t \in T$ and $\beta^0, \ldots, \beta^d \in \mathbb{R}$ and let $\Gamma(t) = \sigma^{d+1} \times F(t) \times \cdots \times F(t)$ with $F(t)$ appearing $n+1$ times in the product. Since f is measurable and $f(t) \in g(t, \Gamma(t))$ for a.e. $t \in T$, then by Theorem 2.5, there exists a measurable function $T \ni t \to (\xi_0(t), \ldots, \xi_n(t), f^0(t), \ldots, f^{\mathrm{d}}(t)) \in \Gamma(t)$ such that $f(t) = g(t, (\xi_0(t), \ldots, \xi_n(t), f^0(t), \ldots, f^{\mathrm{d}}(t))$ for a.e. $t \in T$. Let the vectors $f^i(t)$ be the columns of an $d \times (d+1)$-matrix Y. By virtue of Lemma 3.10 there exists a measurable vector function $\xi^* = (\xi_0^*, \ldots, \xi_d^*)$ on T taking values in the vertices of the simplex σ^d such that $\int_T f(t)\mathrm{d}t = \int_T Y(t) \cdot \xi(t)\mathrm{d}t = \int_T Y(t) \cdot \xi^*(t)\mathrm{d}t$. Now $\xi_i^*(T) \subset \{0, 1\}$ for all $i = 0, 1, \ldots, d$ and $\sum_{i=0}^d \xi_i^*(t) = 1$. Let $T_i = \{t \in T : \xi_i^*(t) = 1\}$. Then T_i is measurable and $\bigcup_{i=0}^d T_i = T$ and $T_i \cap T_j = \emptyset$ for $i \ne j$. Define $f^*(t) = f^i(t)$ for $t \in T_i$ for $i = 0, 1, \ldots, d$. It is clear that f^* is measurable and such that $f^*(t) \in F(t)$ and $\int_T f^*(t)\mathrm{d}t = \int_T f(t)\mathrm{d}t$. Then $\int_T F(t)\mathrm{d}t = \int_T \mathrm{co}\,\mathrm{F}(t)\mathrm{d}t$. □

Theorem 3.4 (Aumann). *If* $F : T \to \mathrm{Cl}(\mathbb{R}^d)$ *is measurable and integrably bounded, then* $\int_T F(t)\mathrm{d}t = \int_T \mathrm{co}\, F(t)\mathrm{d}t$ *, and both integrals are nonempty convex, compact subsets of* $\mathbb{R}^d$.

Proof. By virtue of Lemma 3.11, we have $\int_T F(t)\mathrm{d}t = \int_T \mathrm{co}\, F(t)\mathrm{d}t$, and both integrals are nonempty convex subsets of $\mathbb{R}^d$. By virtue of Remark 3.1, a set $S(\mathrm{co}\, F)$ is a weakly sequentially compact subset of $\mathbb{L}(T, \mathbb{R}^d)$. By the definition of the Aumann integral, we have $\int_T \mathrm{co}\, F(t)\mathrm{d}t = J(S(\mathrm{co}\, F))$, where J is a linear and continuous mapping defined on $\mathbb{L}(T, \mathbb{R}^d)$. By the linearity of J , it follows that J is also continuous on $\mathbb{L}(T, \mathbb{R}^d)$ with respect to its weak topology. Therefore, $J(S(\mathrm{co}\, F))$ is a compact subset of $\mathbb{R}^d$. □

Remark 3.8. It can be proved that if $(X, \|\cdot\|)$ is a separable Banach space, T is an interval of the real line, and $F : T \to \mathrm{Cl}(X)$ is measurable and integrably bounded, then $\mathrm{cl}(\int_T F(t)\mathrm{d}t) = \mathrm{cl}(\int_T \mathrm{co}\, F(t)\mathrm{d}t)$, where the closure is taken in the norm topology of X. □

Theorem 3.5. *If $F : T \to \mathrm{Cl}(\mathbb{R}^d)$ is measurable and integrably bounded, then for every $p \in \mathbb{R}^d$ and $A \in \mathcal{L}_T$, one has $\int_A \sigma(p, F(t))\mathrm{d}t = \sigma(p, \int_A F(t)\mathrm{d}t)$.*

Proof. Let us observe that $\sigma(p, F(\cdot))$ is measurable and integrably bounded for every fixed $p \in \mathbb{R}^d$. Then it is integrable and $\int_A \sigma(p, F(t))\mathrm{d}t < \infty$ for every $p \in \mathbb{R}^d$ and $A \in \mathcal{L}_T$. For every $f \in S(F)$ and $p \in \mathbb{R}^d$, we have $\left\langle p, \int_A f(t)\mathrm{d}t \right\rangle = \int_A \langle p, f(t)\rangle\, \mathrm{d}t \le \int_A \sigma(p, F(t))\mathrm{d}t$. Therefore, for every $p \in \mathbb{R}^d$, one has $\sigma(p, \int_A F(t)\mathrm{d}t) \le \int_A \sigma(p, F(t))\mathrm{d}t$. We shall show now that for every $\alpha \in \mathbb{R}$ and $p \in \mathbb{R}^d$ such that $\alpha < \int_A \sigma(p, F(t))d$, there is $f \in S(F)$ such that $\alpha < \sigma(p, \int_A f(t)\mathrm{d}t)$. Indeed, let us take arbitrarily $g \in S(F)$ and define for every $n \ge 1$ a multifunction F_n by setting $F_n(t) = \{x \in F(t) : |x - g(t)| < n\}$. Similarly as in the proof of Theorem 2.5, we can verify that F_n, and hence also $\mathrm{cl}(F_n)$, is measurable. Then $\sigma(p, F_n(\cdot))$ is measurable for every $p \in \mathbb{R}^d$ and $n \ge 1$. It is also integrably bounded. Furthermore, $\sigma(p, F_n(t)) \to \sigma(p, F(t))$ for $t \in T$ as $n \to \infty$. Then $\int_A \sigma(p, F_n(t))\mathrm{d}t \to \int_A \sigma(p, F(t))\mathrm{d}t$ for every $p \in \mathbb{R}^d$ as $n \to \infty$. Thus we have $\alpha < \int_A \sigma(p, F_n(t))\mathrm{d}t$ for n large enough. Then there exists an integrable function $\varphi : T \to \mathbb{R}$ such that $\alpha < \int_A \varphi(t)\mathrm{d}t$ and $\varphi(t) < \sigma(p, F_n(t))$ for a.e. $t \in T$. Let $G(t) = \{x \in F(t) : \langle p, x\rangle > \varphi(t)\}$ for $t \in T$. It is clear that $G(t) \ne \emptyset$ and that G has a measurable graph. Therefore, by virtue of Remark 2.9, there exists a measurable selector f of G, and hence also of F, such that $\varphi(t) < \langle p, f(t)\rangle$. Thus $\int_A \varphi(t)\mathrm{d}t < \left\langle p, \int_A f(t)\mathrm{d}t \right\rangle$. Hence it follows that $\alpha < \left\langle p, \int_A f(t)\mathrm{d}t \right\rangle$. Now taking in particular $\alpha_n = \int_A \sigma(p, F(t))\mathrm{d}t - 1/n$ for $n \ge 1$, we can select $f_n \in S(F)$ such that $\alpha_n < \sigma(p, \int_A f_n(t)\mathrm{d}t) \le \sigma(p, \int_A F(t)\mathrm{d}t)$ for every $p \in \mathbb{R}^d$ and $n \ge 1$, which implies that $\int_A \sigma(p, F(t))\mathrm{d}t \le \sigma(p, \int_A F(t)\mathrm{d}t)$ for every $p \in \mathbb{R}^d$ and $A \in \mathcal{L}_T$. □

Remark 3.9. The above results are also true for measurable and p-integrably bounded multifunctions with $p \ge 1$. □

4 Set-Valued Stochastic Processes

Similarly as in Chap. 1, we assume that we are given a complete filtered probability space $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\ge 0}$ satisfying the usual conditions. By a set-valued random variable, we mean an $\mathcal{F}$-measurable multifunction $\mathcal{Z} : \Omega \to \mathrm{Cl}(\mathbb{R}^d)$. If $\mathcal{Z} \in \mathcal{A}(\Omega, \mathbb{R}^d)$, then the Aumann integral $\int_\Omega \mathcal{Z}\mathrm{d}P$ is denoted by $E[\mathcal{Z}]$ and is said to be the mean value of the set-valued random variable $\mathcal{Z}$. A set-valued random variable $\mathcal{Z} \in \mathcal{A}(\Omega, \mathbb{R}^d)$ is said to be Aumann integrable. Immediately from properties of measurable set-valued mappings, the following results, dealing with set-valued random variables, follow.

Lemma 4.1. *Let $\mathcal{Z} : \Omega \to \mathrm{Cl}(\mathbb{R}^d)$ be an Aumann integrable set-valued random variable. Then*

(i) *$S(\mathcal{Z})$ is a closed decomposable subset of $\mathbb{L}^p(\Omega, \mathcal{F}, \mathbb{R}^d)$ and $S(\overline{\mathrm{co}}\,\mathcal{Z}) = \overline{\mathrm{co}}\,S(\mathcal{Z})$.*

(ii) *$\mathcal{Z}$ is p-integrably bounded if and only if $S(\mathcal{Z})$ is a bounded subset of $\mathbb{L}^p(\Omega, \mathcal{F}, \mathbb{R}^d)$.*

(iii) *If $\mathcal{Z}$ is p-integrably bounded, then $\mathrm{Int}[S(\mathcal{Z})] = \emptyset$ and $S(\mathcal{Z}) \neq \mathbb{L}^p(\Omega, \mathcal{F}, \mathbb{R}^r)$.*

(iv) *There exists a sequence $(z_n)_{n=1}^{\infty}$ of d-dimensional random variables such that $z_n(\omega) \in \mathcal{Z}(\omega)$ and $\mathcal{Z}(\omega) = \mathrm{cl}\{z_n(\omega) : n \geq 1\}$ for $n \geq 1$ and $\omega \in \Omega$. If $\{z_n : n \geq 1\} \subset S(\mathcal{Z})$, then $S(\mathcal{Z}) = \overline{\mathrm{dec}}\{z_n(\omega) : n \geq 1\}$.*

(v) *If $(z_n)_{n=1}^{\infty} \subset S(\mathcal{Z})$ is such that $\mathcal{Z}(\omega) = \mathrm{cl}\{z_n(\omega) : n \geq 1\}$ for $\omega \in \Omega$, then for every $z \in S(\mathcal{Z})$ and every $\varepsilon > 0$, there exist a partition $(A_k)_{k=1}^{N} \in \Pi(\Omega, \mathcal{F})$ and a family $(z_{n_k})_{k=1}^{N} \subset \{z_n : n \geq 1\}$ such that $E|z - \sum_{k=1}^{N} \mathbb{1}_{A_k} z_{n_k}| \leq \varepsilon$.*

(vi) *If F and G are Aumann integrable set-valued random variables such that $S(F) = S(G)$, then $F(\omega) = G(\omega)$ for a.e. $\omega \in \Omega$.*

(vii) *If $\mathcal{Z}$ is convex-valued and square integrably bounded, then $S(\mathcal{Z})$ is a decomposable, convex, and weakly compact subset of $\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{R}^d)$.*

(viii) *If F and G are convex-valued and integrably bounded set-valued random variables, then $S(F + G) = S(F) + S(G)$.*

A family $\Phi = (\Phi_t)_{t \geq 0}$ of set-valued random variables $\Phi_t : \Omega \to \mathrm{Cl}(\mathbb{R}^q)$ is called a set-valued stochastic process. Similarly as in the case of point-valued stochastic processes, a set-valued process $\Phi = (\Phi_t)_{t \geq 0}$ can also be defined as a set-valued mapping $\Phi : \mathbb{R}^+ \times \Omega \ni (t, \omega) \to \Phi_t(\omega) \in \mathrm{Cl}(\mathbb{R}^q)$ such that $\Phi(t, \cdot)$ is a set-valued random variable for every $t \geq 0$. If such a multifunction Φ is $\beta(\mathbb{R}^+) \otimes \mathcal{F}$-measurable, then a set-valued process Φ is said to be measurable. If furthermore, for every $t \geq 0$, the set-valued mapping Φ_t is $\mathcal{F}_t$-measurable, then Φ is said to be $\mathbb{F}$-nonanticipative. It is easy to see that Φ is $\mathbb{F}$-nonanticipative if and only if it is $\Sigma_{\mathbb{F}}$-measurable, where $\Sigma_{\mathbb{F}} = \{A \in \beta_T \otimes \mathcal{F} : A^t \in \mathcal{F}_t \text{ for } t \in T\}$, and A^t denotes the t-section of a set $A \subset T \times \Omega$. Given $p \geq 1$, we call a set-valued process $\Phi = (\Phi_t)_{t \geq 0}$ p-integrably bounded if there exists $m \in \mathbb{L}^p(\mathbb{R}^+ \times \Omega, \mathbb{R}^+)$ such that $\|\Phi_t(\omega)\| \leq m(t, \omega)$ for a.e. $(t, \omega) \in \mathbb{R}^+ \times \Omega$. A set-valued process $\Phi = (\Phi_t)_{t \geq 0}$ is said to be bounded if there exists a number $M > 0$ such that $\|\Phi_t(\omega)\| \leq M$ for a.e. $(t, \omega) \in \mathbb{R}^+ \times \Omega$. It is clear that every bounded set-valued process is p-integrably bounded for every $p \geq 1$. Similarly as above, by $S(\Phi)$ we denote the subtrajectory integrals of a set-valued stochastic process $\Phi : \mathbb{R}^+ \times \Omega \to \mathrm{Cl}(\mathbb{R}^q)$, i.e., the set of all measurable and $\mathrm{d}t \times P$-integrable selectors of Φ. By $S_{\mathbb{F}}(\Phi)$ we denote the subset of $S(\Phi)$ containing all $\mathbb{F}$-nonanticipative elements of $S(\Phi)$. If Φ is an p-integrably bounded set-valued process defined on $[0, T] \times \Omega$, its subtrajectory integrals will be denoted by $S(\Phi)$ for every $p \geq 1$. In this case, $S(\Phi) \subset \mathbb{L}^p([0, T] \times \Omega, \beta_T \otimes \mathcal{F}_T, \mathbb{R}^q)$. Similarly, if $\Phi : [0, T] \times \Omega \to \mathrm{Cl}(\mathbb{R}^q)$ is $\mathbb{F}$-nonanticipative and square integrably bounded, then $S_{\mathbb{F}}(\Phi) \subset \mathbb{L}^2([0, T] \times \Omega, \Sigma_{\mathbb{F}}, \mathbb{R}^q)$. Similarly as above, Φ is said to be Aumann

(Itô) integrable if $S(\Phi) \neq \emptyset$ ($S_{\mathbb{F}}(\Phi) \neq \emptyset$). We shall consider set-valued stochastic processes with $q = d$ and $q = d \times m$.

Let us denote by $\mathcal{M}(\mathbb{R}^{+} \times \Omega, \mathbb{R}^{d})$ and $\mathcal{M}_{\mathbb{F}}(\mathbb{R}^{+} \times \Omega, \mathbb{R}^{d})$ the spaces of all measurable and $\mathbb{F}$-nonanticipative, respectively set-valued stochastic, processes on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ with values in $\mathrm{Cl}(\mathbb{R}^{d})$. Similarly, the space of all $\mathbb{F}$-nonanticipative processes on $(\Omega, \mathcal{F}, \mathbb{F}, P)$ with values in $\mathrm{Cl}(\mathbb{R}^{d\times m})$ will be denoted by $\mathcal{M}_{\mathbb{F}}(\mathbb{R}^{+} \times \Omega, \mathbb{R}^{d\times m})$. We denote by $\mathcal{L}^{2}(\Omega, \mathbb{R}^{d})$ the space of all (equivalence classes of) set-valued random variables $\mathcal{Z} : \Omega \to \mathrm{Cl}(\mathbb{R}^{d})$ such that $E\|\mathcal{Z}\|^{2} < \infty$, where $\|\mathcal{Z}\|(\omega) = \sup\{|x| : x \in \mathcal{Z}(\omega)\}$ for a.e. $\omega \in \Omega$. Elements of the space $\mathcal{L}^{2}(\Omega, \mathbb{R}^{d})$ are called $\mathbb{R}^{d}$-set-valued square integrably bounded random variables. We shall consider $\mathcal{L}^{2}(\Omega, \mathbb{R}^{d})$ as a metric space with a metric H defined by $H(\mathcal{Z}_{1}, \mathcal{Z}_{2}) = [Eh^{2}(\mathcal{Z}_{1}(\cdot), \mathcal{Z}_{2}(\cdot))]^{1/2}$ for $\mathcal{Z}_{1}, \mathcal{Z}_{2} \in \mathcal{L}^{2}(\Omega, \mathbb{R}^{d})$. Similarly as in the case of $\mathbb{L}^{2}(\Omega, \mathcal{F}, \mathbb{R}^{d})$, it can be verified that $(\mathcal{L}^{2}(\Omega, \mathbb{R}^{d}), H)$ is a complete metric space. By $\mathcal{L}^{2}_{\mathbb{F}}(\mathbb{R}^{+} \times \Omega, \mathbb{R}^{d})$ and $\mathcal{L}^{2}_{\mathbb{F}}(\mathbb{R}^{+} \times \Omega, \mathbb{R}^{d\times m})$ we shall denote the spaces of all square integrably bounded elements of spaces $\mathcal{M}_{\mathbb{F}}(\mathbb{R}^{+} \times \Omega, \mathbb{R}^{d})$ and $\mathcal{M}_{\mathbb{F}}(\mathbb{R}^{+} \times \Omega, \mathbb{R}^{d\times m})$, respectively. Similarly as above, the spaces $\mathcal{L}^{2}_{\mathbb{F}}(\mathbb{R}^{+} \times \Omega, \mathbb{R}^{d})$ and $\mathcal{L}^{2}_{\mathbb{F}}(\mathbb{R}^{+} \times \Omega, \mathbb{R}^{d\times m})$ will be considered metric spaces with metric d_{H} defined by $d_{H}(\Phi, \Psi) = [E\int_{0}^{\infty} h^{2}(\Phi_{t}, \Psi_{t})\mathrm{d}t]^{1/2}$ for every $\Phi = (\Phi_{t})_{t\geq 0}, \Psi = (\Psi_{t})_{t\geq 0} \in \mathcal{L}^{2}_{\mathbb{F}}(\mathbb{R}^{+} \times \Omega, \mathbb{R}^{d})$ or $\Phi = (\Phi_{t})_{t\geq 0}, \Psi = (\Psi_{t})_{t\geq 0} \in \mathcal{L}^{2}_{\mathbb{F}}(\mathbb{R}^{+} \times \Omega, \mathbb{R}^{d\times m})$. It can be verified that $(\mathcal{L}^{2}_{\mathbb{F}}(\mathbb{R}^{+} \times \Omega, \mathbb{R}^{d}), d_{H})$ is a complete metric space. For fixed $T > 0$, we define $\mathcal{L}^{2}_{\mathbb{F}}(T, \Omega, \mathbb{R}^{d}) = \{(\mathbb{1}_{[0,T]}\Phi_{t})_{t\geq 0} : (\Phi_{t})_{t\geq 0} \in \mathcal{L}^{2}_{\mathbb{F}}(\mathbb{R}^{+} \times \Omega, \mathbb{R}^{d})$. The space $\mathcal{L}^{2}_{\mathbb{F}}(T, \Omega, \mathbb{R}^{d\times m})$ is defined similarly. We shall consider $\mathcal{L}^{2}_{\mathbb{F}}(T, \Omega, \mathbb{R}^{d})$ with the metric d_{H}, which in this case, is defined by $d_{H}(\Phi, \Psi) = [E\int_{0}^{T} h^{2}(\Phi_{t}, \Psi_{t})\mathrm{d}t]^{1/2}$ for $\Phi, \Psi \in \mathcal{L}^{2}_{\mathbb{F}}(T, \Omega, \mathbb{R}^{d})$. We shall also consider spaces $\mathcal{L}^{4}_{\mathbb{F}}(T, \Omega, \mathbb{R}^{d})$ and $\mathcal{L}^{4}_{\mathbb{F}}(T, \Omega, \mathbb{R}^{d\times m})$, defined in a similar way. In what follows, stochastic processes Φ and Ψ belonging to $\mathcal{L}^{2}_{\mathbb{F}}(T, \Omega, \mathbb{R}^{d})$ and $\mathcal{L}^{2}_{\mathbb{F}}(T, \Omega, \mathbb{R}^{d\times m})$ will be written as families $\Phi = (\Phi_{t})_{0\leq t\leq T}$ and $\Psi = (\Psi_{t})_{0\leq t\leq T}$, respectively. We shall also consider metric spaces $\mathrm{Cl}[\mathbb{L}^{2}([0,T] \times \Omega, \Sigma_{\mathbb{F}}, \mathbb{R}^{d})]$ and $\mathrm{Cl}[\mathbb{L}^{2}([0,T] \times \Omega, \Sigma_{\mathbb{F}}, \mathbb{R}^{d\times m})]$ with Hausdorff metrics denoted in both cases by D. Given a sequence $(F^{n})_{n=1}^{\infty}$ of set-valued stochastic processes, $F^{n} = (F^{n}_{t})_{0\leq t\leq T} \in \mathcal{L}^{2}_{\mathbb{F}}(T, \Omega, \mathbb{R}^{d})$ is said to be uniformly integrably bounded if there exists $m \in \mathbb{L}^{2}([0,T] \times \Omega, \Sigma_{\mathbb{F}}, \mathbb{R}^{+})$ such that $\|F^{n}_{t}(\omega)\| \leq m_{t}(\omega)$ for $n \geq 1$ and a.e. $(t, \omega) \in [0,T] \times \Omega$. It is said to be uniformly integrable if

$$\lim_{C\to\infty} \sup_{n\geq 1} \int\int_{\{(t,\omega):\|F^{n}_{t}(\omega)\|>C\}} \|F^{n}_{t}(\omega)\|^{2}\mathrm{d}t\mathrm{d}P = 0 .$$

It is clear that every uniformly integrably bounded sequence $(F^{n})_{n=1}^{\infty}$ of set-valued stochastic processes of $\mathcal{L}^{2}_{\mathbb{F}}(T, \Omega, \mathbb{R}^{d})$ is also uniformly integrable. It is easy to see that every sequence $(\varphi^{n})_{n=1}^{\infty}$ of $\mathbb{F}$-nonanticipative selectors φ^{n} of a uniformly integrable sequence $(F^{n})_{n=1}^{\infty} \subset \mathcal{L}^{2}_{\mathbb{F}}(T, \Omega, \mathbb{R}^{d})$ is uniformly integrable. Finally, let us observe that every sequence $(F^{n})_{n=1}^{\infty}$ of set-valued stochastic processes of $\mathcal{L}^{2}_{\mathbb{F}}(T, \Omega, \mathbb{R}^{d})$ converging in the d_{H}-metric topology to $F \in \mathcal{L}^{2}_{\mathbb{F}}(T, \Omega, \mathbb{R}^{d})$ is uniformly integrable.

Lemma 4.2. *Let J and $\mathcal{J}$ be linear continuous mappings defined on $\mathbb{L}^2([0,T]\times\Omega,\Sigma_{\mathbb{F}},\mathbb{R}^d)$ and $\mathbb{L}^2([0,T]\times\Omega,\Sigma_{\mathbb{F}},\mathbb{R}^{d\times m})$, respectively, with values at $\mathbb{L}^2(\Omega,\mathcal{F}_T,\mathbb{R}^d)$. If $(\Phi^n)_{n=1}^\infty$ and $(\Psi^n)_{n=1}^\infty$ are sequences of $\mathcal{L}^2_{\mathbb{F}}(T,\Omega,\mathbb{R}^d)$ and $\mathcal{L}^2_{\mathbb{F}}(T,\Omega,\mathbb{R}^{d\times m})$ converging in the d_H-metric topology to $\Phi\in\mathcal{L}^2_{\mathbb{F}}(T,\Omega,\mathbb{R}^d)$, and $\Psi\in\mathcal{L}^2_{\mathbb{F}}(T,\Omega,\mathbb{R}^{d\times m})$, respectively, then*

(i) $\lim_{n\to\infty}[\max\{D(S_{\mathbb{F}}(\Phi^n),S_{\mathbb{F}}(\Phi)),D(S_{\mathbb{F}}(\Psi^n),S_{\mathbb{F}}(\Psi))\}]=0$;
(ii) $\lim_{n\to\infty}[\max\{H(J(S_{\mathbb{F}}(\Phi^n)),J(S_{\mathbb{F}}(\Phi)),H(\mathcal{J}(S_{\mathbb{F}}(\Psi^n)),\mathcal{J}(S_{\mathbb{F}}(\Psi)))\}]=0.$

Proof. By Theorem 3.1, for every $\varphi\in S_{\mathbb{F}}(\Phi^n))$, one has $E[\int_0^T\inf\{\|\varphi_t(\omega)-x\|^2 : x\in\Phi(t,\omega)\}\mathrm{d}t]=\inf\{E\int_0^T\|\varphi_t-f_t\|^2\mathrm{d}t : f\in S_{\mathbb{F}}(\Phi)\}=\mathrm{Dist}^2(\varphi,S_{\mathbb{F}}(\Phi))$. Similarly, for every $f\in S_{\mathbb{F}}(\Phi)$, we get $\mathrm{Dist}^2(f,S_{\mathbb{F}}(\Phi^n))=E\int_0^T\inf\{\|f_t(\omega)-x\|^2 : x\in\Phi^n_t(\omega)\}\mathrm{d}t$. Hence it follows that $D(S_{\mathbb{F}}(\Phi^n),S_{\mathbb{F}}(\Phi))\le d_H(\Phi^n,\Phi)$ for every $n\ge 1$, which implies $D(S_{\mathbb{F}}(\Phi^n),S_{\mathbb{F}}(\Phi))\to 0$ as $n\to\infty$. In a similar way, we also get $D(S_{\mathbb{F}}(\Psi^n),S_{\mathbb{F}}(\Psi))\to 0$ as $n\to\infty$.

It is easy to see that (ii) follows immediately from (i) and the properties of the mappings J and $\mathcal{J}$. Indeed, let us observe first that by (i), continuity of J and boundedness of $S_{\mathbb{F}}(\Phi)$ and $S_{\mathbb{F}}(\Phi_n))$, there exists $M>0$ such that $(E|J(\varphi)-J(\psi)|^2)^{1/2}\le M(\int_0^T E|\varphi-\psi|^2\mathrm{d}t)^{1/2}$ for $n\ge 1$, $\varphi\in S_{\mathbb{F}}(\Phi)$ and $\psi\in S_{\mathbb{F}}(\Phi_n))$. Suppose now that (ii) is not satisfied and let $A=J[S_{\mathbb{F}}(\Phi)]$ and $A_n=J[S_{\mathbb{F}}(\Phi_n)]$ for $n\ge 1$. There exist $\bar\varepsilon>0$ and an increasing subsequence $(n_k)_{k=1}^\infty$ of $(n)_{n=1}^\infty$ such that $\overline{H}(A_{n_k},A)>\bar\varepsilon$ for every $k\ge 1$. Hence it follows that for every $k\ge 1$, there exists $g^k\in A_{n_k}$ such that $\bar\varepsilon/2<(E|g^k-f|^2)^{1/2}$ for every $f\in A$. Let $\varphi^k\in S_{\mathbb{F}}(\Phi_{n_k})$ and $\phi\in S_{\mathbb{F}}(\Phi)$ be such that $g^k=J(\varphi^k)$ for $k\ge 1$ and $f=J(\phi)$. For every $k\ge 1$, one has

$$\bar\varepsilon/2<(E|g^k-f|^2)^{1/2}\le M\left(\int_0^T E|\varphi^k_t-\phi_t|^2\mathrm{d}t\right)^{1/2}.$$

By (i), it follows that for every $\varphi^k\in S(\Phi_{n_k})$, with $k\ge 1$ sufficiently large, there exists $\xi^k\in S_{\mathbb{F}}(\Phi)$ such that $(E\int_0^T|\varphi^k_t-\xi^k_t|^2\mathrm{d}t)^{1/2}\le\bar\varepsilon/2M$. Taking in particular $\phi=\xi^k$ with sufficiently large $k\ge 1$, we obtain

$$\bar\varepsilon/2<(E|g^k-f|^2)^{1/2}\le M\left(\int_0^T E|\varphi^k_t-\xi^k_t|^2\mathrm{d}t\right)^{1/2} M\cdot\bar\varepsilon/2M=\bar\varepsilon/2,$$

a contradiction. Then $H[J(S_{\mathbb{F}}(\Phi^n),J(S_{\mathbb{F}}(\Phi))]\to 0$ as $n\to\infty$. In a similar way, we also get $H[\mathcal{J}(S_{\mathbb{F}}(\Psi^n),\mathcal{J}(S_{\mathbb{F}}(\Psi))]\to 0$ as $n\to\infty$. □

Remark 4.1. If $J(\varphi)=\int_0^T\varphi_t\mathrm{d}t$ and $\mathcal{J}(\psi)=\int_0^T\psi_t\mathrm{d}B_t$ for $\varphi\in\mathbb{L}^2([0,T]\times\Omega,\Sigma_{\mathbb{F}},\mathbb{R}^d)$ and $\psi\in\mathbb{L}^2([0,T]\times\Omega,\Sigma_{\mathbb{F}},\mathbb{R}^{d\times m})$, then $\overline{H}(J[S_{\mathbb{F}}(\Phi^n)],J[S_{\mathbb{F}}(\Phi)])\le\sqrt{T}d_H(\Phi^n,\Phi)$ and $\overline{H}(\mathcal{J}[S_{\mathbb{F}}(\Psi^n)],\mathcal{J}[S_{\mathbb{F}}(\Psi)])\le d_H(\Psi^n,\Psi)$ for every $n\ge 1$.

Proof. For every $u\in J[S_{\mathbb{F}}(\Phi^n)]$, one has $\mathrm{dist}^2(u,J[S_{\mathbb{F}}(\Phi)])\le E|u-v|^2$ for every $v\in J[S_{\mathbb{F}}(\Phi)]$. But $u=\int_0^T\varphi_t\mathrm{d}t$ and $v=\int_0^T\psi_t\mathrm{d}t$ for some $\varphi\in S_{\mathbb{F}}(\Phi^n)$

and $\psi \in S_{\mathbb{F}}(\Phi)$. Therefore, $\mathrm{dist}^2(u, J[S_{\mathbb{F}}(\Phi)]) \leq E|\int_0^T(\varphi - \psi)\mathrm{d}t|^2$ for every $\psi \in S_{\mathbb{F}}(\Phi)$. By Theorem 3.1, we have

$$\inf\left\{E\left|\int_0^T \varphi_t \mathrm{d}t - \int_0^T f t \mathrm{d}t\right|^2 : f \in S_{\mathbb{F}}(\Phi)\right\}$$
$$\leq T\inf\left\{E\int_0^T |\varphi_t - f_t|^2 \mathrm{d}t : f \in S_{\mathbb{F}}(\Phi)\right\}$$
$$= TE\int_0^T \mathrm{dist}^2(\varphi_t, \Phi_t)\mathrm{d}t \leq Td_H^2(\Phi^n, \Phi).$$

□

Thus $\mathrm{dist}^2(u, J[S_{\mathbb{F}}(\Phi)]) \leq Td_H^2(\Phi^n, \Phi)$ for every $u \in J[S_{\mathbb{F}}(\Phi^n)]$, which implies that $\overline{H}(J[S_{\mathbb{F}}(\Phi^n)], J[S_{\mathbb{F}}(\Phi)]) \leq \sqrt{T}d_H(\Phi^n, \Phi)$ for $n \geq 1$. In a similar way, we also get $\overline{H}(\mathcal{J}[S_{\mathbb{F}}(\Psi^n)], \mathcal{J}[S_{\mathbb{F}}(\Psi)]) \leq d_H(\Psi^n, \Psi)$ for every $n \geq 1$.

In what follows, we shall deal with a conditional expectation of set-valued integrals depending on a random parameter. We begin with the general definition of set-valued conditional expectation and its basic properties. Given a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a sub-$\sigma$-algebra $\mathcal{G}$ of $\mathcal{F}$, and a set-valued random variable $\Phi : \Omega \to \mathrm{Cl}(\mathbb{R}^d)$ the following result follows immediately from Theorem 3.2.

Lemma 4.3. *If $\Phi : \Omega \to \mathrm{Cl}(\mathbb{R}^d)$ is a set-valued random variable such that $S(\Phi) \neq \emptyset$, then there exists a unique in the a.s. sense $\mathcal{G}$-measurable set-valued random variable $\Psi : \Omega \to \mathrm{Cl}(\mathbb{R}^d)$ such that $S(\Psi) = \mathrm{cl}_{\mathbb{L}}\{E[\varphi|\mathcal{G}] : \varphi \in S(\Phi)\}$.*

Proof. Let $A \in \mathcal{G} \subset \mathcal{F}$ and $\mathcal{H} = \{E[\varphi|\mathcal{G}] : \varphi \in S(\Phi)\}$. For every $\psi_1, \psi_2 \in \mathcal{H}$, there exist $\varphi_1, \varphi_2 \in S(\Phi)$ such that $\psi_1 = E[\varphi_1|\mathcal{G}]$ and $\psi_2 = E[\varphi_2|\mathcal{G}]$. By the decomposability of $S(\Phi)$, it follows that $E[\mathbb{1}_A\varphi_1 + \mathbb{1}_{\Omega\setminus A}\varphi_2|\mathcal{G}] \in \mathcal{H}$. Then $\mathcal{H}$ is decomposable, because $E[\mathbb{1}_A\varphi_1 + \mathbb{1}_{\Omega\setminus A}\varphi_2|\mathcal{G}] = \mathbb{1}_A\psi_1 + \mathbb{1}_{\Omega\setminus A}\psi_2$. Therefore, $\mathrm{cl}_{\mathbb{L}}(\mathcal{H})$ is a decomposable subset of $\mathbb{L}^p(\Omega, \mathcal{G}, \mathbb{R}^d)$. By virtue of Theorem 3.2, there exists a $\mathcal{G}$-measurable set-valued mapping $\Psi : \Omega \to \mathrm{Cl}(\mathbb{R}^d)$ such that $S(\Psi) = \mathrm{cl}_{\mathbb{L}}(\mathcal{H})$. Suppose there are two $\mathcal{G}$-measurable mappings $\Psi_1, \Psi_2 : \Omega \to \mathrm{Cl}(\mathbb{R}^d)$ such that $S(\Psi_1) = S(\Psi_2) = \mathrm{cl}_{\mathbb{L}}(\mathcal{H})$. By Corollary 3.1, it follows that $\Psi_1 = \Psi_2$ a.s. □

A $\mathcal{G}$-measurable set-valued mapping $\Psi : \Omega \to \mathrm{Cl}(\mathbb{R}^d)$ such that $S(\Psi) = \mathrm{cl}_{\mathbb{L}}\{E[\varphi|\mathcal{G}] : \varphi \in S(\Phi)\}$ is denoted by $E[\Phi|\mathcal{G}]$ and is said to be a $\mathcal{G}$-conditional expectation of a set-valued mapping of $\Phi : \Omega \to \mathrm{Cl}(\mathbb{R}^d)$. Let us observe that for every square integrably bounded convex-valued set-valued random variable $\Phi : \Omega \to \mathrm{Cl}(\mathbb{R}^d)$, the set $S(\Phi)$ is a convex and weakly compact subset of $\mathbb{L}^2(\Omega, \mathbb{R}^d)$. Then $\{E[\varphi|\mathcal{G}] : \varphi \in S(\Phi)\}$ is a closed subset of this space. Indeed, for every $u \in \mathrm{cl}_{\mathbb{L}}\{E[\varphi|\mathcal{G}] : \varphi \in S(\Phi)\}$, there is a sequence $(\varphi_n)_{n=1}^{\infty} \subset S(\Phi)$ such that $E[\varphi_n|\mathcal{G}] \to u$ as $n \to \infty$. Let $(\varphi_{n_k})_{k=1}^{\infty}$ be a subsequence of $(\varphi_n)_{n=1}^{\infty}$ weakly converging to $\varphi \in S(\Phi)$. Therefore, for every $A \in \mathcal{G}$, one has $\int_A E[\varphi_{n_k}|\mathcal{G}]\mathrm{d}P = \int_A \varphi_{n_k}\mathrm{d}P \to \int_A \varphi\mathrm{d}P = \int_A E[\varphi|\mathcal{G}]\mathrm{d}P$ as $k \to \infty$. Then $E[\varphi_{n_k}|\mathcal{G}]$ converges weakly to $E[\varphi|\mathcal{G}]$ as $k \to \infty$, which implies that $u = E[\varphi|\mathcal{G}] \in \{E[\varphi|\mathcal{G}] : \varphi \in S(\Phi)\}$.

Corollary 4.1. *If $\Phi : \Omega \to \mathrm{Cl}(\mathbb{R}^d)$ is a square integrably bounded convex-valued set-valued random variable, then $S(E[\Phi|\mathcal{G}]) = \{E[\varphi|\mathcal{G}] : \varphi \in S(\Phi)\}$.* □

Theorem 4.1. *Let $\Phi : \Omega \to \mathrm{Cl}(\mathbb{R}^d)$ and $\Psi : \Omega \to \mathrm{Cl}(\mathbb{R}^d)$ be $\mathcal{F}$-measurable integrably bounded and let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$. Then*

(i) $E[\mathbb{1}_A E[\Phi|\mathcal{G}]] = E[\mathbb{1}_A \Phi]$ *for every* $A \in \mathcal{G}$.
(ii) $E[\xi\Phi|\mathcal{G}] = \xi E[\Phi|\mathcal{G}]$ *for every* $\xi \in \mathbb{L}^\infty(\Omega, \mathcal{G}, \mathbb{R})$.
(iii) $E[\overline{\mathrm{co}}\,\Phi|\mathcal{G}] = \overline{\mathrm{co}}\,E[\Phi|\mathcal{G}]$.
(iv) $H(E[\Phi|\mathcal{G}], E[\Psi|\mathcal{G}]) \le H(\Phi, \Psi)$, *where* $H(\Phi, \Psi) = E[h(\Phi, \Psi)]$.
(v) $E[\overline{\Phi + \Psi}|\mathcal{G}] = \overline{E[\Phi|\mathcal{G}] + E[\Psi|\mathcal{G}]}$ *a.s.*

Proof. (i) Let $A \in \mathcal{G}$ be fixed. If $u \in S(E[\Phi|\mathcal{G}])$, then there exists a sequence $(\varphi_n)_{n=1}^\infty$ in $S(\Phi)$ such that $\|u - E[\varphi_n|\mathcal{G}]\| \to 0$ as $n \to \infty$. Then $E[\mathbb{1}_A u] = \lim_{n\to\infty} E[\mathbb{1}_A E[\varphi_n|\mathcal{G}]] = \lim_{n\to\infty} E[\mathbb{1}_A \varphi_n]$. Hence by the compactness of the Aumann integral $E[\mathbb{1}_A \Phi]$, it follows that $E[\mathbb{1}_A u] \in E[\mathbb{1}_A \Phi]$. Thus $E[\mathbb{1}_A E[\Phi|\mathcal{G}]] \subset E[\mathbb{1}_A \Phi]$. Let $\mathcal{H} = \{E[\varphi|\mathcal{G}] : \varphi \in S(\phi)\}$. Then $E[\mathbb{1}_A \mathcal{H}] = \{E[\mathbb{1}_A E[\varphi|\mathcal{G}] : \varphi \in S(\phi)\} = E[\mathbb{1}_A \Phi]$. Hence it follows that $E[\mathbb{1}_A \Phi] \subset E[\mathbb{1}_A \mathrm{cl}_{\mathbb{L}}(\mathcal{H})] = E[\mathbb{1}_A E[\Phi|\mathcal{G}]]$. Therefore, $E[\mathbb{1}_A E[\Phi|\mathcal{G}]] = E[\mathbb{1}_A \Phi]$ for every $A \in \mathcal{G}$.

(ii) Let $\xi \in \mathbb{L}^\infty(\Omega, \mathcal{G}, \mathbb{R})$. We have to show that $S(E[\xi\Phi|\mathcal{G}]) = S(\xi E[\Phi|\mathcal{G}])$. By the definition of a set-valued conditional expectation, we have $S(E[\xi\Phi|\mathcal{G}]) = \mathrm{cl}_{\mathbb{L}}(\{E[f|\mathcal{G}] : f \in S(\xi\Phi)\})$ and $S(\xi E[\Phi|\mathcal{G}]) = \xi S(E[\Phi|\mathcal{G}]) = \xi \mathrm{cl}_{\mathbb{L}}(\{E[\varphi|\mathcal{G}] : \varphi \in S(\Phi)\})$. Let $u \in \xi\mathrm{cl}_{\mathbb{L}}(\{E[\varphi|\mathcal{G}] : \varphi \in S(\Phi)\})$ and $(\varphi_n)_{n=1}^\infty$ be a sequence of $S(\Phi)$ such that $\|\xi E[\varphi_n|\mathcal{G}] - u\| \to 0$ as $n \to \infty$. But $\xi E[\varphi_n|\mathcal{G}] = E[\xi\varphi_n|\mathcal{G}]$ for $n \ge 1$. Then $\|E[\xi\varphi|\mathcal{G}] - u\| \to 0$ as $n \to \infty$. We also have $\xi\varphi_n \in S(\xi\Phi)$ for $n \ge 1$. Therefore, $E[\xi\varphi_n|\mathcal{G}] \in \{E[f|\mathcal{G}] : f \in S(\xi\Phi)\}$ for $n \ge 1$, which implies that $u \in \mathrm{cl}_{\mathbb{L}}\{E[f|\mathcal{G}] : f \in S(\xi\Phi)\}$. Thus

$$\xi\mathrm{cl}_{\mathbb{L}}(\{E[\varphi|\mathcal{G}] : \varphi \in S(\Phi)\}) \subset \mathrm{cl}_{\mathbb{L}}(\{E[f|\mathcal{G}] : f \in S(\xi\Phi)\}).$$

Let $v \in \mathrm{cl}_{\mathbb{L}}\{E[f|\mathcal{G}] : f \in S(\xi\Phi)\}$ and $(\varphi_n)_{n=1}^\infty \subset S(\Phi)$ be such that $\|E[\xi\varphi_n|\mathcal{G}] - v\| \to 0$ as $n \to \infty$. Hence it follows that $\|\xi E[\varphi_n|\mathcal{G}] - v\| \to 0$ as $n \to \infty$. Similarly as above, we get $\xi E[\varphi_n|\mathcal{G}] \in \xi\{E[\varphi|\mathcal{G}] : \varphi \in S(\Phi)\} \subset \xi\mathrm{cl}_{\mathbb{L}}(\{E[\varphi|\mathcal{G}] : \varphi \in S(\Phi)\})$ for every $n \ge 1$. Therefore, $v \in \xi\mathrm{cl}_{\mathbb{L}}(\{E[\varphi|\mathcal{G}] : \varphi \in S(\Phi)\})$. Then $\mathrm{cl}_{\mathbb{L}}(\{E[f|\mathcal{G}] : f \in S(\xi\Phi)\}) \subset \mathrm{cl}_{\mathbb{L}}(\{E[\varphi|\mathcal{G}] : \varphi \in S(\Phi)\})$, which implies that $S(E[\xi\Phi|\mathcal{G}]) = S(\xi E[\Phi|\mathcal{G}])$.

(iii) Let $G = E[\Phi|\mathcal{G}]$. By Lemma 3.3, we obtain $S(E[\overline{\mathrm{co}}\,\Phi|\mathcal{G}]) = \mathrm{cl}_{\mathbb{L}}\{E[\varphi|\mathcal{G}] : \varphi \in \overline{\mathrm{co}}\,S(\Phi)\} = \overline{\mathrm{co}}\{E[\varphi|\mathcal{G}] : \varphi \in S(\Phi)\} = \overline{\mathrm{co}}\,S(G) = S(\overline{\mathrm{co}}\,G)$. Hence, by Corollary 3.1, it follows $E[\overline{\mathrm{co}}\,\Phi|\mathcal{G}] = \overline{\mathrm{co}}\,E[\Phi|\mathcal{G}]$.

(iv) Let $A = \{\omega \in \Omega : \sup[\mathrm{dist}(y, E[\Psi|\mathcal{G}](\omega)) : y \in E[\Phi|\mathcal{G}](\omega)] \geq \sup[\mathrm{dist}(y, E[\Phi|\mathcal{G}](\omega)) : y \in E[\Psi|\mathcal{G}](\omega)]\}$. We have $A \in \mathcal{G}$ and

$$\begin{aligned}
H(E[\Phi|\mathcal{G}], E[\Psi|\mathcal{G}]) &= E[h(E[\Phi|\mathcal{G}], E[\Psi|\mathcal{G}])] \\
&= E[\mathbb{1}_A \sup[\mathrm{dist}(y, E[\Psi|\mathcal{G}](\omega)) : y \in E[\Phi|\mathcal{G}](\omega)] \\
&\quad + E[\mathbb{1}_{\Omega\setminus A} \sup[\mathrm{dist}(E[y, E[\Phi|\mathcal{G}](\omega)) : y \in E[\Psi|\mathcal{G}](\omega)] \\
&= \sup E[\mathbb{1}_A \sup[\mathrm{dist}(E[\varphi|\mathcal{G}], E[\Psi|\mathcal{G}]) : \varphi \in S(\Phi)] \\
&\quad + E[\mathbb{1}_{\Omega\setminus A} \sup[\mathrm{dist}(E[\psi|\mathcal{G}], E[\Phi|\mathcal{G}]) : \psi \in S(\Psi)] \\
&\leq \sup_{\varphi \in S(\Phi)} \inf_{\psi \in S(\Psi)} E[\mathbb{1}_A E[|\varphi - \psi||\mathcal{G}]] \\
&\quad + \sup_{\psi \in S(\Psi)} \inf_{\varphi \in S(\Phi)} E[\mathbb{1}_{\Omega\setminus A} E[|\varphi - \psi||\mathcal{G}]] \\
&= \sup_{\varphi \in S(\Phi)} \inf_{\psi \in S(\Psi)} E[\mathbb{1}_A |\varphi - \psi|] \\
&\quad + \sup_{\psi \in S(\Psi)} \inf_{\varphi \in S(\Phi)} E[\mathbb{1}_{\Omega\setminus A} |\varphi - \psi|] \\
&= \int_A \sup[\mathrm{dist}(x, \Psi(\omega)) : x \in \Phi(\omega)] \mathrm{d}\mathbb{P} \\
&\quad + \int_{\Omega\setminus A} \sup[\mathrm{dist}(x, \Phi(\omega)) : x \in \Psi(\omega)] \mathrm{d}\mathbb{P} \\
&= \int_\Omega h(\Phi(\omega), \Psi(\omega)) \mathrm{d}\mathbb{P} = H(\Phi, \Psi).
\end{aligned}$$

(v) By the definition of a multivalued conditional expectation, we have $S(E[\overline{\Phi + \Psi}|\mathcal{G}]) = \mathrm{cl}_{\mathbb{L}}\{E[g|\mathcal{G}] : g \in S(\overline{\Phi + \Psi})$. By virtue of Lemma 3.4, we have

$$\begin{aligned}
S(E[\overline{\Phi + \Psi}|\mathcal{G}]) &= \mathrm{cl}_{\mathbb{L}}(\{E[\phi|\mathcal{G}] + E[\psi|\mathcal{G}] : \phi \in S(\Phi), \psi \in S(\Psi)\}) \\
&= \overline{S(E[\Phi|\mathcal{G}]) + S(E[\Psi|\mathcal{G}])} = S(\overline{E[\Phi|\mathcal{G}] + E[\Psi|\mathcal{G}]}),
\end{aligned}$$

which by Corollary 3.1, implies that $E[\overline{\Phi + \Psi}|\mathcal{G}] = \overline{E[\Phi|\mathcal{G}] + E[\Psi|\mathcal{G}]}$ a.s. □

Remark 4.2. It can be proved that if $\Phi \in \mathcal{A}(\Omega, \mathcal{F}, \mathbb{R}^d)$ is convex-valued and $\mathcal{T}$ is sub-σ-algebra of $\mathcal{G} \subset \mathcal{F}$, then $E[\Phi|\mathcal{T}]$ taken on the base space $(\Omega, \mathcal{F}, P)$ is equal to $E[\Phi|\mathcal{T}]$ taken on the base space $(\Omega, \mathcal{G}, P)$ and $E[E[\Phi|\mathcal{G}]|\mathcal{T}] = E[\Phi|\mathcal{T}]$, P-a.s. □

5 Notes and Remarks

The definitions and results of the first two sections of this chapter are mainly based on Aubin and Frankowska [12], Hu and Papageorgiou [41], Aubin and Cellina [5], Kisielewicz [49], Kuratowski [69], Hildenbrand [40] and Klein, and Thomson [63]. In particular, Michael's continuous selection theorem is taken from Aubin and Cellina [5] and Kisielewicz [49], whereas Theorem 2.2 comes from Kisielewicz [57]. The proofs of the Kuratowski and Ryll-Nardzewski measurable selection theorem and the Carathèodory selection theorem are taken from Hu and Papageorgiou [41]. The existence of measurable selectors for measurable multifunctions has been considered first by Kuratowski and Ryll-Nardzewski in [70]. The existence of Carathéodory selections has been considered by Rybiński in [91], Fryszkowski in [32], and Kucia and Nowak in [66]. The proof of Theorem 2.3, dealing with the existence of Lipschitz-type selectors, is taken from Hu and Papageorgiou [41]. The idea of this proof is due to Przesławski [90]. The proofs of Lemmas 1.1 and 1.2, Remark 1.1, and Corollary 1.2 can be found in Kuratowski [69] and Hildenbrand [40], respectively. Figures 2.1–2.4 are taken from Aubin and Cellina [5] and Kisielewicz [49]. The proof of Remark 2.9 can be found in Hu and Papageorgiou [41]. The definition and properties of Aumann integrals are taken from Hiai and Umegaki [39] and Kisielewicz [49]. The first results dealing with Aumann integrals are due to Aumann [14]. The existence of continuous selections of multifunctions with decomposable values was proved by Fryszkowski [32]. The sketch of the proof of this theorem given in Sect. 2 is taken from Hu and Papageorgiou [41]. The definition and properties of conditional expectation of set-valued mappings are taken from Hiai and Umegaki [39]. More information on the Hukuhara difference can be found in Hukuhara [42].

Chapter 3
Set-Valued Stochastic Integrals

This chapter is devoted to basic notions of the theory of set-valued stochastic integrals. In Sect. 1, we present properties of functional set-valued stochastic integrals defined, like Aumann integrals, as images of subtrajectory integrals of set-valued stochastic processes by some linear mappings with values in $\mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$. The set-valued stochastic integrals defined in Sect. 2 are understood as certain set-valued random variables. Throughout this chapter, we shall deal with set-valued stochastic processes belonging to the spaces $\mathcal{M}(T, \Omega, \mathbb{R}^d)$, $\mathcal{M}_{\mathbb{F}}(T, \Omega, \mathbb{R}^d)$, and $\mathcal{M}_{\mathbb{F}}(T, \Omega, \mathbb{R}^{d\times m})$ and their subspaces $\mathcal{L}(T, \Omega, \mathbb{R}^d)$, $\mathcal{L}^2_{\mathbb{F}}(T, \Omega, \mathbb{R}^d)$, and $\mathcal{L}^2_{\mathbb{F}}(T, \Omega, \mathbb{R}^{d\times m})$ defined in Chap. 2. All of them are defined on a given filtered probability space $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}_T, \mathbb{F}, P)$ with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ satisfying the usual conditions and such that there exists an m-dimensional $\mathbb{F}$-Brownian motion $B = (B_t)_{t\geq 0}$ defined on this space. A given $F \in \mathcal{M}(T, \Omega, \mathbb{R}^d)$ is said to be Aumann integrable if $S(F) \neq \emptyset$. Similarly, processes $\Phi \in \mathcal{M}_{\mathbb{F}}(T, \Omega, \mathbb{R}^d)$ and $\Psi \in \mathcal{M}_{\mathbb{F}}(T, \Omega, \mathbb{R}^{d\times m})$ are said to be Itô integrable if $S_{\mathbb{F}}(\Phi) \neq \emptyset$ and $S_{\mathbb{F}}(\Psi) \neq \emptyset$.

1 Functional Set-Valued Stochastic Integrals

Let $J : \mathbb{L}^2([0,T] \times \Omega, \beta_T \otimes \mathcal{F}_T, \mathbb{R}^d) \to \mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$ and $\mathcal{J} : \mathbb{L}^2([0,T] \times \Omega, \Sigma_{\mathbb{F}}, \mathbb{R}^{d\times m}) \to \mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$ be mappings defined by $J(\varphi)(\omega) = (\int_0^T \varphi_t \mathrm{d}t)(\omega)$ and $\mathcal{J}(\psi)(\omega) = (\int_0^T \psi_t \mathrm{d}B_t)(\omega)$ for a.e. $\omega \in \Omega$, $\varphi \in \mathbb{L}^2([0,T] \times \Omega, \beta_T \otimes \mathcal{F}_T, \mathbb{R}^d)$ and $\psi \in \mathbb{L}^2([0,T] \times \Omega, \Sigma_{\mathbb{F}}, \mathbb{R}^{d\times m})$, respectively, where $\Sigma_{\mathbb{F}}$ denotes the σ-algebra of all $\mathbb{F}$-nonanticipative subsets of $[0,T] \times \Omega$. The following lemma follows immediately from the properties of the Lebesgue and Itô integrals.

Lemma 1.1. *The mappings J and $\mathcal{J}$ are linear and continuous with respect to the norm topologies of* $\mathbb{L}^2([0,T] \times \Omega, \beta_T \otimes \mathcal{F}_T, \mathbb{R}^r)$, $\mathbb{L}^2([0,T] \times \Omega, \Sigma_{\mathbb{F}}, \mathbb{R}^d)$, $\mathbb{L}^2([0,T] \times \Omega, \Sigma_{\mathbb{F}}, \mathbb{R}^{d\times m})$, *and* $\mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$, *respectively.*

M. Kisielewicz, *Stochastic Differential Inclusions and Applications*,
Springer Optimization and Its Applications 80, DOI 10.1007/978-1-4614-6756-4_3,

Proof. The linearity of J and $\mathcal{J}$ follows immediately from the properties of the Lebesgue and Itô integrals, respectively. For every $\varphi \in \mathbb{L}^2([0,T]\times\Omega, \beta_T\otimes\mathcal{F}_T, \mathbb{R}^d)$ and $\psi \in \mathbb{L}^2([0,T]\times\Omega, \Sigma_\mathbb{F}, \mathbb{R}^{d\times m})$, one has $E|J(\varphi)|^2 \le TE\int_0^T|\varphi_t|^2\mathrm{d}t$ and $E|\mathcal{J}(\psi)|^2 = E\int_0^T|\psi_t|^2\mathrm{d}t$. Hence it follows that J and $\mathcal{J}$ are continuous. □

For given Aumann and Itô integrable set-valued stochastic processes $F \in \mathcal{M}(T,\Omega,\mathbb{R}^d)$, $\Phi \in \mathcal{M}_\mathbb{F}(T,\Omega,\mathbb{R}^d)$, and $\Psi \in \mathcal{M}_\mathbb{F}(T,\Omega,\mathbb{R}^{d\times m})$, their functional set-valued integrals are defined as images of subtrajectory integrals $S(F)$, $S_\mathbb{F}(\Phi)$, and $S_\mathbb{F}(\Psi)$ by linear mappings J and $\mathcal{J}$, respectively, i.e., as subsets $J[S(F)]$, $J[S_\mathbb{F}(\Phi)]$, and $\mathcal{J}[S_\mathbb{F}(\Psi)]$ of the space $\mathbb{L}^2(\Omega,\mathcal{F}_T,\mathbb{R}^d)$. It is clear that $J[S_\mathbb{F}(\Phi) \subset J[S(\Phi)]$ for every Itô integrable process $\Phi \in \mathcal{M}(T,\Omega,\mathbb{R}^d)$.

Corollary 1.1. *For every $F \in \mathcal{L}(T,\Omega,\mathbb{R}^d)$, $\Phi \in \mathcal{L}^2_\mathbb{F}(T,\Omega,\mathbb{R}^d)$, and $\Psi \in \mathcal{L}^2_\mathbb{F}(T,\Omega,\mathbb{R}^{d\times m})$, the functional set-valued stochastic integrals $J[S(\mathrm{co}\,F)]$, $J[S_\mathbb{F}(\mathrm{co}\,\Phi)]$, and $\mathcal{J}[S_\mathbb{F}(\mathrm{co}\,\Psi)]$ are nonempty convex sequentially weakly compact and weakly compact subsets of $\mathbb{L}(\Omega,\mathcal{F}_T,\mathbb{R}^d)$ and $\mathbb{L}^2(\Omega,\mathcal{F}_T,\mathbb{R}^d)$, respectively.*

Proof. By virtue of Remark 3.1 of Chap. 2, the subtrajectory integrals $S(\mathrm{co}\,\Phi)$, $S_\mathbb{F}(\mathrm{co}\,\Phi)$, and $S_\mathbb{F}(\mathrm{co}\,\Psi)$ are nonempty convex sequentially weakly compact and weakly compact subsets of $\mathbb{L}([0,T]\times\Omega, \beta_T\otimes\mathcal{F}_T, \mathbb{R}^d)$, $\mathbb{L}^2([0,T]\times\Omega,\Sigma_\mathbb{F},\mathbb{R}^d)$, and $\mathbb{L}^2([0,T]\times\Omega,\Sigma_\mathbb{F},\mathbb{R}^{d\times m})$, respectively. Then by virtue of Lemma 1.1, the sets $J[S(\mathrm{co}\,\Phi)]$, $J[S_\mathbb{F}(\mathrm{co}\,\Phi)]$, and $\mathcal{J}[S_\mathbb{F}(\mathrm{co}\,\Psi)]$ are nonempty convex sequentially weakly compact and weakly compact subsets of $\mathbb{L}(\Omega,\mathcal{F}_T,\mathbb{R}^d)$ and $\mathbb{L}^2(\Omega,\mathcal{F}_T,\mathbb{R}^d)$, respectively. □

Lemma 1.2. *For every Itô integrable set-valued process $\Psi \in \mathcal{M}_\mathbb{F}(T,\Omega,\mathbb{R}^{d\times m})$, a set-valued stochastic integral $\mathcal{J}[S_\mathbb{F}(\Psi)]$ is a closed subset of $\mathbb{L}^2(\Omega,\mathcal{F}_T,\mathbb{R}^d)$.*

Proof. Let $(u_n)_{n=1}^\infty$ be a sequence of $\mathcal{J}[S_\mathbb{F}(\Psi)]$ converging to u in the norm topology of $\mathbb{L}^2(\Omega,\mathcal{F}_T,\mathbb{R}^d)$, and let $(\psi^n)_{n=1}^\infty$ be a sequence of $S_\mathbb{F}(\Psi)$ such that $u_n = \int_0^T \psi_t^n\mathrm{d}B_t$ for $n \ge 1$. For every $n, m \ge 1$, we have

$$|u_n - u_m\|^2 = E\left|\int_0^T(\psi_t^n - \psi_t^m)\mathrm{d}B_t\right|^2 = \int_0^T E|\psi_t^n - \psi_t^m|^2\mathrm{d}t.$$

Therefore, $(\psi^n)_{n=1}^\infty$ is a Cauchy sequence of $\mathbb{L}^2([0,T]\times\Omega,\Sigma_\mathbb{F},\mathbb{R}^{d\times m})$. Thus there exists $\psi \in \mathbb{L}^2([0,T]\times\Omega,\Sigma_\mathbb{F},\mathbb{R}^{d\times m})$ such that $E\int_0^T|\psi_t^n - \psi_t|^2\mathrm{d}t \to 0$ as $n\to\infty$. By the closedness of $S_\mathbb{F}(\Psi)$, we have $\psi \in S_\mathbb{F}(\Psi)$. Let $v = \int_0^T\psi_t\mathrm{d}B_t$. We have $v \in \mathcal{J}[S_\mathbb{F}(\Psi)]$ and

$$\begin{aligned} E|u-v|^2 &\le 2E|u-u_n|^2 + 2E\left|\int_0^T(\psi_t^n - \psi)\mathrm{d}B_t\right|^2 \\ &= 2E|u-u_n|^2 + E\int_0^T|\psi_t^n - \psi_t|^2\,\mathrm{d}t. \end{aligned}$$

Therefore, $v = u$ and $u \in \mathcal{J}[S_\mathbb{F}(\Psi)]$. □

Given the above set-valued processes F, Φ, and Ψ, and $0 \le s < t \le T$, the sets $J[\mathbb{1}_{[s,t]}S(F)]$, $J[\mathbb{1}_{[s,t]}S_{\mathbb{F}}(\Phi)]$, and $\mathcal{J}[\mathbb{1}_{[s,t]}S_{\mathbb{F}}(\Psi)]$ are denoted by $J_{st}[S(F)]$, $J_{st}[S_{\mathbb{F}}(\Phi)]$, and $\mathcal{J}_{st}[S_{\mathbb{F}}(\Phi)]$, respectively, and are said to be the functional set-valued stochastic integrals of F, Φ, and Ψ, respectively, on the interval $[s,t]$.

Lemma 1.3. *For every Aumann integrable set-valued process $F \in \mathcal{M}(T,\Omega,\mathbb{R}^d)$, there exists an $\mathcal{F}_T$-measurable set-valued mapping $\Lambda : \Omega \to \mathrm{Cl}(\mathbb{R}^d)$ such that $\mathrm{cl}_{\mathbb{L}}\{J[S(F)]\} = S_T(\Lambda)$, where $S_T(\Lambda)$ denotes the set of all $\mathcal{F}_T$-measurable selectors of Λ.*

Proof. We shall show that $J[S(F)]$ is a decomposable subset of $\mathbb{L}(\Omega,\mathcal{F}_T,\mathbb{R}^d)$. Indeed, for every $u, v \in J[S(F)]$, there exist $f \in S(F)$ and $g \in S(F)$ such that $u = \int_0^T f_t \mathrm{d}t$ and $v = \int_0^T g_t \mathrm{d}t$. For every $A \in \mathcal{F}_T$, one has $\mathbb{1}_A u + \mathbb{1}_{A^\sim} v = \int_0^T [\mathbb{1}_A f_t + \mathbb{1}_{A^\sim} g_t]\mathrm{d}t$, where $A^\sim = \Omega \setminus A$. But $\mathbb{1}_A = \mathbb{1}_{[0,T]} \cdot \mathbb{1}_A = \mathbb{1}_{[0,T]\times A}$ and $\mathbb{1}_{A^\sim} = \mathbb{1}_{[0,T]} \cdot \mathbb{1}_{A^\sim} = \mathbb{1}_{[0,T]\times A^\sim} = \mathbb{1}_{([0,T]\times A)^\sim}$, because $([0,T] \times A)^\sim = ([0,T]\times\Omega)\setminus([0,T]\times A) = [0,T]\times(\Omega\setminus A) = [0,T]\times A^\sim$. By the decomposability of $S(F)$, we get $\mathbb{1}_{[0,T]\times A} f + \mathbb{1}_{([0,T]\times A)^\sim} g \in S(F)$, which implies that $\mathbb{1}_A u + \mathbb{1}_{A^\sim} v = \int_0^T [\mathbb{1}_{[0,T]\times A} f_t + \mathbb{1}_{([0,T]\times A)^\sim} g_t]\mathrm{d}t \in J[S(F)]$. Then $\mathrm{cl}_{\mathbb{L}}\{J[(S(F))]\}$ is a closed decomposable subset of $\mathbb{L}(\Omega,\mathcal{F}_T,\mathbb{R}^d)$, which by virtue of Theorem 3.2 of Chap. 2, implies that there exists an $\mathcal{F}_T$-measurable set-valued mapping $\Lambda : \Omega \to \mathrm{Cl}(\mathbb{R}^d)$ such that $\mathrm{cl}_{\mathbb{L}}\{J[S(\Phi)]\} = S_T(\Lambda)$. □

Remark 1.1. If $F \in \mathcal{L}(T,\Omega,\mathbb{R}^d)$, then the multifunction Λ defined above is integrably bounded.

Proof. By the integrably boundedness of F, it follows that $J[S(F)]$ is a bounded subset of $\mathbb{L}(\Omega,\mathcal{F}_T,\mathbb{R}^d)$. Hence, by Corollary 3.3 of Chap. 2, it follows that Λ is integrably bounded. □

Remark 1.2. In the general case, the above procedure cannot be applied to the integrals $J[S_{\mathbb{F}}(\Phi)]$ and $\mathcal{J}[S_{\mathbb{F}}(\Psi)]$, because they are not decomposable subsets of $\mathbb{L}^2(\Omega,\mathcal{F}_T,\mathbb{R}^d)$. This can be seen from the following examples. □

Example 1.1. Let $\Phi(t,\omega) =: [0,1]$ for $t \in [0,1]$ and $\omega \in \Omega$ and suppose $J[S_{\mathbb{F}}(\Phi)]$ is decomposable. Then for every $A \in \mathcal{F}_1$ and $\varphi_1, \varphi_2 \in S_{\mathbb{F}}(\Phi)$, one has $\mathbb{1}_A J(\varphi_1) + \mathbb{1}_{A^\sim} J(\varphi_2) \in J[S_{\mathbb{F}}(\Phi)]$, where $A^\sim = \Omega \setminus A$. Taking, in particular, $\varphi_1 = 1, \varphi_2 = 0$, we get $\mathbb{1}_A J(\varphi_1) \in J[S_{\mathbb{F}}(\Phi)]$. Let $\psi = \mathbb{1}_{A\times[0,1]}$ for $A \in \mathcal{F}_1 \setminus \mathcal{F}_t$ with fixed $t \in [0,1)$. We have $J(\psi) = \mathbb{1}_A J(\varphi_1)$, which implies $J(\psi) \in J[S_{\mathbb{F}}(\Phi)]$. But ψ is not $\mathbb{F}$-nonanticipative, because a random variable $\psi(t,\cdot)$ is not $\mathcal{F}_t$-measurable. Therefore, $\psi \notin S_{\mathbb{F}}(\Phi)$ and $J(\psi) \in J[S_{\mathbb{F}}(\Phi)]$, a contradiction. Thus $J[S_{\mathbb{F}}(\Phi)]$ is not decomposable.

Example 1.2. Let $\Psi(t,\omega) =: [0,1]$ for $t \ge 0$ and $\omega \in \Omega$ and suppose $\mathcal{J}[S_{\mathbb{F}}(\Psi)]$ is decomposable. Then for every $A \in \mathcal{F}_1$ and every $u_1, u_2 \in \mathcal{J}[S_{\mathbb{F}}(\Psi)]$, one has $\mathbb{1}_A u_1 + \mathbb{1}_{A^\sim} u_2 \in \mathcal{J}[S_{\mathbb{F}}(\Psi)]$, where $A^\sim = \Omega \setminus A$. Suppose that $u_1 = \int_0^1 \mathrm{d}B_t = B_1$ and $u_2 = \int_0^1 0\mathrm{d}B_t = 0$ a.s. and let $A = \{\omega \in \Omega : B_1 \ge \varepsilon\}$ for $\varepsilon > 0$. We have $A \in \mathcal{F}_1$. On the other hand, by the definition of $\mathcal{J}[S_{\mathbb{F}}(\Psi)]$, there exists $\psi \in S_{\mathbb{F}}(\Psi)$

such that $\mathbb{1}_A u_1 + \mathbb{1}_{A\sim} u_2 = \mathbb{1}_A \cdot B_1 = \int_0^1 \psi_t \mathrm{d}B_t$. Then $E(\mathbb{1}_A \cdot B_1) = 0$. But $E(\mathbb{1}_A \cdot B_1) = \int_{\{B_1 \geq \varepsilon\}} B_1 \mathrm{d}P \geq \varepsilon P(\{B_1 \geq \varepsilon\}) > 0$, which contradicts the last equality.

Immediately from Theorem 3.2 of Chap. 2, we also obtain the following result.

Lemma 1.4. *For arbitrary Itô integrable set-valued processes* $\Phi \in \mathcal{M}_{\mathbb{F}}(T, \Omega, \mathbb{R}^d)$ *and* $\Psi \in \mathcal{M}_{\mathbb{F}}(T, \Omega, \mathbb{R}^{d\times m})$*, there exist* $\mathcal{F}_T$*-measurable set-valued random variables* $\Gamma, \mathcal{Z} : \Omega \to \mathrm{Cl}(\mathbb{R}^d)$ *such that* $\overline{\mathrm{dec}}\{J[S_{\mathbb{F}}(\Phi)]\} = S_T(\Gamma)$ *and* $\overline{\mathrm{dec}}\{\mathcal{J}[S_{\mathbb{F}}(\Psi)]\} = S_T(\mathcal{Z})$.

Lemma 1.5. *For every* $\Phi \in \mathcal{L}^2_{\mathbb{F}}(T, \Omega, \mathbb{R}^d)$ *and* $\Psi \in \mathcal{L}^2_{\mathbb{F}}(T, \Omega, \mathbb{R}^{d\times m})$*, the set* $J[S_{\mathbb{F}}(\overline{\mathrm{co}}\,\Phi)] + \mathcal{J}[S_{\mathbb{F}}(\overline{\mathrm{co}}\,\Psi)]$ *is a closed subset of* $\mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$.

Proof. Let $(u_n)_{n=1}^{\infty}$ be a sequence of $J[S_{\mathbb{F}}(\overline{\mathrm{co}}\,\Phi)] + \mathcal{J}[S_{\mathbb{F}}(\overline{\mathrm{co}}\,\Psi)]$ converging in the norm topology of $\mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$ to $u \in \mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$. Let $\phi^n \in S_{\mathbb{F}}(\overline{\mathrm{co}}\,\Phi)$ and $\psi^n \in S_{\mathbb{F}}(\overline{\mathrm{co}}\,\Psi)$ be such that $u_n = \int_0^T \phi_t^n \mathrm{d}t + \int_0^T \psi_t^n \mathrm{d}B_t$ for every $n \geq 1$. By the weak compactness of $S_{\mathbb{F}}(\overline{\mathrm{co}}\,\Phi)$, there exist $\phi \in S_{\mathbb{F}}(\overline{\mathrm{co}}\,\Phi)$ and a subsequence, denoted again by $(\phi^n)_{n=1}^{\infty}$, of $(\phi^n)_{n=1}^{\infty}$ weakly converging to ϕ. Then the sequence $(u_n - \int_0^T \phi_t^n \mathrm{d}t)_{n=1}^{\infty}$ converges weakly to $u - \int_T \phi_t$ as $n \to \infty$. Hence it follows that the sequence $(\int_0^T \psi_t^n \mathrm{d}B_t)_{n=1}^{\infty}$ weakly converges to $u - \int_0^T \phi_t \mathrm{d}t \in \mathcal{J}[S_{\mathbb{F}}(\overline{\mathrm{co}}\,\Psi)]$. Then there exists $\psi \in S_{\mathbb{F}}(\overline{\mathrm{co}}\,\Psi)$ such that $u - \int_0^T \phi_t \mathrm{d}t = \int_0^T \psi_t \mathrm{d}B_t$, which implies that $u \in J[S_{\mathbb{F}}(\overline{\mathrm{co}}\,\Phi)] + \mathcal{J}[S_{\mathbb{F}}(\overline{\mathrm{co}}\,\Psi)]$. □

Lemma 1.6. *For every* $\Phi \in \mathcal{L}^2_{\mathbb{F}}(T, \Omega, \mathbb{R}^d)$ *and* $\Psi \in \mathcal{L}^2_{\mathbb{F}}(T, \Omega, \mathbb{R}^{d\times m})$*, one has* $J[S_{\mathbb{F}}(\Phi)] + \mathcal{J}[S_{\mathbb{F}}(\Psi)] \subset \mathrm{cl}_{\mathbb{L}}\{J[S_{\mathbb{F}}(\Phi)]\} + \mathcal{J}[S_{\mathbb{F}}(\Psi)] \subset \mathrm{cl}_{\mathbb{L}}\{J[S_{\mathbb{F}}(\Phi)] + \mathcal{J}[S_{\mathbb{F}}(\Psi)]\} \subset J[S_{\mathbb{F}}(\overline{\mathrm{co}}\,\Phi)] + \mathcal{J}[S_{\mathbb{F}}(\overline{\mathrm{co}}\,\Psi)]$*, where the closures are taken in the norm topology of* $\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{R}^d)$.

Proof. It is enough to verify that

$$\mathrm{cl}_{\mathbb{L}}\{J[S_{\mathbb{F}}(\Phi)]\} + \mathcal{J}[S_{\mathbb{F}}(\Psi)] \subset \mathrm{cl}_{\mathbb{L}}\{J[S_{\mathbb{F}}(\Phi)] + \mathcal{J}[S_{\mathbb{F}}(\Psi)]\}.$$

The rest of the above inclusions follow immediately from the properties of the space $\mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$, the definitions and properties of functional set-valued stochastic integrals, and Lemma 1.5. Let $u \in \mathrm{cl}_{\mathbb{L}}\{J[S_{\mathbb{F}}(\Phi)]\} + \mathcal{J}[S_{\mathbb{F}}(\Psi)]$. There exists a sequence $(\phi^n)_{n=1}^{\infty}$ of $S_{\mathbb{F}}(\Phi)$ and $\psi \in S_{\mathbb{F}}(\Psi)$ such that $u = \lim_{n\to\infty}\left(\int_0^T \phi_t^n \mathrm{d}t\right) + \int_0^T \psi_t \mathrm{d}B_t$, where the limit is taken in the norm topology of $\mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$. Taking $\psi^n = \psi$ for every $n \geq 1$, we obtain $u = \lim_{n\to\infty}\left(\int_0^T \phi_t^n \mathrm{d}t + \int_0^T \psi_t^n \mathrm{d}B_t\right)$. Therefore, $u \in \mathrm{cl}_{\mathbb{L}}\{J[S_{\mathbb{F}}(\Phi)] + \mathcal{J}[S_{\mathbb{F}}(\Psi)]\}$. □

Lemma 1.7. *Let* $\mathcal{P}_{\mathbb{F}}$ *be a filtered separable probability space with a filtration* $\mathbb{F} = (\mathcal{F}_t)_{t\in T}$ *satisfying the usual conditions. For every* $F \in \mathcal{L}(T, \Omega, \mathbb{R}^d)$ *and* $\Phi \in \mathcal{L}^2_{\mathbb{F}}(T, \Omega, \mathbb{R}^d)$*, one has* $J[S(\overline{\mathrm{co}}\,F)] = \mathrm{cl}_{\mathbb{L}}\{J[S(F)]\}$ *and* $J[S_{\mathbb{F}}(\overline{\mathrm{co}}\,\Phi)] = \mathrm{cl}_{\mathbb{L}}\{J[S_{\mathbb{F}}(\Phi)]\}$.

Proof. We shall prove the second equality. The first one can be verified similarly. Put $H = \mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$ and let us observe that $S_{\mathbb{F}}(\Phi)$ is a subset of the space $\mathbb{L}^2([0,T], H)$ of all Bochner measurable functions $\phi : T \to H$ such that $\int_0^T \|\phi_t\| \mathrm{d}t < \infty$, where $\|\phi_t\|^2 = E|\phi_t|^2$. Indeed, every $\phi \in S_{\mathbb{F}}(\Phi)$ is $\beta_T \otimes \mathcal{F}_T$-measurable, $\mathbb{F}$-adapted, and such that $E\int_0^T |\phi_t|^2 \mathrm{d}t < \infty$. By virtue of Lemma 6.3 of Chap. 1, there exists a sequence $(\phi^n)_{n=1}^\infty$ of simple $\mathbb{F}$-nonanticipative processes on $\mathcal{P}_{\mathbb{F}}$ such that $E\int_0^T |\phi_t^n - \phi_t|^2 \mathrm{d}t \to 0$ as $n \to \infty$. Let $\varphi_t^n = \sum_{i=1}^{N-1} \mathbb{1}_{[t_{i-1}, t_i)}(t)\phi_{t_{i-1}}^n + \mathbb{I}_{[t_{N-1}, t_N]}(t)\phi_{t_{N-1}}$ for $n \geq 1$. It is clear that $\varphi^n : [0,T] \ni t \to \varphi_t^n \in H$ is for every $n \geq 1$ a step function such that $\int_0^T E|\varphi_t^n - \phi_t|^2 \mathrm{d}t \to 0$ as $n \to \infty$. Hence it follows that there exists a subsequence $(\varphi^{n_k})_{k=1}^\infty$ of $(\varphi^n)_{n=1}^\infty$ such that $E|\varphi_t^{n_k} - \phi_t|^2 \to 0$ for a.e. $t \in [0,T]$ as $k \to \infty$. Then the vector function $[0,T] \ni t \to \phi_t \in H$ is Bochner measurable such that the sequence $(B)\int_0^T \varphi_t^{n_k} \mathrm{d}t)_{k=1}^\infty$ of Bochner integrals of simple functions $\varphi^{n_k} : [0,T] \to H$ converges in the norm topology of H to $\int_0^T \phi_t \mathrm{d}t$, i.e., $E|\int_0^T \phi_t \mathrm{d}t - (B)\int_0^T \varphi_t^{n_k} \mathrm{d}t|^2 \to 0$ as $k \to \infty$. Thus the integral $\int_0^T \phi_t \mathrm{d}t$ can be regarded as the Bochner integral $(B)\int_0^T \phi_t \mathrm{d}t$ of the vector function $[0,T] \ni t \to \phi_t \in H$.

Let us observe that by the closedness of $S_{\mathbb{F}}(\Phi)$ in $\mathbb{L}^2([0,T] \times \Omega, \Sigma_{\mathbb{F}}, \mathbb{R}^d)$, for every sequence $(\phi^n)_{n=1}^\infty$ of $S_{\mathbb{F}}(\Phi)$ converging in the norm topology of $\mathbb{L}^2([0,T] \times \Omega, \beta_T \otimes \mathcal{F}_T, \mathbb{R}^d)$ to $\phi \in \mathbb{L}^2([0,T] \times \Omega, \beta_T \otimes \mathcal{F}_T, \mathbb{R}^d)$, we have $\phi \in S_{\mathbb{F}}(\Phi)$. But $\int_0^T \|\phi_t^n - \phi_t\|^2 \mathrm{d}t \to 0$ as $n \to \infty$. Therefore, $S_{\mathbb{F}}(\Phi)$ is a closed subset of $\mathbb{L}^2([0,T], H)$. It is also easy to see that $S_{\mathbb{F}}(\Phi)$ is a decomposable subset of $\mathbb{L}^2([0,T], H)$. Indeed, let $f, g \in S_{\mathbb{F}}(\Phi)$ and let A be a measurable subset of $[0,T]$. We have $(\mathbb{1}_A f + \mathbb{1}_{A^\sim} g)_t(\omega) = (\mathbb{1}_A(t) f_t + \mathbb{1}_{A^\sim}(t) g_t)(\omega) = \mathbb{1}_{A\times\Omega}(t,\omega) f_t(\omega) + \mathbb{1}_{(A\times\Omega)^\sim}(t,\omega) g_t(\omega) \in \Phi_t(\omega))$ for a.e. $(t,\omega) \in [0,T]\times\Omega$, and $\mathbb{1}_{A\times\Omega} f + \mathbb{1}_{(A\times\Omega)^\sim})g$ is $\mathbb{F}$-nonanticipative for every measurable set $A \subset [0,T]$. Then $\mathbb{1}_A f + \mathbb{1}_{A^\sim} g \in S_{\mathbb{F}}(\Phi)$ for every $f, g \in S_{\mathbb{F}}(\Phi)$ and every measurable set $A \subset [0,T]$, where $A^\sim = [0,T] \setminus A$. Therefore, by Remark 3.2 of Chap. 2, there exists a set-valued mapping $\mathcal{Z} : T \to Cl(H)$ such that $S(\mathcal{Z}) = S_{\mathbb{F}}(\Phi)$, where $S(\mathcal{Z})$ is the set of all $(\beta_T, \beta(H))$-measurable selectors for $\mathcal{Z}$. Hence it follows that the Aumann integral for $\mathcal{Z}$ can be defined by $\int_0^T \mathcal{Z}(t)\mathrm{d}t = \left\{(B)\int_0^T f_t \mathrm{d}t : f \in S_{\mathbb{F}}(\Phi)\right\}$, which by the equality $J(f) = (B)\int_0^T f_t \mathrm{d}t$, implies that $\int_0^T \mathcal{Z}(t)\mathrm{d}t = J[S_{\mathbb{F}}(\Phi)]$. By the separability of the space $\mathcal{P}_{\mathbb{F}}$, the Banach space $H = \mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$ is separable. Then by virtue of Remark 3.8 of Chap. 2, we have $\mathrm{cl}_{\mathbb{L}}(\int_0^T \overline{\mathrm{co}}\,\mathcal{Z}(t)\mathrm{d}t) = \mathrm{cl}_{\mathbb{L}}\left(\int_0^T \mathcal{Z}(t)\mathrm{d}t\right)$. By Lemma 3.3 of Chap. 2 and Corollary 1.1, one has

$$\mathrm{cl}_{\mathbb{L}}\left(\int_0^T \overline{co}\,\mathcal{Z}(t)\mathrm{d}t\right) = \mathrm{cl}_{\mathbb{L}}\left\{(B)\int_0^T f_t \mathrm{d}t : f \in S(\overline{co}\,\mathcal{Z})\right\}$$
$$= \mathrm{cl}_{\mathbb{L}}\left\{(B)\int_0^T f_t \mathrm{d}t : f \in \overline{\mathrm{co}}\,S(\mathcal{Z})\right\} = \mathrm{cl}_{\mathbb{L}}\{J(f) : f \in S_{\mathbb{F}}(\overline{\mathrm{co}}\,\Phi)\} = J[S_{\mathbb{F}}(\overline{\mathrm{co}}\,\Phi)].$$

Furthermore, we have $\mathrm{cl}_{\mathbb{L}}(\int_0^T \mathcal{Z}(t)\mathrm{d}t) = \mathrm{cl}_{\mathbb{L}}\{J[S_{\mathbb{F}}(\Phi)]\}$. Therefore, $J[S_{\mathbb{F}}(\overline{\mathrm{co}}\,\Phi)] = \mathrm{cl}_{\mathbb{L}}(J[S_{\mathbb{F}}(\Phi)])$. □

Corollary 1.2. *If $\Phi = (\Phi_t)_{t\in T}$ and $\Psi = (\Psi_t)_{t\in T}$ are as in Lemma 1.5, and $\mathcal{P}_{\mathbb{F}}$ is a separable filtered probability space, then $J[S_{\mathbb{F}}(\overline{\mathrm{co}}\,\Phi)] + \mathcal{J}[S_{\mathbb{F}}(\Psi)] = \mathrm{cl}_{\mathbb{L}}\{J[S_{\mathbb{F}}(\Phi)]\} + \mathcal{J}[S_{\mathbb{F}}(\Psi)] \subset \mathrm{cl}_{\mathbb{L}}\{J[S_{\mathbb{F}}(\Phi)] + \mathcal{J}[S_{\mathbb{F}}(\Psi)]\} \subset J[S_{\mathbb{F}}(\overline{\mathrm{co}}\,\Phi)] + \mathcal{J}[S_{\mathbb{F}}(\overline{\mathrm{co}}\,\Psi)]$.* □

Theorem 1.1. *Let $\Phi = (\Phi_t)_{0\le t\le T}$ and $\Psi = (\Psi_t)_{0\le t\le T}$ be $d \times m$-dimensional Itô integrable set-valued processes. Then*

(i) $E\{\mathcal{J}[S_{\mathbb{F}}(\Phi)]\} = \{0\}$. If Φ is square integrably bounded, then $\mathcal{J}[S_{\mathbb{F}}(\Phi)]$ is a bounded subset of $\mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$.
(ii) $\mathcal{J}[S_{\mathbb{F}}(\Phi)]$ is decomposable if and only if it is a singleton.
(iii) $\overline{\mathrm{dec}}\{\mathcal{J}[S_{\mathbb{F}}(\Phi)])\} = \mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$ if and only if $\mathrm{Int}[\overline{\mathrm{dec}}\{\mathcal{J}[S_{\mathbb{F}}(\Phi)]\}] \neq \emptyset$.
(iv) If Φ is convex-valued, then $\mathcal{J}[S_{\mathbb{F}}(\Phi)]$ and $\mathrm{dec}[\mathcal{J}[S_{\mathbb{F}}(\Phi)]]$ are convex, and their closures are weakly closed subsets of $\mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$.
(v) If $(\Omega, \mathcal{F}, P)$ is separable, then there exists a sequence $(\varphi^n)_{n=1}^\infty \subset S_{\mathbb{F}}(\Phi)$ such that $\mathcal{J}[S_{\mathbb{F}}(\Phi)] = \mathrm{cl}_{\mathbb{L}}\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \ge 1\}$ and $\overline{\mathrm{dec}}\{\mathcal{J}[S_{\mathbb{F}}(\Phi)]\} = \overline{\mathrm{dec}}\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \ge 1\}$.
(vi) If Φ is convex-valued and square integrably bounded, then there exists a sequence $(\varphi^n)_{n=1}^\infty \subset S_{\mathbb{F}}(\Phi)$ such that $\mathcal{J}[S_{\mathbb{F}}(\Phi)] = \mathrm{cl}_w\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \ge 1\}$ and $\mathrm{cl}_w\{\mathrm{dec}\{\mathcal{J}[S_{\mathbb{F}}(\Phi)]\}\} = \mathrm{cl}_w[\mathrm{dec}\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \ge 1\}]$.
(vii) If Φ and Ψ are convex-valued and square integrably bounded, then $\mathcal{J}[S_{\mathbb{F}}(\Phi + \Psi)] = \mathcal{J}[S_{\mathbb{F}}(\Phi)] + \mathcal{J}[S_{\mathbb{F}}(\Psi)]$.
(viii) If Φ is convex-valued and square integrably bounded and P is nonatomic, then there exists a sequence $(\varphi^n)_{n=1}^\infty \subset S_{\mathbb{F}}(\Phi)$ such that $\overline{\mathrm{dec}}\{\mathcal{J}[S_{\mathbb{F}}(\Phi)]\} = \overline{\mathrm{co}}[\mathrm{dec}\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \ge 1\}]$.

Proof. (i) By the definition of $\mathcal{J}[S_{\mathbb{F}}(\Phi)]$, one has $E\{\mathcal{J}[S_{\mathbb{F}}(\Phi)]\} = \{E[J(\varphi)] = 0 : \varphi \in S_{\mathbb{F}}(\Phi)\}$. Then $E\{\mathcal{J}[S_{\mathbb{F}}(\Phi)]\} = \{0\}$. If Φ is square integrably bounded, then $E\int_0^T \|\Phi_t\|^2 \mathrm{d}t < \infty$. For every $u \in \mathcal{J}[S_{\mathbb{F}}(\Phi)]$, there exists $\varphi \in S_{\mathbb{F}}(\Phi)$ such that $u = \int_0^T \varphi_t \mathrm{d}B_t$. Therefore, for every $u \in \mathcal{J}[S_{\mathbb{F}}(\Phi)]$, one has

$$E|u|^2 = E\left|\int_0^T \varphi_t \mathrm{d}B_t\right|^2 = E\int_0^T |\varphi_t|^2 \mathrm{d}t \le E\int_0^T \|\Phi_t\|^2 \mathrm{d}t.$$

(ii) It is clear that if $\mathcal{J}[S_{\mathbb{F}}(\Phi)]$ is a singleton, then it is decomposable. Suppose $\mathrm{dec}\{\mathcal{J}[S_{\mathbb{F}}(\Phi)]\} = \mathcal{J}[S_{\mathbb{F}}(\Phi)]$. Then for every $A \in \mathcal{F}_T$ and every $u, v \in \mathcal{J}[S_{\mathbb{F}}(\Phi)]$, one has $\mathbb{1}_A u + \mathbb{1}_{\Omega\setminus A} v \in \mathcal{J}[S_{\mathbb{F}}(\Phi)]$. But $\mathbb{1}_A u + \mathbb{1}_{\Omega\setminus A} v = \mathbb{1}_A(u - v) + v$, $E[v] \in E\{\mathcal{J}[S_{\mathbb{F}}(\Phi)]\} = \{0\}$, and $E[\mathbb{1}_A u + \mathbb{1}_{\Omega\setminus A} v] \in E\{\mathcal{J}[S_{\mathbb{F}}(\Phi)]\} = \{0\}$. Therefore, $E[\mathbb{1}_A \mathcal{J}(\varphi - \psi)] \in E\{\mathcal{J}[S_{\mathbb{F}}(\Phi)]\} = \{0\}$ for every $A \in \mathcal{F}_T$, which implies that $\mathcal{J}(\varphi - \psi) = 0$, because $\mathcal{J}(\varphi - \psi)$ is $\mathcal{F}_T$-measurable. Then for every $u, v \in \mathcal{J}[S_{\mathbb{F}}(\Phi)]$, one has $u = v$. Thus $\mathcal{J}[S_{\mathbb{F}}(\Phi)$ is a singleton.

(iii) If $\overline{\text{dec}}\{\mathcal{J}[S_{\mathbb{F}}(\Phi)]\} = \mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$, then $\text{Int}[\overline{\text{dec}}\{\mathcal{J}[S_{\mathbb{F}}(\Phi)]\}] \neq \emptyset$, because in this case, $\overline{\text{dec}}\{\mathcal{J}[S_{\mathbb{F}}(\Phi)]\}$ is an open set. If $\text{Int}[\overline{\text{dec}}\{\mathcal{J}[S_{\mathbb{F}}(\Phi)]\}] \neq \emptyset$, then for every $u \in \text{Int}[\overline{\text{dec}}\{\mathcal{J}[S_{\mathbb{F}}(\Phi)]\}]$, there exists an open ball $\mathcal{B}(u)$ centered at u such that $\mathcal{B}(u) \subset \text{Int}[\overline{\text{dec}}\{\mathcal{J}[S_{\mathbb{F}}(\Phi)]\}] \subset \overline{\text{dec}}\{\mathcal{J}[S_{\mathbb{F}}(\Phi)]\}$, which implies that $\overline{\text{dec}}\{\mathcal{B}(u)\} \subset \overline{\text{dec}}\{\mathcal{J}[S_{\mathbb{F}}(\Phi)]\}$. Hence it follows that $\overline{\text{dec}}\{\mathcal{J}[S_{\mathbb{F}}(\Phi)]\} = \mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$, because $\overline{\text{dec}}\{\mathcal{B}(u)\} = \mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$.

(iv) If Φ is convex-valued, then $S_{\mathbb{F}}(\Phi)$ is closed and convex, which implies that $\mathcal{J}[S_{\mathbb{F}}(\Phi)]$ is convex. By Lemma 1.2, a set-valued integral $\mathcal{J}[S_{\mathbb{F}}(\Phi)]$ is a closed subset of $\mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$. Therefore, by Mazur's theorem ([4], Theorem 9.11), $\mathcal{J}[S_{\mathbb{F}}(\Phi)]$ is sequentially weakly closed and hence a weakly closed subset of $\mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$. The properties of $\overline{\text{dec}}\{\mathcal{J}[S_{\mathbb{F}}(\Phi)]\}$ follow now immediately from Theorem 3.3 of Chap. 2.

(v) Suppose $(\Omega, \mathcal{F}, P)$ is separable. Then $\mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$ is a separable metric space, and hence its closed subset $\mathcal{J}[S_{\mathbb{F}}(\Phi)]$ is a separable metric space. Thus there exists a sequence $(u_n)_{n=1}^{\infty}$ of $\mathcal{J}[S_{\mathbb{F}}(\Phi)]$ such that $\mathcal{J}[S_{\mathbb{F}}(\Phi)] = \text{cl}_{\mathbb{L}}\{u_n : n \geq 1\}$. By the definition of $\mathcal{J}[S_{\mathbb{F}}(\Phi)$, it follows that there exists a sequence $(\varphi^n)_{n=1}^{\infty}$ of $S_{\mathbb{F}}(\Phi)$ such that $u_n = \mathcal{J}(\varphi^n)$ for every $n \geq 1$, which together with the last equality implies that $\mathcal{J}[S_{\mathbb{F}}(\Phi)] = \text{cl}_{\mathbb{L}^2}\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}$. Hence it follows that $\overline{\text{dec}}\{\mathcal{J}[S_{\mathbb{F}}(\Phi)\} = \overline{\text{dec}}\{\text{cl}_{\mathbb{L}}\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}\}$. We shall show now that $\overline{\text{dec}}\{\text{cl}_{\mathbb{L}}\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}\} = \overline{\text{dec}}\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}$. Indeed, it is clear that $\overline{\text{dec}}\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\} \subset \overline{\text{dec}}\{\text{cl}_{\mathbb{L}}\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}\}$. Let $u \in \overline{\text{dec}}\{\text{cl}_{\mathbb{L}}\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}\}$. By the definition of $\overline{\text{dec}}\{\text{cl}_{\mathbb{L}}\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}$, for every $\varepsilon > 0$ there exist an $\mathcal{F}_T$-measurable partition $(A_k)_{k=1}^N$ of Ω and a family $(v_k)_{k=1}^N \subset \text{cl}_{\mathbb{L}}\{\mathcal{J}(\varphi^n) : n \geq 1\}$ such that $E|u - \sum_{k=1}^N \mathbb{1}_{A_k} v_k|^2 < \varepsilon/4$. For every $k = 1, \ldots, N$, there exists a subsequence $(\varphi^{n_j(k)})_{j=1}^{\infty}$ of $(\varphi^n)_{n=1}^{\infty}$ such that $\max_{1 \leq k \leq N} E|v_k - \mathcal{J}(\varphi^{n_j(k)})|^2 \to 0$ as $j \to \infty$. Thus for every $\varepsilon > 0$, there exists $r_N \geq 1$ such that $\max_{1 \leq k \leq N} E|v_k - J_B(\varphi^{n_j(k)})|^2 < \varepsilon/4N$ for $j \geq r_N$. For every $j \geq r_N$, one obtains

$$
\begin{aligned}
E\left|u - \sum_{k=1}^N \mathbb{1}_{A_k} \mathcal{J}(\varphi^{n_j(k)})\right|^2 &\leq 2E\left|u - \sum_{k=1}^N \mathbb{1}_{A_k} v_k\right|^2 \\
&\quad + 2E\left|\sum_{k=1}^N \mathbb{1}_{A_k} \left(v_k - \mathcal{J}(\varphi^{n_j(k)})\right)\right|^2 \\
&\leq \varepsilon/2 + 2E\sum_{k=1}^N \mathbb{1}_{A_k} \left|v_k - \mathcal{J}(\varphi^{n_j(k)})\right|^2 \leq \varepsilon.
\end{aligned}
$$

Then $u \in \overline{\text{dec}}\{\mathcal{J}(\varphi^n) : n \geq 1\}$, and therefore, $\overline{\text{dec}}\{\text{cl}_{\mathbb{L}}\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}\} \subset \overline{\text{dec}}\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}$. Thus $\overline{\text{dec}}\{\mathcal{J}[S_{\mathbb{F}}(\Phi)]\} = \overline{\text{dec}}\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}$.

(vi) Let Φ be convex-valued and square integrably bounded. By Corollary 1.1, a set-valued integral $\mathcal{J}[S_{\mathbb{F}}(\Phi)]$ is a convex weakly compact subset of $\mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$. Then $\mathcal{J}[S_{\mathbb{F}}(\Phi)]$ together with the induced weak topology, generated by the weak topology of $\mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$, is a separable topological space. Thus there exists a sequence $(u_n)_{n=1}^{\infty}$ of $\mathcal{J}[S_{\mathbb{F}}(\Phi)]$ such that $\mathcal{J}[S_{\mathbb{F}}(\Phi)] = \mathrm{cl}_w\{u_n : n \geq 1\}$. For every $n \geq 1$, there exists $\varphi^n \in S_{\mathbb{F}}(\Phi)$ such that $u_n = \int_0^T \varphi_t^n \mathrm{d}B_t$. Therefore, $\mathcal{J}[S_{\mathbb{F}}(\Phi)] = \mathrm{cl}_w\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}$. Hence it follows that $\mathrm{cl}_w[\mathrm{dec}\{\mathcal{J}[S_{\mathbb{F}}(\Phi)]\}] = \mathrm{cl}_w[\mathrm{dec}\{\mathrm{cl}_w\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}\}]$. We shall show that $\mathrm{cl}_w[\mathrm{dec}\{\mathrm{cl}_w\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}\}] = \mathrm{cl}_w[\mathrm{dec}\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}]$. It is clear that $\mathrm{cl}_w[\mathrm{dec}\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}] \subset \mathrm{cl}_w[\mathrm{dec}\{\mathrm{cl}_w\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}\}]$. Let $u \in \mathrm{cl}_w[\mathrm{dec}\{\mathrm{cl}_w\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}\}]$. There exists a sequence $(u_m)_{m=1}^{\infty}$ of $\mathrm{dec}\{\mathrm{cl}_w[\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}\}]$ weakly converging to u, i.e., such that $|\int_A u_m \mathrm{d}P - \int_A u \mathrm{d}P| \to 0$ for every $A \in \mathcal{F}_T$ as $m \to \infty$. For every $m \geq 1$, there exist an $\mathcal{F}_T$-measurable partition $(A_k^m)_{k=1}^{N_m}$ of Ω and a family $(v_k^m)_{k=1}^{N_m} \subset \mathrm{cl}_w\{\mathcal{J}(\varphi^n) : n \geq 1\}$ such that $u_m = \sum_{k=1}^{N_m} \mathbb{1}_{A_k^m} v_k^m$. For every $m \geq 1$ and $k = 1, \ldots, N_m$, there exists a subsequence $(\varphi^{n_j(k,m)})_{j=1}^{\infty}$ of $(\varphi^n)_{n=1}^{\infty}$ such that $|\int_C \mathcal{J}(\varphi^{n_j(k,m)})\mathrm{d}P - \int_C v_k^m \mathrm{d}P| \to 0$ for every $C \in \mathcal{F}_T$ as $j \to \infty$. By the weak compactness of $S_{\mathbb{F}}(\Phi)$, for every $m \geq 1$ and $k = 1, \ldots, N_m$, there exists a subsequence, still denoted by $(\varphi^{n_j(k,m)})_{j=1}^{\infty}$, of $(\varphi^{n_j(k,m)})_{j=1}^{\infty}$ weakly converging to $\varphi^{k,m} \in S_{\mathbb{F}}(\Phi)$, which implies that $|\int_C \mathcal{J}(\varphi^{n_j(k,m)})\mathrm{d}P - \int_C \varphi^{k,m} \mathrm{d}P| \to 0$ for every $C \in \mathcal{F}_T$ as $j \to \infty$. Now, for every $j, m \geq 1$, one has

$$\begin{aligned}
&\left|\int_A u\mathrm{d}P - \int_A \sum_{k=1}^{N_m} \mathbb{1}_{A_k^m} \mathcal{J}(\varphi^{k,m})\mathrm{d}P\right| \\
&\leq \left|\int_A u\mathrm{d}P - \int_A u_m \mathrm{d}P\right| + \left|\int_A u_m \mathrm{d}P - \int_A \sum_{k=1}^{N_m} \mathbb{1}_{A_k^m} J_B(\varphi^{n_j(k,m)})\mathrm{d}P\right| \\
&\quad + \left|\sum_{k=1}^{N_m} \int_A \mathbb{1}_{A_k^m} J_B(\varphi^{n_j(k,m)})\mathrm{d}P - \sum_{k=1}^{N_m} \int_A \mathbb{1}_{A_k^m} \mathcal{J}(\varphi^{k,m})\mathrm{d}P\right| \leq \left|\int_A u\mathrm{d}P - \int_A u_m \mathrm{d}P\right| \\
&\quad + \left|\sum_{k=1}^{N_m} \int_A \mathbb{1}_{A_k^m} v_k^m \mathrm{d}P - \sum_{k=1}^{N_m} \int_A \mathbb{1}_{A_k^m} \mathcal{J}(\varphi^{n_j(k,m)})\mathrm{d}P\right| + \left|\sum_{k=1}^{N_m} \int_A \mathbb{1}_{A_k^m} J_B(\varphi^{n_j(k,m)})\mathrm{d}P\right. \\
&\quad \left. - \sum_{k=1}^{N_m} \int_A \mathbb{1}_{A_k^m} \mathcal{J}(\varphi^{k,m})\mathrm{d}P\right| \leq \left|\int_A (u - u_m)\,\mathrm{d}P\right| \\
&\quad + \left|\sum_{k=1}^{N_m} \int_{A \cap A_k^m} \left(v_k^m - \mathcal{J}(\varphi^{n_j(k,m)}\right)\mathrm{d}P\right| + \left|\sum_{k=1}^{N_m} \int_{A \cap A_k^m} \left(\mathcal{J}(\varphi^{n_j(k,m)}) - \mathcal{J}(\varphi^{k,m})\right)\mathrm{d}P\right|.
\end{aligned}$$

Hence it follows that

$$\left|\int_A u\mathrm{d}P - \int_A \sum_{k=1}^{N_m} \mathbb{1}_{A_k^m} \mathcal{J}(\varphi^{k,m})\mathrm{d}P\right| \leq \left|\int_A (u - u_m)\,\mathrm{d}P\right|$$

for every $A \in \mathcal{F}_T$ and $m \geq 1$. Then a sequence $\left(\sum_{k=1}^{N_m} \mathbb{1}_{A_k^m} \mathcal{J}(\varphi^{k,m})\right)_{m=1}^{\infty}$ of $\mathrm{dec}\{\mathcal{J}(\varphi^n : n \geq 1)\}$ converges weakly to u. Therefore, $u \in \mathrm{cl}_w[\mathrm{dec}\{\mathcal{J}(\varphi^n : n \geq 1)\}]$. Thus $\mathrm{cl}_w[\mathrm{dec}\{\mathrm{cl}_w\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}\}] \subset \mathrm{cl}_w[\mathrm{dec}\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}]$ and then $\mathrm{cl}_w[\mathrm{dec}\{\mathrm{cl}_w\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}\}] = \mathrm{cl}_w[\mathrm{dec}\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}]$.

(vii) By virtue of Lemma 3.4 of Chap. 2 and the weak compactness of $S_{\mathbb{F}}(\Phi)$ and $S_{\mathbb{F}}(\Psi)$, one gets $S_{\mathbb{F}}(\Phi + \Psi) = S_{\mathbb{F}}(\Phi) + S_{\mathbb{F}}(\Psi)$, because F and G are compact-valued and $S_{\mathbb{F}}(\Phi) + S_{\mathbb{F}}(\Psi)$ is convex and weakly compact. Therefore, $\mathcal{J}[S_{\mathbb{F}}(\Phi+\Psi)] = \mathcal{J}[S_{\mathbb{F}}(\Phi)+S_{\mathbb{F}}(\Psi)]$. For every $u \in \mathcal{J}[S_{\mathbb{F}}(\Phi)+S_{\mathbb{F}}(\Psi)]$, there are $\varphi \in S_{\mathbb{F}}(\Phi)$ and $\psi \in S_{\mathbb{F}}(\Psi)$ such that $u = \mathcal{J}(\varphi)+\mathcal{T}(\psi) \in \mathcal{J}[S_{\mathbb{F}}(\Phi)]+\mathcal{J}[S_{\mathbb{F}}(\Psi)]$. Then $\mathcal{J}[S_{\mathbb{F}}(\Phi)+S_{\mathbb{F}}(\Psi)] \subset \mathcal{J}[S_{\mathbb{F}}(\Phi)]+\mathcal{T}[S_{\mathbb{F}}(\Psi)]$. In a similar way, we also get $\mathcal{J}[S_{\mathbb{F}}(\Phi)] + \mathcal{J}[S_{\mathbb{F}}(\Psi)] \subset \mathcal{J}[S_{\mathbb{F}}(\Phi) + S_{\mathbb{F}}(\Psi)]$. Therefore, $\mathcal{J}[S_{\mathbb{F}}(\Phi + \Psi)] = \mathcal{J}[S_{\mathbb{F}}(\Phi)] + \mathcal{J}[S_{\mathbb{F}}(\Psi)]$.

(viii) Let Φ be convex-valued and square integrably bounded. Assume that P is nonatomic. By virtue of Theorem 3.3 of Chap. 2, one has $\mathrm{cl}_w[\mathrm{dec}\{\mathcal{J}[S_{\mathbb{F}}(\Phi))]\}] = \overline{\mathrm{co}}[\mathrm{dec}\{\mathcal{J}[S_{\mathbb{F}}(\Phi))]\}$ and $\mathrm{cl}_w[\mathrm{dec}\{\mathcal{J}(\varphi^n) : n \geq 1\}] = \overline{\mathrm{co}}[\mathrm{dec}\{\mathcal{J}(\varphi^n) : n \geq 1\}]$. By virtue of (iv), the sets $\mathcal{J}[S_{\mathbb{F}}(\Phi)]$ and $\mathrm{dec}\{\mathcal{J}[S_{\mathbb{F}}(\Phi)]\}$ are convex. Therefore, $\overline{\mathrm{co}}[\mathrm{dec}\{\mathcal{J}[S_{\mathbb{F}}(\Phi))]\}] = \mathrm{cl}_{\mathbb{L}}(\mathrm{co}[\mathrm{dec}\{\mathcal{J}[S_{\mathbb{F}}(\Phi))]\}]) = \overline{\mathrm{dec}}\{\mathcal{J}[S_{\mathbb{F}}(\Phi))]\}$. Hence by (vi), it follows that $\overline{\mathrm{dec}}\{\mathcal{J}[S_{\mathbb{F}}(\Phi))])\} = \overline{\mathrm{co}}[\mathrm{dec}\{\mathcal{J}(\varphi^n) : n \geq 1\}]$.

□

Theorem 1.2. *Let $\Phi = (\Phi_t)_{0\leq t\leq T}$ and $\Psi = (\Psi_t)_{0\leq t\leq T}$ be d-dimensional Itô integrable set-valued processes. Then*

(i) If Φ is square integrably bounded, then $J[S_{\mathbb{F}}(\Phi)]$ is a bounded subset of $\mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$.
(ii) $\overline{\mathrm{dec}}\{J[S_{\mathbb{F}}(\Phi)]\} = \mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$ if and only if $\mathrm{Int}[\overline{\mathrm{dec}}\{J[S_{\mathbb{F}}(\Phi)]\}] \neq \emptyset$.
(iii) If Φ is convex-valued and square integrably bounded, then $\overline{\mathrm{dec}}[J[S_{\mathbb{F}}(\Phi)]$ is convex and is a weakly closed subset of $\mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$.
(iv) If $(\Omega, \mathcal{F}, P)$ is separable, then there exists a sequence $(\varphi^n)_{n=1}^{\infty} \subset S_{\mathbb{F}}(\Phi)$ such that $J[S_{\mathbb{F}}(\Phi)] = \mathrm{cl}_{\mathbb{L}}\{\int_0^T \varphi_t^n \mathrm{d}t : n \geq 1\}$ and $\overline{\mathrm{dec}}\{J[S_{\mathbb{F}}(\Phi)]\} = \overline{\mathrm{dec}}\{\int_0^T \varphi_t^n \mathrm{d}t : n \geq 1\}$.
(v) If Φ is convex-valued and square integrably bounded, then there exists a sequence $(\varphi^n)_{n=1}^{\infty} \subset S_{\mathbb{F}}(\Phi)$ such that $J[S_{\mathbb{F}}(\Phi)] = \mathrm{cl}_w\{\int_0^T \varphi_t^n \mathrm{d}t : n \geq 1\}$ and $\mathrm{cl}_w\{\mathrm{dec}\{J[S_{\mathbb{F}}(\Phi)]\} = \mathrm{cl}_w[\mathrm{dec}\{\int_0^T \varphi_t^n \mathrm{d}t : n \geq 1\}]$.
(vi) If Φ and Ψ are convex-valued and square integrably bounded, then $J[S_{\mathbb{F}}(\Phi+\Psi)] = J[S_{\mathbb{F}}(\Phi)] + J[S_{\mathbb{F}}(\Psi)])$.

(vii) *If Φ is convex-valued and square integrably bounded, and P is nonatomic, then there exists a sequence $(\varphi^n)_{n=1}^{\infty} \subset S_{\mathbb{F}}(\Phi)$ such that $\overline{\mathrm{dec}}\{J[S_{\mathbb{F}}(\Phi)]\} = \overline{\mathrm{co}}[\mathrm{dec}\{\int_0^T \varphi_t^n \mathrm{d}t : n \geq 1\}]$,*

Proof. (i) Immediately from Hölder's inequality, it follows that $E|J(\varphi)|^2 \leq T\int_0^T E|\varphi_t|^2 \mathrm{d}t$ for every $\varphi \in S_{\mathbb{F}}(\Phi)$. If Φ is square integrably bounded, then $\sup\{E|u|^2 : u \in J(\Phi)\} = \sup\{E|J(\varphi)|^2 : \varphi \in S_{\mathbb{F}}(\Phi)\} \leq T\int_0^T E\|\Phi_t\|^2 < \infty$. The proofs of (ii)–(vii) are similar to those of (iii)–(viii) of Theorem 1.1. □

Theorem 1.3. *Let $\Phi \in \mathcal{L}^2_{\mathbb{F}}(\mathbb{R}^+ \times \Omega, \mathbb{R}^d)$, $\Psi \in \mathcal{L}^2_{\mathbb{F}}(\mathbb{R}^+ \times \Omega, \mathbb{R}^{d\times m})$, and let $B = (B_t)_{t\geq 0}$ be an m-dimensional $\mathbb{F}$-Brownian motion on $\mathcal{P}_{\mathbb{F}}$. The set-valued mappings $\Lambda : [0,\infty) \ni t \to \overline{\mathrm{dec}}\{J_{0,t}[S_{\mathbb{F}}(\Phi)]\} \in \mathrm{Cl}[\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{R}^d)]$ and $\Gamma : [0,\infty) \ni t \to \overline{\mathrm{dec}}\{\mathcal{J}_{0,t}[S_{\mathbb{F}}(\Phi)]\} \in \mathrm{Cl}[\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{R}^d)]$ are l.s.c.*

Proof. Let H denote the Hausdorff distance on the space $\mathrm{Cl}[\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{R}^d)]$ of all nonempty closed subsets of $\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{R}^d)$. We shall show that the multivalued mapping $\gamma : [0,\infty) \ni t \to \mathcal{J}_{0,t}[S_{\mathbb{F}}(\Phi)] \in \mathrm{Cl}[\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{R}^d)]$ is H-l.s.c., i.e., that for every fixed $t \geq 0$ and every sequence $(t_n)_{n=1}^{\infty}$ of positive numbers t_n converging to t, we have $\overline{H}(\gamma(t), \gamma(t_n)) \to 0$ as $n \to \infty$. Let $t \geq 0$ be fixed and $(t_n)_{n=1}^{\infty}$ a sequence converging to t. Assume $t_n \geq t$ for every $n \geq 1$. For every $\varphi \in S_{\mathbb{F}}(\Phi)$, one has $\mathrm{dist}^2(\int_0^t \varphi_\tau \mathrm{d}B_\tau, \gamma(t_n)) \leq E|\int_0^t \varphi_\tau \mathrm{d}B_\tau - \int_0^{t_n} \varphi_\tau \mathrm{d}B_\tau|^2 = E|\int_t^{t_n} \varphi_\tau \mathrm{d}B_\tau|^2 = E\int_t^{t_n} |\varphi_\tau|^2 \mathrm{d}\tau \leq \int_t^{t_n} E\|\Phi_\tau\|^2 \mathrm{d}\tau$ for every $n \geq 1$. Therefore, $\overline{H}(\gamma(t), \gamma(t_n)) \to 0$ as $n \to \infty$. In a similar way, we can consider the case in which $t > 0$ and every sequence $(t_n)_{n=1}^{\infty}$ of the interval $[0, t]$ converges to t. Thus γ is H-l.s.c. Hence, similarly as in the proof of Remark 3.7 of Chap. 2, it follows that γ is l.s.c. at every $t \geq 0$. Now the lower semicontinuity of Γ follows immediately from Remark 3.4 of Chap. 2. The lower semicontinuity of Λ can be verified in a similar way. □

Theorem 1.4. *Let $\Phi \in \mathcal{L}^2_{\mathbb{F}}(\mathbb{R}^+ \times \Omega, \mathbb{R}^d)$, $\Psi \in \mathcal{L}^2_{\mathbb{F}}(\mathbb{R}^+ \times \Omega, \mathbb{R}^{d\times m})$, and let $x = (x_t)_{t\geq 0}$ be a d-dimensional $\mathbb{F}$-nonanticipative $\mathbb{L}^2$-continuous stochastic process such that*

$$x_t - x_s \in \mathrm{cl}_{\mathbb{L}}\{J_{st}[S_{\mathbb{F}}(\Phi)] + \mathcal{J}_{st}[S_{\mathbb{F}}(\Psi)]\} \tag{1.1}$$

for every $0 \leq s \leq t < \infty$, where $\mathbb{L}^2 =: \mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{R}^d)$. For every $\varepsilon > 0$, there are $f^\varepsilon \in S_{\mathbb{F}}(\Phi)$ and $g^\varepsilon \in S_{\mathbb{F}}(\Psi)$ such that

$$\sup_{t\geq 0} \left\| x_t - x_0 - \int_0^t f_\tau^\varepsilon \mathrm{d}\tau - \int_0^t g_\tau^\varepsilon \mathrm{d}B_\tau \right\| \leq \varepsilon. \tag{1.2}$$

Proof. Let $\varepsilon > 0$ be fixed. Select $\delta \in (0, 1)$ such that

$$\sup_{t\geq 0} \sup_{t\leq s\leq t+\delta} E|x_t - x_s|^2 \leq (\varepsilon/3)^2,$$

$$\sup_{t\geq 0} E\int_t^{t+\delta} \|\Phi_\tau\|^2 \mathrm{d}\tau \leq (1/4)(\varepsilon/3)^2 \quad \text{and} \quad \sup_{t\geq 0} E\int_t^{t+\delta} \|\Psi_\tau\|^2 \mathrm{d}\tau \leq (1/4)(\varepsilon/3)^2.$$

Let $\tau_0^\varepsilon = 0$ and $\tau_k^\varepsilon = k\delta$ for $k = 1, 2, \ldots,$. Select for every $k = 1, 2, \ldots,$ processes $f_k^\varepsilon \in S_{\mathbb{F}}(\Phi)$ and $g_k^\varepsilon \in S_{\mathbb{F}}(\Psi)$ such that

$$\left\| x_{\tau_k^\varepsilon} - x_{\tau_{k-1}^\varepsilon} - \int_{\tau_{k-1}^\varepsilon}^{\tau_k^\varepsilon} f_k^\varepsilon \mathrm{d}\tau - \int_{\tau_{k-1}^\varepsilon}^{\tau_k^\varepsilon} g_k^\varepsilon \mathrm{d}B_\tau \right\| \leq \frac{\varepsilon}{6 \cdot 2^k}.$$

Define $f^\varepsilon = \mathbb{1}_{[0,\tau_1^\varepsilon]} f_1^\varepsilon + \sum_{k=2}^\infty \mathbb{1}_{(\tau_{k-1}^\varepsilon, \tau_k^\varepsilon]} f_k^\varepsilon$ and $g^\varepsilon = \mathbb{1}_{[0,\tau_1^\varepsilon]} g_1^\varepsilon + \sum_{k=2}^\infty \mathbb{1}_{(\tau_{k-1}^\varepsilon, \tau_k^\varepsilon]} g_k^\varepsilon$. By the decomposability of $S_{\mathbb{F}}(\Phi)$ and $S_{\mathbb{F}}(\Psi)$, we have $f^\varepsilon \in S_{\mathbb{F}}(\Phi)$ and $g^\varepsilon \in S_{\mathbb{F}}(\Psi)$. Now we get

$$\sup_{t\geq 0} \left\| x_t - x_0 - \left(\int_0^t f_\tau^\varepsilon \mathrm{d}\tau + \int_0^t g_\tau^\varepsilon \mathrm{d}B_\tau \right) \right\| \leq \left(\sup_{k\geq 1} \sup_{\tau_{k-1}^\varepsilon \leq t \leq \tau_k^\varepsilon} E|x_t - x_{\tau_{k-1}^\varepsilon}|^2 \right)^{\frac{1}{2}}$$
$$+ \left(\sup_{k\geq 1} \sup_{\tau_{k-1}^\varepsilon \leq t \leq \tau_k^\varepsilon} E \left| \int_{\tau_{k-1}^\varepsilon}^t f_\tau^\varepsilon \mathrm{d}\tau + \int_{\tau_{k-1}^\varepsilon}^t g_\tau^\varepsilon \mathrm{d}B_\tau \right|^2 \right)^{\frac{1}{2}}$$
$$+ \sup_{k\geq 1} \left\| \sum_{i=1}^{k-1} \left[x_{\tau_i^\varepsilon} - x_{\tau_{i-1}^\varepsilon} - \int_{\tau_{i-1}^\varepsilon}^{\tau_i^\varepsilon} f_\tau^\varepsilon \mathrm{d}\tau - \int_{\tau_{i-1}^\varepsilon}^{\tau_i^\varepsilon} g_\tau^\varepsilon \mathrm{d}B_\tau \right] \right\|.$$

By the definition of τ_k^ε, we get $\sup_{k\geq 1} \sup_{\tau_{k-1}^\varepsilon \leq t \leq \tau_k^\varepsilon} E|x_t - x_{\tau_{k-1}^\varepsilon}|^2 \leq (\varepsilon/3)^2$ and

$$\sup_{k\geq 1} \sup_{\tau_{k-1}^\varepsilon \leq t \leq \tau_k^\varepsilon} E \left| \int_{\tau_{k-1}^\varepsilon}^t f_\tau^\varepsilon \mathrm{d}\tau + \int_{\tau_{k-1}^\varepsilon}^t g_\tau^\varepsilon \mathrm{d}B_\tau \right|^2$$
$$\leq 2 \cdot \sup_{k\geq 1} \sup_{\tau_{k-1}^\varepsilon \leq t \leq \tau_k^\varepsilon} (t - \tau_{k-1}^\varepsilon) E \int_{\tau_{k-1}^\varepsilon}^t \|\Phi_\tau)\|^2 \mathrm{d}\tau + 2 \cdot \sup_{k\geq 1} \sup_{\tau_{k-1}^\varepsilon \leq t \leq \tau_k^\varepsilon} E \int_{\tau_{k-1}^\varepsilon}^t \|\Psi_\tau)\|^2 \mathrm{d}\tau \leq \frac{1}{2} \left(\frac{\varepsilon}{3}\right)^2$$
$$+ \frac{1}{2} \left(\frac{\varepsilon}{3}\right)^2 = \left(\frac{\varepsilon}{3}\right)^2.$$

Moreover,

$$\sup_{1\leq k \leq N_\varepsilon} \left\| \sum_{i=1}^{k-1} \left[x_{\tau_i^\varepsilon} - x_{\tau_{i-1}^\varepsilon} - \int_{\tau_{i-1}^\varepsilon}^{\tau_i^\varepsilon} f_{\tau d\tau}^\varepsilon - \int_{\tau_{i-1}^\varepsilon}^{\tau_i^\varepsilon} g_\tau^\varepsilon \mathrm{d}B_\tau \right] \right\| \leq \sum_{i=1}^\infty \frac{\varepsilon}{6 \cdot 2^k} < \frac{\varepsilon}{3}.$$

Therefore, (1.2) is satisfied because

$$\sup_{t\geq 0} \left\| x_t - x_0 - \left(\int_0^t f_\tau^\varepsilon \mathrm{d}\tau + \int_0^t g_\tau^\varepsilon \mathrm{d}B_\tau \right) \right\| \leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$ □

Remark 1.2. Immediately from Theorem 1.4, it follows that the above selection theorem is also true for every $\Phi \in \mathcal{L}_{\mathbb{F}}^2(T, \Omega, \mathbb{R}^d)$ and $\Psi \in \mathcal{L}_{\mathbb{F}}^2(T, \Omega, \mathbb{R}^{d\times m})$. □

Remark 1.3. Immediately from Lemma 1.6, it follows that Theorem 1.4 remains true if instead of (1.1), we assume that

$$x_t - x_s \in J_{st}[S_{\mathbb{F}}(\Phi)] + \mathcal{J}_{st}[S_{\mathbb{F}}(\Psi)] \tag{1.3}$$

or

$$x_t - x_s \in \mathrm{cl}_{\mathbb{L}}\,\{J_{st}[S_{\mathbb{F}}(\Phi)]\} + \mathcal{J}_{st}[S_{\mathbb{F}}(\Psi)] \tag{1.4}$$

for every $0 \le s \le t \le T$. If furthermore, $\mathcal{P}_{\mathbb{F}}$ is a separable filtered probability space, then (1.4) is by Lemma 1.7 equivalent to

$$x_t - x_s \in J_{st}[S_{\mathbb{F}}(\overline{\mathrm{co}}\,\Phi)] + \mathcal{J}_{st}[S_{\mathbb{F}}(\Psi)] \tag{1.5}$$

for every $0 \le s \le t \le T$. □

Theorem 1.5. *If the assumptions of Theorem 1.4 are satisfied and* Φ *and* Ψ *are convex-valued, then (1.1) is satisfied if and only if there exist* $f \in S_{\mathbb{F}}(\Phi)$ *and* $g \in S_{\mathbb{F}}(\Psi)$ *such that* $x_t = x_0 + \int_0^t f_\tau \mathrm{d}\tau + \int_0^t g_\tau \mathrm{d}B_\tau$ *a.s. for every* $t \ge 1$.

Proof. It is clear that if there exist $f \in S_{\mathbb{F}}(\Phi)$ and $g \in S_{\mathbb{F}}(\Psi)$ such that $x_t = x_0 + \int_0^t f_\tau \mathrm{d}\tau + \int_0^t g_\tau \mathrm{d}B_\tau$ a.s. for every $t \ge 0$, then (1.1) is satisfied for every $0 \le s \le t < \infty$.

Suppose (1.1) is satisfied. By virtue of Theorem 1.4, there exist sequences $(f^n)_{n=1}^\infty$ and $(g^n)_{n=1}^\infty$ of $S_{\mathbb{F}}(\Phi)$ and $S_{\mathbb{F}}(\Psi)$, respectively, such that $\sup_{t\ge 0} \|x_t - x_0 - \int_0^t f_\tau^n \mathrm{d}\tau - \int_0^t g_\tau^n \mathrm{d}B_\tau\| \to 0$ as $n \to \infty$. Hence in particular, it follows that the sequence $(\int_0^t f_\tau^n \mathrm{d}\tau + \int_0^t g_\tau^n \mathrm{d}B_\tau)_{n=1}^\infty$ converges weakly to $x_t - x_0 \in \mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$ uniformly with respect to $t \ge 0$. By the weak compactness of $S_{\mathbb{F}}(\Phi)$ and $S_{\mathbb{F}}(\Psi)$, sequences $(f^n)_{n=1}^\infty$ and $(g^n)_{n=1}^\infty$ have weak cluster points $f \in S_{\mathbb{F}}(\Phi)$ and $g \in S_{\mathbb{F}}(\Psi)$, respectively. From this and the properties of the mappings J and $\mathcal{J}$, it follows that $\int_0^t f_\tau \mathrm{d}\tau + \int_0^t g_\tau \mathrm{d}B_\tau$ is a weak cluster point of the sequence $(\int_0^t f_\tau^n \mathrm{d}\tau + \int_0^t g_\tau^n \mathrm{d}B_\tau)_{n=1}^\infty$ for every $t \ge 0$. Therefore, $x_t - x_0$ is a modification of $\int_0^t f_\tau \mathrm{d}\tau + \int_0^t g_\tau \mathrm{d}B_\tau$, which implies that $x_t = x_0 + \int_0^t f_\tau \mathrm{d}\tau + \int_0^t g_\tau \mathrm{d}B_\tau$ a.s. for every $t \ge 0$. □

Remark 1.4. If Φ, Ψ and $x = (x_t)_{t\ge 0}$ are as in Theorem 1.4 and condition (1.1) ((1.3), (1.4), (1.5)) is satisfied, then there exist $f \in S_{\mathbb{F}}(\overline{\mathrm{co}}\,\Phi)$ and $g \in S_{\mathbb{F}}(\overline{\mathrm{co}}\,\Psi)$ such that $x_t = x_0 + \int_0^t f_\tau \mathrm{d}\tau + \int_0^t g_\tau \mathrm{d}B_\tau$ a.s. for every $t \ge 0$. In that case, $x = (x_t)_{t\in T}$ possesses an $\mathbb{F}$-nonanticipative continuous modification. □

2 Set-Valued Stochastic Integrals

For fixed $T > 0$ and Aumann and Itô integrable set-valued processes $F \in \mathcal{M}(T, \Omega, \mathbb{R}^d)$, $\Phi \in \mathcal{M}_{\mathbb{F}}(T, \Omega, \mathbb{R}^d)$, and $\Psi \in \mathcal{M}_{\mathbb{F}}(T, \Omega, \mathbb{R}^{d\times m})$, we define set-valued stochastic integrals, denoted by $(\mathcal{A})\int_0^T F_t \mathrm{d}t$, $\int_0^T \Phi_t \mathrm{d}t$ and $\int_0^T \Psi_t \mathrm{d}B_t$, respectively, as $\mathcal{F}_T$-measurable set-valued random variables such that $S_T[(\mathcal{A})\int_0^T F_t\mathrm{d}t] = \mathrm{cl}_{\mathbb{L}}\{J[S(F)]\}$, $S_T[\int_0^T \Phi_t\mathrm{d}t] = \overline{\mathrm{dec}}\{J[S_{\mathbb{F}}(\Psi)]\}$ and $S_T[\int_0^T \Psi_t\mathrm{d}B_t] = \overline{\mathrm{dec}}\{\mathcal{J}[S_{\mathbb{F}}(\Psi)]\}$, where the closures are taken in the norm topology of the space $\mathbb{L}^p(\Omega, \mathcal{F}_T, \mathbb{R}^d)$ with $p = 1$ and $p = 2$, respectively. It is clear that for every $\Phi \in \mathcal{L}^2_{\mathbb{F}}(T, \Omega, \mathbb{R}^d)$, one has $\int_0^T \Phi_t\mathrm{d}t \subset (\mathcal{A})\int_0^T \Phi_t\mathrm{d}t$ a.s. As usual, for every $0 \le s < t \le T$, the stochastic set-valued integrals $(\mathcal{A})\int_s^t F_\tau\mathrm{d}\tau$, $\int_s^t \Phi_\tau\mathrm{d}\tau$, and $\int_s^t \Psi_\tau\mathrm{d}B_\tau$ are defined by setting $(\mathcal{A})\int_s^t F_\tau\mathrm{d}\tau = (\mathcal{A})\int_0^T \mathbb{1}_{[s,t]}F_\tau\mathrm{d}\tau$, $\int_s^t \Phi_\tau\mathrm{d}\tau = \int_0^T \mathbb{1}_{[s,t]}\Phi_\tau\mathrm{d}\tau$, and $\int_s^t \Psi_\tau\mathrm{d}B_\tau = \int_0^T \mathbb{1}_{[s,t]}\Psi_\tau\mathrm{d}B_\tau$. It is clear that they are $\mathcal{F}_t$-measurable set-valued random variables such that $S_t[(\mathcal{A})\int_s^t F_\tau\mathrm{d}\tau] = \mathrm{cl}_{\mathbb{L}}\{J_{st}[S(F)]\}$, $S_t[\int_s^t \Phi_\tau\mathrm{d}\tau] = \overline{\mathrm{dec}}\{J_{s,t}[S_{\mathbb{F}}(\Psi)]\}$, and $S_t[\int_s^t \Psi_\tau\mathrm{d}B_\tau] = \overline{\mathrm{dec}}\{\mathcal{J}_{st}[S_{\mathbb{F}}(\Psi)]\}$.

For given Aumann and Itô integrable set-valued processes $F \in \mathcal{M}(\mathbb{R}^+\times\Omega, \mathbb{R}^d)$, $\Phi \in \mathcal{M}_{\mathbb{F}}(\mathbb{R}^+ \times \Omega, \mathbb{R}^d)$, and $\Psi \in \mathcal{M}_{\mathbb{F}}(\mathbb{R}^+ \times \Omega, \mathbb{R}^{d\times m})$, the set-valued processes $((\mathcal{A})\int_0^t F_\tau\mathrm{d}\tau)_{t\ge 0}$, $(\int_0^t \Phi_\tau\mathrm{d}\tau)_{t\ge 0}$, and $(\int_0^t \Psi_\tau\mathrm{d}B_\tau)_{t\ge 0}$ are said to be indefinite set-valued stochastic integrals of F, Φ, and Ψ, respectively.

Theorem 2.1. *For fixed $T > 0$ and Aumann and Itô integrable set-valued processes $F \in \mathcal{M}(T, \Omega, \mathbb{R}^d)$, $\Phi \in \mathcal{M}_{\mathbb{F}}(T, \Omega, \mathbb{R}^d)$, and $\Psi \in \mathcal{M}_{\mathbb{F}}(T, \Omega, \mathbb{R}^{d\times m})$, set-valued stochastic integrals $(\mathcal{A})\int_0^T F_t\mathrm{d}t$, $\int_0^T \Phi_t\mathrm{d}t$ and $\int_0^T \Psi_t\mathrm{d}B_t$ exist and are $\mathcal{F}_T$-measurable closed-valued set-valued random variables such that*

(i) $(\mathcal{A})\int_0^T F_t\mathrm{d}t$ and $\int_0^T \Phi_t\mathrm{d}t$ are integrably and square integrably bounded, respectively, if F and Φ are integrably and square integrably bounded.

(ii) If $(\Omega, \mathcal{F}, \mathbb{F}, P)$ is a filtered separable probability space and F and Φ are integrably and square integrably bounded, respectively, then $(\mathcal{A})\int_0^T \overline{\mathrm{co}}\,F_t\mathrm{d}t = (\mathcal{A})\int_0^T F_t\mathrm{d}t$ and $\int_0^T \overline{\mathrm{co}}\,\Phi_t\mathrm{d}t = \int_0^T \Phi_t\mathrm{d}t$.

(iii) For every $\mathbb{F}$-nonanticipative d-dimensional stochastic process $x = (x_t)_{0\le t\le T}$,

$$x_t(\omega) - x_s(\omega) \in \mathrm{cl}\left\{\left(\int_s^t \Phi_\tau\mathrm{d}\tau\right)(\omega) + \left(\int_s^t \Psi_\tau\mathrm{d}B_\tau\right)(\omega)\right\} \tag{2.1}$$

is satisfied for every $0 \le s \le t \le T$ and a.e. $\omega \in \Omega$ if and only if $x_t - x_s \in \mathrm{cl}_{\mathbb{L}}\left[\overline{\mathrm{dec}}\{J_{st}[S_{\mathbb{F}}(\Phi)]\} + \overline{\mathrm{dec}}\{\mathcal{J}_{st}[S_{\mathbb{F}}(\Psi)]\}\right]$ for every $0 \le s < t \le T$.

(iv) If $\Phi \in \mathcal{L}^2_{\mathbb{F}}(\mathbb{R}^+ \times \Omega, \mathbb{R}^d)$ and $\Psi \in \mathcal{L}^2_{\mathbb{F}}(T, \Omega, \mathbb{R}^{d\times m})$ are convex-valued, then (2.1) is satisfied for every $0 \le s \le t \le T$ and a.e. $\omega \in \Omega$ if and only if $x_t - x_s \in \overline{\mathrm{dec}}\{J_{st}[S_{\mathbb{F}}(\Phi)]\} + \overline{\mathrm{dec}}\{\mathcal{J}_{st}[S_{\mathbb{F}}(\Psi)]\}$ for every $0 \le s \le t \le T$.

Proof. The existence of stochastic set-valued integrals $(\mathcal{A})\int_0^T F_t\mathrm{d}t$, $\int_0^T \Phi_t\mathrm{d}t$, and $\int_0^T \Psi_t\mathrm{d}B_t$ follows from Lemma 1.3 and Lemma 1.4, respectively. Immediately from the definitions of integrals $(\mathcal{A})\int_0^T F_t\mathrm{d}t$, $\int_0^T \Phi_t\mathrm{d}t$ and $\int_0^T \Psi_t\mathrm{d}B_t$, it follows that they are $\mathcal{F}_T$-measurable set-valued random variables with closed values in $\mathbb{R}^d$.

(i) By Corollary 1.1 and Remark 3.5 of Chap. 2, the sets $J[S(\overline{\mathrm{co}}\,F)]$ and $\overline{\mathrm{dec}}\{J[S_{\mathbb{F}}(\overline{\mathrm{co}}\,\Phi)]\}$ are convex sequentially weakly compact subsets of the Banach spaces $\mathbb{L}(\Omega, \mathcal{F}, \mathbb{R}^d)$ and $\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{R}^d)$, respectively. Therefore,

$S_T[(\mathcal{A})\int_0^T F_t dt]$ and $S_T[\int_0^T \Phi_t dt]$ are bounded subsets of $\mathbb{L}(\Omega, \mathcal{F}, \mathbb{R}^d)$ and $\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{R}^d)$, respectively. Hence by Corollary 3.3 of Chap. 2, it follows that the multifunctions $(\mathcal{A})\int_0^T F_t dt$ and $\int_0^T \Phi_t dt$ are integrably and square integrably bounded, respectively.

(ii) By virtue of Lemma 1.7 and definitions of closed decomposable hull and stochastic set-valued integrals, one has $\overline{\text{dec}}[\text{cl}_{\mathbb{L}}\{J[S_{\mathbb{F}}(\Phi)]\}] = \overline{\text{dec}}[\{J[S_{\mathbb{F}}(\overline{\text{co}}\Phi)]\}]$. Similarly as in the proof of (v) of Theorem 1.1, one has $\overline{\text{dec}}[\text{cl}_{\mathbb{L}}\{J[S_{\mathbb{F}}(\Phi)]\}] = \overline{\text{dec}}[\{J[S_{\mathbb{F}}(\Phi)]\}]$. Then $\overline{\text{dec}}\{J[S_{\mathbb{F}}(\Phi)]\} = \overline{\text{dec}}\{J[S_{\mathbb{F}}(\overline{\text{co}}\,\Phi)]\}$. By the definition of set-valued stochastic integrals, we have $S_T[\int_0^T \Phi_t dt] = \overline{\text{dec}}\{J[S_{\mathbb{F}}(\Phi)]\}$ and $S_T[\int_0^T \overline{\text{co}}\,\Phi_t dt] = \overline{\text{dec}}\{J[S_{\mathbb{F}}(\overline{\text{co}}\,\Phi)]\}$. Therefore, $S_T[\int_0^T \overline{\text{co}}\,\Phi_t dt] = S_T[\int_0^T \Phi_t dt]$. Hence, by Corollary 3.1 of Chap. 2, the second equality of (ii) follows. The first equality of (ii) can be obtained similarly.

(iii) It is clear that (2.1) is satisfied for every $0 \le s \le t \le T$ and a.e. $\omega \in \Omega$ if and only if $x_t - x_s$ is an $\mathcal{F}_t$-measurable selector of $\text{cl}_{\mathbb{L}}\{\int_s^t \Phi_\tau d\tau + \int_s^t \Psi_\tau dB_\tau\}$. By virtue of Lemma 3.4 of Chap. 2, the set of all such selectors is equal to $\text{cl}_{\mathbb{L}}\{S_t[\int_s^t \Phi_\tau d\tau] + S_t[\int_s^t \Psi_\tau dB_\tau]\}$, because $S_t[\int_s^t \Phi_\tau d\tau] \ne \emptyset$ and $S_t[\int_s^t \Psi_\tau dB_\tau] \ne \emptyset$. By the definition of stochastic set-valued integrals, one has $\text{cl}_{\mathbb{L}}\{\int_s^t \Phi_\tau d\tau + \int_s^t \Psi_\tau dB_\tau\} = \text{cl}_{\mathbb{L}}\{\{\overline{\text{dec}}\{J_{st}[S_{\mathbb{F}}(\Psi)]\} + \overline{\text{dec}}\{\mathcal{J}_{st}[S_{\mathbb{F}}(\Psi)]\}\}$. Therefore, (iii) is satisfied.

(iv) Immediately from Theorem 3.3 of Chap. 2 and Remark 3.5 of Chap. 2, it follows that $\overline{\text{dec}}\{J_{st}[S_{\mathbb{F}}(\Phi)]\} + \overline{\text{dec}}\{\mathcal{J}_{st}[S_{\mathbb{F}}(\Psi)]\}$ is a closed subset of $\mathbb{L}^2(\Omega, \mathcal{F}_t, \mathbb{R}^d)$. Indeed, let $(u_n)_{n=1}^\infty$ be a sequence of $\text{cl}_{\mathbb{L}}[\overline{\text{dec}}\{J_{st}[S_{\mathbb{F}}(\Psi)]\} + \overline{\text{dec}}\{\mathcal{J}_{st}[S_{\mathbb{F}}(\Psi)]\}]$ converging to $u \in \mathbb{L}^2(\Omega.\mathcal{F}_T, \mathbb{R}^n)$, and let $(a_n)_{n=1}^\infty$ and $(b_n)_{n=1}^\infty$ be sequences of $\overline{\text{dec}}\{J_{st}[S_{\mathbb{F}}(\Phi)]\}$ and $\overline{\text{dec}}\{\mathcal{J}_{st}[S_{\mathbb{F}}(\Psi)]\}$, respectively, such that $u_n = a_n + b_n$ for $n = 1, 2, \ldots$. By the properties of a set-valued process Φ, it follows that $J_{st}[S_{\mathbb{F}}(\Phi)]$ is convex and integrably bounded. Then, by virtue of Remark 3.5 of Chap. 2, $\overline{\text{dec}}\{J_{st}[S_{\mathbb{F}}(\Phi)]\}$ is sequentially weakly compact. Therefore, there exist a subsequence $(a_{n_k})_{k=1}^\infty$ and $a \in \overline{\text{dec}}\{J_{st}[S_{\mathbb{F}}(\Phi)]\}$ such that $(u_{n_k} - a_{n_k})_{k=1}^\infty$ converges weakly to $u - a$, which implies that a subsequence $(b_{n_k})_{k=1}^\infty$ of $(b_n)_{n=1}^\infty$ converges weakly to $u - a$. By virtue of Theorem 3.3 of Chap. 2, the set $\overline{\text{dec}}\{\mathcal{J}_{st}[S_{\mathbb{F}}(\Psi)]\}$ is sequentially weakly closed. Then $u - a \in \overline{\text{dec}}\{\mathcal{J}_{st}[S_{\mathbb{F}}(\Psi)]\}$, which implies that $u = a + (u - a)$ belongs to $\overline{\text{dec}}\{J_{st}[S_{\mathbb{F}}(\Phi)]\} + \overline{\text{dec}}\{\mathcal{J}_{st}[S_{\mathbb{F}}(\Psi)]\}$. From this and (iii), it follows that (2.1) is satisfied for every $0 \le s < t \le T$ and a.e. $\omega \in \Omega$ if and only if $x_t - x_s$ belongs to $\overline{\text{dec}}\{J_{st}[(S_{\mathbb{F}}(\Psi))]\} + \overline{\text{dec}}\{\mathcal{J}_{st}[S_{\mathbb{F}}(\Psi)]\}$ for every $0 \le s \le t \le T$.

□

Theorem 2.2. *Let $B = (B_t)_{t \ge 0}$ be an m-dimensional $\mathbb{F}$-Brownian motion, $\Phi = (\Phi_t)_{0 \le t \le T}$, and let $\Psi = (\Psi_t)_{0 \le t \le T}$ be $d \times m$-dimensional Itô integrable processes on $\mathcal{P}_{\mathbb{F}}$. Then*

(i) $S_T(\int_0^T \Phi_t dB_t) \ne \mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$ *if* $Int[S_T(\int_0^T \Phi_t dB_t)] = \emptyset$.

(ii) $\int_0^T \Phi_t \mathrm{d}B_t$ *is convex-valued if* Φ *is convex-valued.*
(iii) If $(\Omega, \mathcal{F}, P)$ *is separable, then there exists a sequence* $(\varphi^n)_{n=1}^\infty \subset S_{\mathbb{F}}(\Phi)$ *such that* $(\int_0^T \Phi_t \mathrm{d}B_t)(\omega) = \mathrm{cl}\{(\int_0^T \varphi_t^n \mathrm{d}B_t)(\omega) : n \geq 1\}$ *for a.e.* $\omega \in \Omega$.
(iv) If Φ *is convex-valued abnd square integrably bounded, and* P *is nonatomic, then there exists a sequence* $(\varphi^n)_{n=1}^\infty \subset S_{\mathbb{F}}(\Phi)$ *such that* $(\int_0^T \Phi_t \mathrm{d}B_t)(\omega) = \overline{\mathrm{co}}\{(\int_0^T \varphi_t^n \mathrm{d}B_t)(\omega) : n \geq 1\}$ *for a.e.* $\omega \in \Omega$.

Proof. (i) The result follows immediately from (iii) of Theorem 1.1.
(ii) By the definition of $\int_0^T \Phi_t \mathrm{d}B_t$ and Theorem 3.3 of Chap. 2, one gets

$$S_T\left(\int_0^T \Phi_t \mathrm{d}B_t\right) = \overline{\mathrm{dec}}\mathcal{J}[S_{\mathbb{F}}(\Phi)]) = \overline{\mathrm{co}}\left[\overline{\mathrm{dec}}\mathcal{J}[S_{\mathbb{F}}(\Phi)]\right] = \overline{\mathrm{co}}\left[S_T\left(\int_0^T \Phi_t \mathrm{d}B_t\right)\right].$$

But $S_T\left(\int_0^T \Phi_t \mathrm{d}B_t\right)$ is a closed convex subset of $\mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$. Therefore, by virtue of Lemma 3.3 of Chap. 2, the above equalities imply

$$S_T\left(\int_0^T \Phi_t \mathrm{d}B_t\right) = \overline{\mathrm{co}}\left[S_T\left(\int_0^T \Phi_t \mathrm{d}B_t\right)\right] = S_T\left(\overline{\mathrm{co}}\int_0^T \Phi_t \mathrm{d}B_t\right).$$

Hence, by (vi) of Lemma 4.1 of Chap. 2, it follows that $\int_0^T \Phi_t \mathrm{d}B_t = \overline{\mathrm{co}}\int_0^T \Phi_t \mathrm{d}B_t$ a.s.
(iii) By the definition of $\int_0^T \Phi_t \mathrm{d}B_t$ and (v) of Theorem 1.1, one has $S_T(\int_0^T \Phi_t \mathrm{d}B_t) = \overline{\mathrm{dec}}\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}$, where $(\varphi^n)_{n=1}^\infty \subset S_{\mathbb{F}}(\Phi)$ is such that $\mathcal{J}[S_{\mathbb{F}}(\Phi)] = \mathrm{cl}_{\mathbb{L}}\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}$. Let $\Lambda : \Omega \to \mathrm{Cl}(\mathbb{R}^r)$ be a set-valued random variable defined by $\Lambda(\omega) = \mathrm{cl}\{(\int_0^T \varphi_t^n \mathrm{d}B_t)(\omega) : n \geq 1\}$ for $\omega \in \Omega$. By (iv) of Lemma 4.1 of Chap. 2, one has $S_T(\Lambda) = \overline{\mathrm{dec}}\{\mathcal{J}(\varphi^n) : n \geq 1\}$. Then $S_T(\int_0^T \Phi_t \mathrm{d}B_t) = \overline{\mathrm{dec}}\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\} = S_T(\Lambda)$, which by (vi) of Lemma 4.1 of Chap. 2, implies that $\int_0^T \Phi_t \mathrm{d}B_t = \Lambda$ a.s. Thus $\int_0^T \Phi_t \mathrm{d}B_t = \mathrm{cl}\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}$ a.s.
(iv) By (viii) of Theorem 1.1, one has $\overline{\mathrm{dec}}\{\mathcal{J}[S_{\mathbb{F}}(\Phi)]\} = \overline{\mathrm{co}}[\mathrm{dec}\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}]$, where $(\varphi^n)_{n=1}^\infty \subset S_{\mathbb{F}}(\Phi)$ is such that $\mathcal{J}[S_{\mathbb{F}}(\Phi)] = \mathrm{cl}_w\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}$. From this and the definition of the integral $\int_0^T \Phi_t \mathrm{d}B_t$, it follows that $S_T(\int_0^T \Phi_t \mathrm{d}B_t) = \overline{\mathrm{co}}[\mathrm{dec}\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}]$. Let $G(\omega) = \overline{\mathrm{co}}[\mathrm{cl}\{(\int_0^T \varphi_t^n \mathrm{d}B_t)(\omega) : n \geq 1\}]$ for a.e. $\omega \in \Omega$. Similarly as above, for every $n \geq 1$ we have $\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\} \subset S_T(G)$. But $S_T(G)$ is a closed, convex, and decomposable subset of $\mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$. Therefore, $\overline{\mathrm{co}}[\mathrm{dec}\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}] \subset S_T(G)$. On the other hand, by (i) of Lemma 4.1 of Chap. 2, one has $S_T(G) = S_T(\overline{\mathrm{co}}[\mathrm{cl}\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}]) = \overline{\mathrm{co}}[S_T(\mathrm{cl}\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\})]$. Then for every $u \in S_T(G)$, there exists a sequence $(u_m)_{m=1}^\infty$ of $\mathrm{co}[S_T(\mathrm{cl}\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\})]$ such that $E|u - u_m|^2 \to 0$ as $m \to \infty$.

Therefore, for every $\varepsilon > 0$, there exists $m_\varepsilon \geq 1$ such that $E|u - u_m|^2 \leq \varepsilon/2$ for every $m \geq m_\varepsilon$. Now for every $m \geq m_\varepsilon$, there exist $\lambda_1^m, \ldots, \lambda_{k(m)}^m \in [0, 1]$ and $v_1^m, \ldots, v_{k(m)}^m \in S_T(\mathrm{cl}\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\})$ such that $\sum_{k=1}^{k(m)} \lambda_k^m = 1$ and $u_m = \sum_{k=1}^{k(m)} \lambda_k^m v_k^m$ a.s. By (v) of Lemma 4.1 of Chap. 2, for every $\eta > 0$, $\varepsilon > 0$, $m \geq m_\varepsilon$ and $k = 1, \ldots, k(m)$, there exist an $\mathcal{F}_T$-measurable partition $(A_j^{k,m})_{j=1}^{N(k,m)}$ of Ω and a family $(\varphi^{n_j(k,m)})_{j=1}^{N(k,m)} \subset \{\varphi^n : n \geq 1\}$ such that $E|v_k^m - \sum_{j=1}^{N(k,m)} \mathbb{1}_{A_j^{k,m}} \int_0^T \varphi_t^{n_j(k,m)} \mathrm{d}B_t|^2 \leq \eta/2k(m_\varepsilon)M(\varepsilon)$, where $M(\varepsilon) = \sum_{k=1}^{k(m_\varepsilon)} (\lambda_k^{m_\varepsilon})^2$. Hence for every $m \geq m_\varepsilon$, it follows that

$$
\begin{aligned}
&E\left|u - \sum_{k=1}^{k(m)} \lambda_k^m \left(\sum_{j=1}^{N(k,m)} \mathbb{1}_{A_j^{k,m}} \int_0^T \varphi_t^{n_j(k,m)} \mathrm{d}B_t\right)\right|^2 \\
&\leq 2E|u - u_m|^2 + 2E\left|\sum_{k=1}^{k(m)} \lambda_k^m v_k^m - \sum_{k=1}^{k(m)} \lambda_k^m \left(\sum_{j=1}^{N(k,m)} \mathbb{1}_{A_j^{k,m}} \int_0^T \varphi_t^{n_j(k,m)} \mathrm{d}B_t\right)\right|^2 \\
&\leq \varepsilon + 2E\left[\sum_{k=1}^{k(m)} \lambda_k^m \left|v_k^m - \sum_{j=1}^{N(k,m)} \mathbb{1}_{A_j^{k,m}} \int_0^T \varphi_t^{n_j(k,m)} \mathrm{d}B_t\right|\right]^2 = \varepsilon + 2E\,[\langle \lambda^m, \xi^m \rangle]^2,
\end{aligned}
$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product in $\mathbb{R}^{k(m)}$, $\lambda^m = (\lambda_1^m, \ldots, \lambda_{k(m)}^m)$, and $\xi^m = (\xi_1^m, \ldots, \xi_{k(m)}^m)$ with $\xi_k^m = |v_k^m - \sum_{j=1}^{N(k,m)} \mathbb{1}_{A_j^{k,m}} \int_0^T \varphi_t^{n_j(k,m)} \mathrm{d}B_t|$ for $k = 1, \ldots, k(m)$ and $m \geq m_\varepsilon$. But $E[\langle \lambda^m, \xi^m \rangle]^2 \leq E[\langle \lambda^m, \lambda^m \rangle \langle \xi^m, \xi^m \rangle] = |\lambda^m|^2 E[|\xi^m|^2] \leq |\lambda^m|^2 k(m)\eta/2k(m_\varepsilon)M(\varepsilon)$ for $m \geq m_\varepsilon$. In particular, we have $E[\langle \lambda^{m_\varepsilon}, \xi^{m_\varepsilon} \rangle]^2 \leq |\lambda^{m_\varepsilon}|^2 k(m_\varepsilon)\eta/2k(m_\varepsilon)M(\varepsilon) = \eta/2$ for every $\varepsilon > 0$ and $\eta > 0$. Then for every $\varepsilon > 0$ and $\eta > 0$, one gets

$$
E\left|u - \sum_{k=1}^{k(m_\varepsilon)} \lambda_k^{m_\varepsilon} \left(\sum_{j=1}^{N(k,m_\varepsilon)} \mathbb{1}_{A_j^{k,m_\varepsilon}} \int_0^T \varphi_t^{n_j(k,m_\varepsilon)} \mathrm{d}B_t\right)\right|^2 \leq \varepsilon + \eta.
$$

Taking in particular $\varepsilon = 1/2r$ and $\eta = 1/2r$ for $r \geq 1$, we obtain a sequence $(z_r)_{r=1}^\infty$ of $\mathrm{co}[dec\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}$ such that $E|z_r - u|^2 \to 0$ as $r \to \infty$. Therefore, $u \in \overline{\mathrm{co}}[dec\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}$, which implies that $S_T(G) \subset \overline{\mathrm{co}}[dec\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}$. Thus $S_T(G) = \overline{\mathrm{co}}[dec\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}$. But by (viii) of Theorem 1.1 and the definition of $\int_0^T \Phi_t \mathrm{d}B_t$, one has $S_T(\int_0^T \Phi_t \mathrm{d}B_t) = \overline{\mathrm{co}}[\mathrm{dec}\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}]$. Therefore, $S_T(\int_0^T \Phi_t \mathrm{d}B_t) = S_T(G)$, which by (vi) of Lemma 4.1 of Chap. 2, implies that $\int_0^T \Phi_t \mathrm{d}B_t = G$ a.s. Thus $(\int_0^T \Phi_t \mathrm{d}B_t)(\omega) = \overline{\mathrm{co}}[\mathrm{cl}\{(\int_0^T \varphi_t^n \mathrm{d}B_t)$

$(\omega) : n \geq 1\}] = \overline{\mathrm{co}}\{(\int_0^T \varphi_t^n \mathrm{d}B_t)(\omega) : n \geq 1\}$ for a.e. $\omega \in \Omega$, because $\overline{\mathrm{co}}(\overline{A}) = \overline{\mathrm{co}}(A)$ for $A \subset \mathbb{R}^d$.

□

Remark 2.1. In what follows, for a compact set $K \subset \mathbb{R}^d$, we shall write $\mathrm{co}(K)$ instead of $\overline{\mathrm{co}}(K)$, because for a compact set $K \subset \mathbb{R}^d$, a set $\mathrm{co}(K)$ is compact. □

Theorem 2.3. *Let $\Phi = (\Phi_t)_{0\leq t\leq T}$ and $\Psi = (\Psi_t)_{0\leq t\leq T}$ be d-dimensional Itô integrable set-valued processes on $\mathcal{P}_{\mathbb{F}}$. Then*

(i) $\int_0^T \Phi_t \mathrm{d}t$ is square integrably bounded if Φ is square integrably bounded.
(ii) $\int_0^T \Phi_t \mathrm{d}t$ is convex-valued if Φ is convex-valued.
(iii) If $(\Omega, \mathcal{F}, P)$ is separable, then there exists a sequence $(\varphi^n)_{n=1}^\infty \subset S_{\mathbb{F}}(\Phi)$ such that $(\int_0^T \Phi_t \mathrm{d}t)(\omega) = \mathrm{cl}\{(\int_0^T \varphi_t^n \mathrm{d}t)(\omega) : n \geq 1\}$ for a.e. $\omega \in \Omega$.
(iv) If Φ is convex-valued and square integrably bounded, and P is nonatomic, then there exists a sequence $(\varphi^n)_{n=1}^\infty \subset S_{\mathbb{F}}(\Phi)$ such that $(\int_0^T \Phi_t \mathrm{d}t)(\omega) = \overline{\mathrm{co}}\{(\int_0^T \varphi_t^n \mathrm{d}t)(\omega) : n \geq 1\}$ for a.e. $\omega \in \Omega$.
(v) If Φ and Ψ are convex-valued and square integrably bounded, and P is nonatomic, then $\int_0^T (\Phi + \Psi)_t \mathrm{d}t = \int_0^T \Phi_t \mathrm{d}t + \int_0^T \Psi_t \mathrm{d}t$ a.s.

Proof. (i) By the definition of $\int_0^T \Phi_t \mathrm{d}t$, one has $\sup\{E|u|^2 : u \in S(\int_0^T \Phi_t \mathrm{d}t)\} = \sup\{E|u|^2 : u \in \mathrm{dec}\{J[S_{\mathbb{F}}(\Phi)]\}\}$. For every $u \in \mathrm{dec}[J_{\mathbb{F}}(\Phi)]$, there exist a partition $(A_k)_{k=1}^N \in \Pi(\Omega, \mathcal{F}_T)$ and a family $(\varphi^k)_{k=1}^N \subset S_{\mathbb{F}}(\Phi)$ such that $u = \sum_{k=1}^N \mathbb{1}_{A_k} J(\varphi^k)$. Hence it follows that $E|u|^2 \leq E[\max_{1\leq k\leq N} |J(\varphi^k)|^2]$. But $\max_{1\leq k\leq N} |J(\varphi^k)|^2 \leq \int_0^T \|\Phi_t\|^2 \mathrm{d}t$. Then $E|u|^2 \leq \int_0^T E\|\Phi_t\|^2 \mathrm{d}t < \infty$ for every $u \in \mathrm{dec}\{J[S_{\mathbb{F}}(\Phi)]\}$, which implies that $S_T(\int_0^T \Phi_t dt)$ is a bounded subset of $\mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$. Therefore, by (ii) of Lemma 4.1 of Chap. 2, $\int_0^T \Phi_t \mathrm{d}t$ is square integrably bounded. Conditions (ii)–(iv) can be verified similarly to the verification of (ii)–(iv) of Theorem 2.2.

(v) Let $(\varphi^n)_{n=1}^\infty \subset S_{\mathbb{F}}(\Phi)$ and $(\psi^m)_{m=1}^\infty \subset S_{\mathbb{F}}(\Psi)$ be such that $J[S_{\mathbb{F}}(\Phi)] = \mathrm{cl}_w\{\int_0^T \varphi_t^n \mathrm{d}t : n \geq 1\}$ and $J[S_{\mathbb{F}}(\Psi)] = \mathrm{cl}_w\{\int_0^T \psi_t^m \mathrm{d}t : m \geq 1\}$. By (vi) of Theorem 1.2, one has $J[S_{\mathbb{F}}(\Phi + \Psi)] = J[S_{\mathbb{F}}(\Phi)] + J[S_{\mathbb{F}}(\Psi)]$. Therefore, for every $n, m \geq 1$, one has $J(\varphi^n + \psi^m) \in J[S_{\mathbb{F}}(\Phi + \Psi)]$, which implies that $\mathrm{cl}_w\{\int_0^T (\varphi_t^n + \psi_t^m)\mathrm{d}t : n, m \geq 1\} \subset J[S_{\mathbb{F}}(\Phi + \Psi)]$, because $J[S_{\mathbb{F}}(\Phi + \Psi)]$ is a weakly compact subset of $\mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$. For every $u \in J[S_{\mathbb{F}}(\Phi + \Psi)]$, there exists $(\varphi, \psi) \in S_{\mathbb{F}}(\Phi) \times S_{\mathbb{F}}(\Psi)$ such that $u = J(\varphi + \psi)$. On the other hand, by the properties of sequences $(\varphi^n)_{n=1}^\infty$ and $(\psi^m)_{m=1}^\infty$, there exist subsequences $(\varphi^{n_k})_{k=1}^\infty$ and $(\psi^{n_k})_{k=1}^\infty$ of $(\varphi^n)_{n=1}^\infty$ and $(\psi^m)_{m=1}^\infty$, respectively, such that $(J(\varphi^{n_k}))_{k=1}^\infty$ and $(J(\psi^{n_k}))_{k=1}^\infty$ converge weakly to $J(\varphi)$ and $J(\psi)$, respectively. Therefore, $(J(\varphi^{n_k} + \psi^{n_k}))_{k=1}^\infty$ converges weakly to $J(\varphi + \psi) = u$, which implies that $u \in \mathrm{cl}_w\{\int_0^T (\varphi_t^n + \psi^m)\mathrm{d}t : n, m \geq 1\}$. Then $J[S_{\mathbb{F}}(\Phi + \Psi)] \subset \mathrm{cl}_w\{\int_0^T (\varphi_t^n + \psi^m)\mathrm{d}t : n, m \geq 1\}$. Now, by virtue of (iv), we get $\int_0^T (\Phi_t + \Psi_t)\mathrm{d}t = \overline{\mathrm{co}}\{\int_0^T (\varphi_t^n + \psi_t^m)\mathrm{d}t : n, m \geq 1\} = \overline{\mathrm{co}}\{\int_0^T \varphi_t^n \mathrm{d}t + \int_0^T \psi_t^m \mathrm{d}t : n, m \geq 1\}$ a.s. Let $A(\omega) = \{(\int_0^T \varphi_t^n \mathrm{d}t)(\omega) : n \geq 1\}$

and $B(\omega) = \{(\int_0^T \psi_t^m \mathrm{d}t)(\omega) : m \geq 1\}$ for fixed $\omega \in \Omega$. It is clear that $\overline{A(\omega)}$ and $\overline{B(\omega)}$ are compact subsets of $\mathbb{R}^d$ for a.e. $\omega \in \Omega$, because for every $n, m \geq 1$ and a.e. $\omega \in \Omega$, one has $|\int_0^T \varphi_t^n \mathrm{d}t| \leq \int_0^T |\varphi_t^n|\mathrm{d}t \leq \int_0^T m_t(\omega)\mathrm{d}t$ and $|\int_0^T \psi_t^m \mathrm{d}t| \leq \int_0^T |\psi_t^m|\mathrm{d}t \leq \int_0^T m_t(\omega)\mathrm{d}t$, where $m \in \mathbb{L}^2([0,T] \times \Omega, \Sigma_{\mathbb{F}}, \mathbb{R}^+)$ is such that $\max\{\|\Phi_t(\omega)\|, \|\Psi_t(\omega)\|\} \leq m_t(\omega)$ for a.e. $(t,\omega) \in [0,T] \times \Omega$, and therefore, $\overline{A(\omega)}$ and $\overline{B(\omega)}$ are bounded subsets of $\mathbb{R}^d$ for a.e. $\omega \in \Omega$. Then $\overline{A(\omega) + B(\omega)} = \overline{A(\omega)} + \overline{B(\omega)}$, which implies that $\overline{\mathrm{co}}\{\overline{A(\omega)} + \overline{B(\omega)}\} = \overline{\mathrm{co}}\{\overline{A(\omega)}\} + \overline{\mathrm{co}}\{\overline{B(\omega)}\}$, because $\overline{\mathrm{co}}(\overline{A(\omega)})$ is compact. Thus $\overline{\mathrm{co}}\{\int_0^T \varphi_t^n \mathrm{d}t + \int_0^T \psi_t^m \mathrm{d}t : n, m \geq 1\} = \overline{\mathrm{co}}\{(\int_0^T \varphi_t^n \mathrm{d}t)(\omega) : n \geq 1\} + \overline{\mathrm{co}}\{(\int_0^T \psi_t^m \mathrm{d}t)(\omega) : m \geq 1\}$. Hence, by virtue of (iv), it follows that $\int_0^T (\Phi_t + \Psi_t)\mathrm{d}t = \int_0^T \Phi_t \mathrm{d}t + \int_0^T \Psi_t \mathrm{d}t$ a.s., because $\overline{\mathrm{co}}\{(\int_0^T \varphi_t^n \mathrm{d}t)(\omega) : n \geq 1\} = \int_0^T \Phi_t \mathrm{d}t$ and $\overline{\mathrm{co}}\{(\int_0^T \psi_t^m \mathrm{d}t)(\omega) : m \geq 1\} = \int_0^T \Psi_t \mathrm{d}t$ a.s. □

Theorem 2.4. *Let $B = (B_t)_{t\geq 0}$ be an m-dimensional $\mathbb{F}$-Brownian motion and $\Phi = (\Phi_t)_{0\leq t\leq T}$ an $r \times m$-dimensional Itô integrable set-valued process. Then*

(i) If Φ is square integrably bounded, then $S_T(\int_0^T \Phi_t \mathrm{d}t) \neq \overline{\mathrm{dec}}\{\mathcal{J}(\mathbb{L}^2_{\mathbb{F}})\}$, where $\mathbb{L}^2_{\mathbb{F}} =: \mathbb{L}^2([0,T] \times \Omega, \Sigma_{\mathbb{F}}, \mathbb{R}^{d\times m})$.

(ii) If $(\Omega, \mathcal{F}, P)$ is separable, then

$$E\left\|\int_0^T \Phi_t \mathrm{d}B_t\right\|^2 \leq \sup_{(\varphi^{n_k})_{k=1}^N \subset \{\varphi^n : n\geq 1\}} E\left[\max_{1\leq k\leq N}\left|\int_0^T \varphi_t^{n_k} \mathrm{d}B_t\right|^2\right]$$

for every sequence $(\varphi^n)_{n=1}^\infty \subset S_{\mathbb{F}}(\Phi)$ such that $\int_0^T \Phi_t \mathrm{d}B_t = \mathrm{cl}\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}$ a.s.

(iii) If Φ is convex-valued and square integrably bounded, and P is nonatomic, then

$$E\left\|\int_0^T \Phi_t \mathrm{d}B_t\right\|^2 \leq \sup_{(\varphi^{n_k})_{k=1}^N \subset \{\varphi^n : n\geq 1\}} E\left[\max_{1\leq k\leq N}\left|\int_0^T \varphi_t^{n_k} \mathrm{d}B_t\right|^2\right]$$

for every sequence $(\varphi^n)_{n=1}^\infty \subset S_{\mathbb{F}}(\Phi)$ such that $\int_0^T \Phi_t \mathrm{d}B_t = \overline{\mathrm{co}}\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}$ a.s.

Proof. (i) By virtue of (iii) of Lemma 4.1 of Chap. 2, one has $S_{\mathbb{F}}(\Phi) \neq \mathbb{L}^2_{\mathbb{F}}$. From this and the definition of $\int_0^T \Phi_t \mathrm{d}B_t$, the result follows.

(ii) Let $(\varphi^n)_{n=1}^\infty$ be any sequence of $S_{\mathbb{F}}(\Phi)$ such that $\int_0^T \Phi_t \mathrm{d}B_t = \mathrm{cl}\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}$ a.s. By virtue of Corollary 3.2 of Chap. 2, one has

$$E\left\|\int_0^T \Phi_t \mathrm{d}B_t\right\|^2 = E\left[\sup\left\{|x|^2 : x \in \left(\int_0^T \Phi_t \mathrm{d}B_t\right)(\omega)\right\}\right]$$
$$= \sup\left\{E|u|^2 : u \in S_T\left(\int_0^T \Phi_t \mathrm{d}B_t\right)\right\}.$$

By (v) of Theorem 1.1, we have $S_T(\int_0^T \Phi_t \mathrm{d}B_t) = \overline{\mathrm{dec}}\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}$. Therefore,

$$\sup\left\{E|u|^2 : u \in S_T\left(\int_0^T \Phi_t \mathrm{d}B_t\right)\right\}$$
$$= \sup\left\{E|u|^2 : u \in \overline{\mathrm{dec}}\left\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\right\}\right\}$$
$$= \sup\left\{E|u|^2 : u \in \mathrm{dec}\left\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\right\}\right\}$$
$$= \sup\left\{E\left|\sum_{k=1}^N \mathbb{1}_{A_k} \int_0^T \varphi_t^{n_k} \mathrm{d}B_t\right|^2 : (A_k)_{k=1}^N \in \Pi(\Omega, \mathcal{F}_T), (\varphi^{n_k})_{k=1}^N \subset \{\varphi^n : n \geq 1\}\right\}$$
$$\leq \sup_{(\varphi^{n_k})_{k=1}^N \subset \{\varphi^n : n \geq 1\}} E\left[\max_{1\leq k\leq N}\left|\int_0^T \varphi_t^{n_k} \mathrm{d}B_t\right|^2\right].$$

The result follows from this and the first inequality.

(iii) Let $(\varphi^n)_{n=1}^\infty$ be any sequence of $S_{\mathbb{F}}(\Phi)$ such that $\int_0^T \Phi_t \mathrm{d}B_t = \overline{\mathrm{co}}[\mathrm{cl}\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}]$ a.s. By the proof of (iv) of Theorem 2.2, it follows that in this case, we have $S_T(\int_0^T \Phi_t \mathrm{d}B_t) = \overline{\mathrm{co}}[\mathrm{dec}\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}]$. Therefore, similarly as above, we get

$$E\left\|\int_0^T \Phi_t \mathrm{d}B_t\right\|^2 = \sup\left\{E|u|^2 : u \in \overline{\mathrm{co}}\left[\mathrm{dec}\left\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\right\}\right]\right\}$$
$$= H\left(\overline{\mathrm{co}}\left[\mathrm{dec}\left\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\right\}\right], \{0\}\right) \leq H\left(\mathrm{dec}\left\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\right\}, \{0\}\right)$$
$$= \sup\left\{E|u|^2 : u \in \mathrm{dec}\left\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\right\}\right\},$$

where H denotes the Hausdorff metric on $\mathrm{Cl}(\mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$. Hence, similarly as in case (ii), it follows that

$$E\left\|\int_0^T \Phi_t \mathrm{d}B_t\right\|^2 \leq \sup_{(\varphi^{n_k})_{k=1}^N \subset \{\varphi^n : n \geq 1\}} E\left[\max_{1\leq k\leq N}\left|\int_0^T \varphi_t^{n_k} \mathrm{d}B_t\right|^2\right].$$

□

Corollary 2.1. *If the assumptions of Theorem 2.4 are satisfied, then* $\int_0^T \Phi_t \mathrm{d}B_t$ *is square integrably bounded if and only if there exists a sequence* $(\varphi^n)_{n=1}^\infty$ *of* $S_{\mathbb{F}}(\Phi)$ *with properties defined in (ii) or (iii) of Theorem 2.4 such that the sequence* $(\int_0^T \varphi_t^n \mathrm{d}B_t)_{n=1}^\infty$ *of d-dimensional random variables is square integrably bounded.*

Proof. If $\int_0^T \Phi_t \mathrm{d}B_t$ is square integrably bounded, then by (ii) of Lemma 4.1 of Chap. 2, the set $S_T(\int_0^T \Phi_t \mathrm{d}B_t)$ is a nonempty bounded subset of $\mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$. Therefore, for every sequence $(\varphi^n)_{n=1}^\infty \subset S_{\mathbb{F}}(\Phi)$ such that $\int_0^T \Phi_t \mathrm{d}B_t = \mathrm{cl}\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}$ a.s., the set $\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}$ is a bounded subset of this space, which implies that the sequence $(\int_0^T \varphi_t^n \mathrm{d}B_t)_{n=1}^\infty$ is square integrably bounded. Conversely, if there exists a sequence $(\varphi^n)_{n=1}^\infty \subset S_{\mathbb{F}}(\Phi)$ such that $\int_0^T \Phi_t \mathrm{d}B_t = \mathrm{cl}\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}$ a.s., and the sequence $(\int_0^T \varphi_t^n \mathrm{d}B_t)_{n=1}^\infty$ is square integrably bounded, then there exists $m \in \mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^+)$ such that $|\int_0^T \varphi_t^n \mathrm{d}B_t| \leq m$ a.s. for every $n \geq 1$. But $S_T(\int_0^T \Phi_t \mathrm{d}B_t) = \overline{\mathrm{dec}}\{\int_0^T \varphi_t^n \mathrm{d}B_t : n \geq 1\}$. Therefore, for every $u \in S_T(\int_0^T \Phi_t \mathrm{d}B_t)$ and $\varepsilon > 0$, there exist a partition $(A_k^\varepsilon)_{k=1}^{N_\varepsilon} \in \Pi(\Omega, \mathcal{F}_T)$ and a family $(\varphi_k^\varepsilon)_{k=1}^N \subset \{\varphi^n : n \geq 1\}$ such that $E|u - \sum_{k=1}^{N_\varepsilon} \mathbb{1}_{A_k^\varepsilon} \varphi_k^\varepsilon|^2 \leq \varepsilon$. Thus $E|u|^2 \leq 2\varepsilon + 2E[\sum_{k=1}^{N_\varepsilon} \mathbb{1}_{A_k^\varepsilon} |\int_0^T (\varphi_k^\varepsilon)_t \mathrm{d}B_t|^2] \leq 2\varepsilon + 2E[m^2]$. Therefore, for every $u \in S_T(\int_0^T \Phi_t \mathrm{d}B_t)$, one has $E|u|^2 \leq 2E[m^2]$. Then $S(\int_0^T \Phi_t \mathrm{d}B_t)$ is a bounded subset of $\mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$, which implies that $\int_0^T \Phi_t \mathrm{d}B_t$ is square integrably bounded. □

Immediately from Theorem 1.3 and the definition of set-valued stochastic integrals, it follows that for every $\Phi \in \mathcal{L}_{\mathbb{F}}^2(\mathbb{R}^+ \times \Omega, \mathbb{R}^d)$, $\Psi \in \mathcal{L}_{\mathbb{F}}^2(\mathbb{R}^+ \times \Omega, \mathbb{R}^{d \times m})$, and m-dimensional $\mathbb{F}$-Brownian motion $B = (B_t)_{t \geq 0}$ on $\mathcal{P}_{\mathbb{F}}$, the set-valued mappings $[0, \infty) \ni t \to S_T(\int_0^t \Phi_\tau \mathrm{d}\tau) \in \mathrm{Cl}[\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{R}^d)]$ and $[0, \infty) \ni t \to S_T(\int_0^t \Psi_\tau \mathrm{d}B_\tau) \in \mathrm{Cl}[\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{R}^d)]$ are l.s.c. This, by Michael's continuous selection theorem, implies that if Φ and Ψ are convex-valued, then the above set-valued mappings possess $\mathbb{L}^2$-continuous selectors.

It is natural to ask whether it is possible to get, for set-valued stochastic integrals, an approximation selection theorem similar to Theorem 1.4. In the general case, the answer to such a question is negative, because in the proof of Theorem 1.4, the boundedness of the set-valued integral $\int_0^t \Psi_\tau \mathrm{d}B_\tau$ is essential. We can get only the following result.

Theorem 2.5. *Let* $B = (B_t)_{t \geq 0}$ *be an m-dimensional* $\mathbb{F}$*-Brownian motion on a filtered probability space* $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}_T, \mathbb{F}, P)$ *with a filtration* $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ *satisfying the usual conditions and let* $(\Phi, \Psi) \in \mathcal{L}_{\mathbb{F}}^2(T, \Omega, \mathbb{R}^n) \times \mathcal{L}_{\mathbb{F}}^2(T, \Omega, \mathbb{R}^{d \times m})$. *For every* $\varepsilon > 0$ *and a continuous d-dimensional stochastic process* $x = (x_t)_{t \geq 0}$ *such that*

$$x_t(\omega) - x_s(\omega) \in \mathrm{cl}\left\{\left(\int_s^t \Phi_t \mathrm{d}t\right)(\omega) + \left(\int_s^t \Psi_t \mathrm{d}B_t\right)(\omega)\right\} \tag{2.2}$$

for every $0 \le s \le t \le T$ *and a.e.* $\omega \in \Omega$*, there exist* $\delta_\varepsilon \in (0, \varepsilon]$*, a positive integer* $N_\varepsilon \in [T/\delta_\varepsilon, T/\delta_\varepsilon + 1)$*, and partitions* $(A_i^k)_{i=1}^{m_k} \subset \mathcal{F}_{k\delta_\varepsilon}$ *of* Ω *for* $k = 1, \ldots, N_\varepsilon$ *such that for every partition* $(B_i^k)_{i=1}^{m_k} \subset \mathcal{F}_{(k-1)\delta_\varepsilon}$ *of* Ω *with* $k = 1, \ldots, N_\varepsilon$*, there exists a pair of processes* $(\varphi^\varepsilon, \psi^\varepsilon) \in S_{\mathbb{F}}(\Phi) \times S_{\mathbb{F}}(\Psi)$ *such that*

$$\sup_{0\le t\le T} \left\| x_t - x_0 - \int_0^t \varphi_\tau^\varepsilon \mathrm{d}\tau - \int_0^t \psi_\tau^\varepsilon \mathrm{d}B_t \right\| \tag{2.3}$$
$$\le \varepsilon + \sup_{1\le j\le N_\varepsilon} \left\| \sum_{k=1}^{j} \sum_{i=1}^{m_k} (\mathbb{1}_{A_i^k} - \mathbb{1}_{B_i^k}) \left(\int_{(k-1)\delta_\varepsilon}^{k\delta_\varepsilon} \phi_\tau^\varepsilon \mathrm{d}\tau + \int_{(k-1)\delta_\varepsilon}^{k\delta_\varepsilon} \psi_\tau^\varepsilon \mathrm{d}B_\tau \right) \right\|.$$

Proof. By Theorem 2.1, the relation (2.2) is equivalent to the relation $x_t - x_s \in \mathrm{cl}_{\mathbb{L}}[\overline{dec}\{J_{st}[S_{\mathbb{F}}(\Phi)]\} + \overline{dec}\{\mathcal{J}_{st}[S_{\mathbb{F}}(\Psi)]\}]$ for every $0 \le s < t \le T$. Let $\varepsilon \in (0,1)$ and select $\delta_\varepsilon \in (0, \varepsilon]$ such that $\sup_{0\le s\le t\le T} \sup_{t\le s\le t+\delta_\varepsilon} E|x_t - x_s|^2 \le (\varepsilon/4)^2$, $\sup_{0\le t\le T} E\int_t^{t+\delta_\varepsilon} \|\Phi_\tau\|^2 \mathrm{d}\tau \le (1/4T)(\varepsilon/4)^2$ and $\sup_{0\le t\le T} E\int_t^{t+\delta_\varepsilon} \|\Psi_\tau\|^2 \mathrm{d}\tau \le (1/4)(\varepsilon/4)^4$. Put $\tau_0^\varepsilon = 0$, $\tau_k^\varepsilon = k\delta_\varepsilon$ for $k = 1, 2, \ldots, N_\varepsilon - 1$, where N_ε is such that $(N_\varepsilon - 1)\delta_\varepsilon < T \le N_\varepsilon\delta_\varepsilon$. For simplicity, assume that N_ε is such that $T = N_\varepsilon\delta_\varepsilon$. For every $k = 1, 2, \ldots, N_\varepsilon$, there are $f_k^\varepsilon \in \overline{dec}\{J_{\tau_{k-1}^\varepsilon \tau_k^\varepsilon}[S_{\mathbb{F}}(\Phi)]\}$ and $g_k^\varepsilon \in \overline{dec}\{\mathcal{J}_{\tau_{k-1}^\varepsilon \tau_k^\varepsilon}[S_{\mathbb{F}}(\Psi)]\}$ such that $\|x_{\tau_k^\varepsilon} - x_{\tau_{k-1}^\varepsilon} - f_k^\varepsilon - g_k^\varepsilon\| \le \varepsilon/4N_\varepsilon$, where $\|\cdot\|$ denotes the norm of $\mathbb{L}^2(\Omega, \mathcal{F}_{\tau_k^\varepsilon}, \mathbb{R}^d)$. By the definition of closed decomposable hulls of subsets of $\mathbb{L}^2(\Omega, \mathcal{F}_{\tau_k^\varepsilon}, \mathbb{R}^d)$, for every $k = 1, 2, \ldots, N_\varepsilon$ there are partitions $\{A_1^k, \ldots, A_{m_k}^k\} \subset \mathcal{F}_{\tau_k^\varepsilon}$ and $\{C_1^k, \ldots, C_{n_k}^k\} \subset \mathcal{F}_{\tau_k^\varepsilon}$ of Ω and families $\{\varphi_1^k, \ldots, \varphi_{m_k}^k\} \subset S_{\mathbb{F}}(\Phi)$ and $\{\psi_1^k, \ldots, \psi_{n_k}^k\} \subset S_{\mathbb{F}}(\Psi)$ such that $\|f_k - \sum_{i=1}^{m_k} \mathbb{1}_{A_i^k} \int_{\tau_{k-1}^\varepsilon}^{\tau_k^\varepsilon} (\varphi_i^k)_\tau \mathrm{d}\tau\| \le \varepsilon/8N_\varepsilon$ and $\|g_k - \sum_{i=1}^{n_k} \mathbb{1}_{C_i^k} \int_{\tau_{k-1}^\varepsilon}^{\tau_k^\varepsilon} (\psi_i^k)_\tau \mathrm{d}B_\tau\| \le \varepsilon/8N_\varepsilon$. For simplicity, we can assume that $m_k = n_k$ and $A_i^k = C_i^k$ for every $i = 1, 2, \ldots, m_k$ and $k = 1, 2, \ldots, N_\varepsilon$. Let $(B_i^k)_{i=1}^{m_k} \subset \mathcal{F}_{\tau_{k-1}^\varepsilon}$ be a partition of Ω for every $k = 1, 2, \ldots, N_\varepsilon$ and let φ^ε and ψ^ε be defined by

$$\varphi^\varepsilon = \sum_{k=1}^{N_\varepsilon - 1} \sum_{i=1}^{m_k} \mathbb{1}_{[\tau_{k-1}^\varepsilon, \tau_k^\varepsilon)} \mathbb{1}_{B_i^k} \varphi_i^k + \sum_{i=1}^{m_{N_\varepsilon}} \mathbb{1}_{[\tau_{N_\varepsilon - 1}^\varepsilon, T]} \mathbb{1}_{B_i^{N_\varepsilon}} \varphi_i^{N_\varepsilon}$$

and

$$\psi^\varepsilon = \sum_{k=1}^{N_\varepsilon - 1} \sum_{i=1}^{m_k} \mathbb{1}_{[\tau_{k-1}^\varepsilon, \tau_k^\varepsilon)} \mathbb{1}_{B_i^k} \psi_i^k + \sum_{i=1}^{m_{N_\varepsilon}} \mathbb{1}_{[\tau_{N_\varepsilon - 1}^\varepsilon, T]} \mathbb{1}_{B_i^{N_\varepsilon}} \psi_i^{N_\varepsilon}.$$

It is clear that φ^ε and ψ^ε are $\mathbb{F}$–nonanticipative selectors of Φ and Ψ, respectively, and

$$
\begin{aligned}
&\sup_{0\le t\le T}\left[E\left|x_t-x_0-\left(\int_0^t\varphi_\tau^\varepsilon \mathrm{d}\tau+\int_0^t\psi_\tau^\varepsilon \mathrm{d}B_\tau\right)\right|\right]^{\frac{1}{2}}\\
&\quad\le \sup_{1\le k\le N^\varepsilon}\sup_{\tau_{k-1}^\varepsilon\le t\le\tau_k^\varepsilon}\left[E|x_t-x_{\tau_{k-1}^\varepsilon}|^2\right]^{\frac{1}{2}}\\
&\qquad+\sup_{1\le k\le N_\varepsilon}\sup_{\tau_{k-1}^\varepsilon\le t\le\tau_k^\varepsilon}\left[E\left|\int_{\tau_{k-1}^\varepsilon}^t\varphi_\tau^\varepsilon \mathrm{d}\tau+\int_{\tau_{k-1}^\varepsilon}^t\psi_\tau^\varepsilon \mathrm{d}B_\tau\right|^2\right]^{\frac{1}{2}}+\sum_{k=1}^{N_\varepsilon}\left[|x_{\tau_k^\varepsilon}-x_{\tau_{k-1}^\varepsilon}-f_k^\varepsilon-g_k^\varepsilon|^2\right]^{\frac{1}{2}}\\
&\qquad+\sum_{k=1}^{N_\varepsilon}\left[E\left|f_k^\varepsilon-g_k^\varepsilon-\sum_{i=1}^{m_k}\mathbb{1}_{A_i^k}\int_{\tau_{k-1}^\varepsilon}^{\tau_k^\varepsilon}(\varphi_i^k)_\tau \mathrm{d}\tau-\sum_{i=1}^{m_k}\mathbb{1}_{A_i^k}\int_{\tau_{k-1}^\varepsilon}^{\tau_k^\varepsilon}(\psi_i^k)_\tau \mathrm{d}B_\tau\right|^2\right]^{\frac{1}{2}}\\
&\qquad+\sup_{1\le j\le N_\varepsilon}\left\|\sum_{k=1}^{j}\sum_{i=1}^{m_k}(\mathbb{1}_{A_i^k}-\mathbb{1}_{B_i^k})\left(\int_{\tau_{k-1}^\varepsilon}^{\tau_k^\varepsilon}(\varphi_i^k)_\tau \mathrm{d}\tau+\int_{\tau_{k-1}^\varepsilon}^{\tau_k^\varepsilon}(\psi_i^k)_\tau \mathrm{d}B_\tau\right)\right\|.
\end{aligned}
$$

By the definitions of τ_k^ε, f_k and g_k, we get $\sup_{1\le k\le N^\varepsilon}\sup_{\tau_{k-1}^\varepsilon\le t\le\tau_k^\varepsilon}E|x_t-x_{\tau_{k-1}^\varepsilon}|^2\le(\varepsilon/4)^2$, $\sum_{k=1}^{N_\varepsilon}\|x_{\tau_k^\varepsilon}-x_{\tau_{k-1}^\varepsilon}-f_k^\varepsilon-g_k^\varepsilon\|\le\sum_{k=1}^{N_\varepsilon}\varepsilon/4N_\varepsilon=\varepsilon/4$ and

$$
\sup_{1\le k\le N_\varepsilon}\sup_{\tau_{k-1}^\varepsilon\le t\le\tau_k^\varepsilon}E\left|\int_{\tau_{k-1}^\varepsilon}^t\varphi_\tau^\varepsilon \mathrm{d}\tau+\int_{\tau_{k-1}^\varepsilon}^t\psi_\tau^\varepsilon \mathrm{d}B_\tau\right|^2\le
$$

$$
\begin{aligned}
2T\sup_{1\le k\le N_\varepsilon}\sup_{\tau_{k-1}^\varepsilon\le t\le\tau_k^\varepsilon}E\left[\int_{\tau_{k-1}^\varepsilon}^t\|\Phi_\tau)\|^2\mathrm{d}\tau\right]+2\sup_{1\le k\le N_\varepsilon}\sup_{\tau_{k-1}^\varepsilon\le t\le\tau_k^\varepsilon}E\left[\int_{\tau_{k-1}^\varepsilon}^t\|\Psi_\tau)\|^2\mathrm{d}\tau\right]\\
\le\frac{1}{2}\left(\frac{\varepsilon}{4}\right)^2+\frac{1}{2}\left(\frac{\varepsilon}{4}\right)^4\le\left(\frac{\varepsilon}{4}\right)^2.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
&\sup_{0\le t\le T}\left\|x_t-x_0-\int_0^t\varphi_\tau^\varepsilon \mathrm{d}\tau-\int_0^t\psi_\tau^\varepsilon \mathrm{d}B_\tau\right\|\\
&\quad\le\varepsilon+\sup_{1\le j\le N_\varepsilon}\left\|\sum_{k=1}^{j}\sum_{i=1}^{m_k}(\mathbb{1}_{A_i^k}-\mathbb{1}_{B_i^k})\left(\int_{\tau_{k-1}^\varepsilon}^{\tau_k^\varepsilon}(\varphi_i^k)_\tau \mathrm{d}\tau+\int_{\tau_{k-1}^\varepsilon}^{\tau_k^\varepsilon}(\psi_i^k)_\tau \mathrm{d}B_\tau\right)\right\|.
\end{aligned}
$$

□

Remark 2.2. If $(\Phi,\Psi)\in\mathcal{L}_{\mathbb{F}}^4(T,\Omega,\mathbb{R}^n)\times\mathcal{L}_{\mathbb{F}}^4(T,\Omega,\mathbb{R}^{n\times d})$ then

$$\sup_{1\le j\le N_\varepsilon}\left\|\sum_{k=1}^{j}\sum_{i=1}^{m_k}(\mathbb{1}_{A_i^k}-\mathbb{1}_{B_i^k})\left(\int_{(k-1)\delta_\varepsilon}^{k\delta_\varepsilon}\phi_\tau^\varepsilon d\tau+\int_{(k-1)\delta_\varepsilon}^{k\delta_\varepsilon}\psi_\tau^\varepsilon dB_\tau\right)\right\|$$
$$\le\left(\sqrt{\delta_\varepsilon}+d\sqrt[4]{36\delta_\varepsilon}\right)\sum_{k=1}^{N_\varepsilon}\sum_{i=1}^{m_k}\sqrt[4]{P(A_i^k\triangle B_i^k)}\sqrt[4]{E\int_{(k-1)\delta_\varepsilon}^{k\delta_\varepsilon}max(\|\Phi_\tau\|^4,\|\Psi_\tau\|^4)d\tau}.$$

□

It is natural to look for some additional conditions by which it is possible to select for every $k=1,2,\dots,N_\varepsilon$, a partition $(B^k)_{i=1}^{m_k}\subset\mathcal{F}_{(k-1)\delta_\varepsilon}$ of Ω such that for appropriately selected multifunctions Φ and Ψ, one has

$$\lim_{\varepsilon\to 0}\sum_{k=1}^{N_\varepsilon}\sum_{i=1}^{m_k}\sqrt[4]{P(A_i^k\triangle B_i^k)}\sqrt[4]{E\int_{(k-1)\delta_\varepsilon}^{k\delta_\varepsilon}max(\|\Phi_\tau\|^4,\|\Psi_\tau\|^4)d\tau}<\infty$$

or

$$\lim_{\varepsilon\to 0}\sum_{k=1}^{N_\varepsilon}\sum_{i=1}^{m_k}\sqrt[4]{P(A_i^k\triangle B_i^k)}\sqrt[4]{E\int_{(k-1)\delta_\varepsilon}^{k\delta_\varepsilon}max(\|\Phi_\tau\|^4,\|\Psi_\tau\|^4)d\tau}=\infty$$

"very slowly." Such conditions may depend on some type continuity of the filtration $\mathbb{F}$ regarded as a multifunction defined on the interval $[0,T]$ with values in the space of all nonempty closed subsets of a metric space $(\mathcal{S}(P),\rho)$ associated to the measure space $(\Omega,\mathcal{F}_T,P)$. It is known (see [35], p. 169) that $(\mathcal{S}(P),\rho)$ is a complete metric space. Given a filtration $\mathbb{F}=(\mathcal{F}_t)_{0\le t\le T}$ on the probability space $(\Omega,\mathcal{F}_T,P)$, we can consider $\mathcal{F}_t$ for every $t\in[0,T]$ a closed subset of $\mathcal{S}(P)$, because the metric space $(\mathcal{S}_t(P),\rho)$ associated to the probability space $(\Omega,\mathcal{F}_t,P)$ is for every $t\ge 0$, a complete metric space, and $\mathcal{S}_t(P)\subset\mathcal{S}(P)$. Therefore, we can treat the filtration $\mathbb{F}$ as a multifunction $[0,T]\ni t\to\mathcal{F}_t\in \mathrm{Cl}(\mathcal{S}(P))$. Let h denote the Hausdorff metric defined on $\mathrm{Cl}(\mathcal{S}(P))$. We can now define the following types of continuity of the filtration $\mathbb{F}$.

The filtration $\mathbb{F}$ is said to be h-continuous if the multifunction $[0,T]\ni t\to\mathcal{F}_t\in \mathrm{Cl}(\mathcal{S}(P))$ is continuous with respect to the Hausdorff metric h on $\mathrm{Cl}(\mathcal{S}(P))$. It is called Hölder h-continuous with exponential $\alpha>0$ if there exists a number $\beta>0$ such that $h(\mathcal{F}_t,\mathcal{F}_s)\le\beta|t-s|^\alpha$ for every $t,s\in[0,T]$. It can be easily verified that every h-continuous filtration $\mathbb{F}$ of the probability space $(\Omega,\mathcal{F},P)$ is continuous.

Remark 2.3. If the assumptions of Theorem 2.5 are satisfied with $(\Phi,\Psi)\in\mathcal{L}_{\mathbb{F}}^4(T,\Omega,\mathbb{R}^d)\times\mathcal{L}_{\mathbb{F}}^4(T,\Omega,\mathbb{R}^{d\times m})$ and the filtration $\mathbb{F}$ is Hölder h-continuous with exponential $\alpha=3$, then for every $\varepsilon>0$, there exist $\delta_\varepsilon\in(0,\varepsilon)$, a positive integer m_ε and a pair of processes $(\varphi^\varepsilon,\psi^\varepsilon)\in S_{\mathbb{F}}(\Phi)\times S_{\mathbb{F}}(\Psi)$ such that

$$\sup_{0\le t\le T}\left\|x_t-x_0-\int_0^t\varphi_\tau^\varepsilon d\tau-\int_0^t\psi_\tau^\varepsilon dB_t\right\|\le\varepsilon\left\{1+m_\varepsilon\beta[3\sqrt{6}d(T+2\delta_\varepsilon)+T+\delta_\varepsilon^3\sqrt{\delta_\varepsilon}]\right\}.$$

In particular, if (Φ, Ψ) and the filtration $\mathbb{F}$ are such that

$$M_\varepsilon =: sup\left\{m_\varepsilon\beta\left[3\sqrt{6}d(T+2\delta_\varepsilon)+T+\delta_\varepsilon^3\sqrt{\delta_\varepsilon}\right] : \varepsilon\in(0,1)\right\}<\infty,$$

then immediately from (2.3) it follows

$$\sup_{0\le t\le T}\left\|x_t-x_0-\int_0^t\varphi_\tau^\varepsilon \mathrm{d}\tau-\int_0^t\psi_\tau^\varepsilon \mathrm{d}B_t\right\|\le\varepsilon(1+M_\varepsilon).$$

□

3 Conditional Expectation of Set-Valued Integrals Depending on Random Parameters

Given an Aumann integrable set-valued process $F\in\mathcal{M}(T,\Omega,\mathbb{R}^d)$, by $\int_0^T F_t(\cdot)\mathrm{d}t$ we denote the set-valued mapping $\Omega\ni\omega\to\int_0^T F_t(\omega)\mathrm{d}t\in\mathrm{Cl}(\mathbb{R}^d)$ defined by

$$\int_0^T F_t(\omega)\mathrm{d}t=\left\{\int_0^T f_t(\omega)\mathrm{d}t\,:\,(f_t)_{0\le t\le T}\in S(F)\right\}$$

for a.e. $\omega\in\Omega$. By the properties of Aumann integrals, it follows that if $F\in\mathcal{L}(T,\Omega,\mathbb{R}^d)$, then $\int_0^T F_t(\omega)dt$ is a compact convex subset of $\mathbb{R}^d$ for a.e. $\omega\in\Omega$. We shall consider properties of a conditional expectation of set-valued random variables of the form $\Omega\ni\omega\to\int_0^T F_t(\omega)\mathrm{d}t\in\mathrm{Cl}(\mathbb{R}^d)$. It will be shown that for every convex-valued process $F\in\mathcal{L}(T,\Omega,\mathbb{R}^d)$, one has $\int_0^T F_t(\omega)\mathrm{d}t=(\mathcal{A})(\int_0^T F_t dt)(\omega)$ for a.e. $\omega\in\Omega$. We begin with the following lemmas.

Lemma 3.1. *For every $F\in\mathcal{L}(T,\Omega,\mathbb{R}^d)$, the set-valued mapping $\int_0^T F_t(\cdot)\mathrm{d}t$ defined by $\Omega\ni\omega\to\int_0^T F_t(\omega)\mathrm{d}t\in\mathrm{Cl}(\mathbb{R}^d)$ is $\mathcal{F}_T$-measurable with compact convex values.*

Proof. Let us observe that the set-valued integral $\int_0^T F_t(\omega)\mathrm{d}t$ is an Aumann integral depending on the random parameter $\omega\in\Omega$. Therefore, by virtue of Theorem 3.4 of Chap. 2, for every $\omega\in\Omega$, it is a nonempty compact convex subset of $\mathbb{R}^d$, which by Remark 2.3 of Chap. 2, implies that for the $\mathcal{F}_T$-measurability of the set-valued mapping $\Omega\ni\omega\to\int_0^T F_t(\omega)\mathrm{d}t\in\mathrm{Cl}(\mathbb{R}^d)$, it is enough to verify that the function $\Omega\ni\omega\to\sigma(p,\int_0^T F_t(\omega)\mathrm{d}t)\in\mathbb{R}$ is $\mathcal{F}_T$-measurable for every $p\in\mathbb{R}^d$. By the measurability of F and its integrably boundedness, the function $[0,T]\times\Omega\ni(t,\omega)\to\sigma(p,\mathrm{co}\,F_t(\omega))\subset\mathbb{R}$ is measurable for every $p\in\mathbb{R}^d$. By virtue of Theorem 3.5 of Chap. 2, for every $p\in\mathbb{R}^d$, one has $\sigma(p,\int_0^T F_t(\omega)\mathrm{d}t))=\int_0^T\sigma(p,\mathrm{co}\,F_t(\omega))\mathrm{d}t$ for every $\omega\in\Omega$. Hence, by Fubini's theorem, the $\mathcal{F}_T$-measurability of the function $\Omega\ni\omega\to\sigma(p,\int_0^T F_t(\omega)\mathrm{d}t)\in\mathbb{R}$ follows for every $p\in\mathbb{R}^d$. Therefore, $\int_0^T F_t(\cdot)\mathrm{d}t$ is $\mathcal{F}_T$-measurable. □

Lemma 3.2. *Let $F \in \mathcal{L}(T,\Omega,\mathbb{R}^d)$. The subtrajectory integrals $S_T[\int_0^T F_t(\cdot)\mathrm{d}t]$ of $\int_0^T F_t(\cdot)\mathrm{d}t$ form a nonempty convex sequentially weakly compact subset of the space $\mathbb{L}(\Omega,\mathcal{F}_T,\mathbb{R}^d)$, and $S_T[\int_0^T F_t(\cdot)\mathrm{d}t] = J[S(\mathrm{co}\,F)]$.*

Proof. Immediately from Remark 3.1 of Chap. 2, by the definition of $J[S(\mathrm{co}\,F)]$ and the equality $\int_0^T F_t(\omega)\mathrm{d}t = \int_0^T \mathrm{co}\,F_t(\omega)\mathrm{d}t$ for a.e. $\omega \in \Omega$, it follows that $S[\int_0^T F_t(\cdot)\mathrm{d}t]$ is a nonempty convex sequentially weakly compact subset of the space $\mathbb{L}(\Omega,\mathcal{F}_T,\mathbb{R}^d)$, and $J[S(\mathrm{co}\,F)] \subset S_T[\int_0^T \mathrm{co}\,F_t(\cdot)\mathrm{d}t] = S_T[\int_0^T F_t(\cdot)\mathrm{d}t]$.

Assume now that $\varphi \in S_T[\int_0^T F(t,\cdot)\mathrm{d}t]$. Then for every $A \in \mathcal{F}_T$, one has $E_A\varphi \in E_A\Phi$, where $\Phi = \int_0^T F_t(\cdot)\mathrm{d}t$, $E_A\varphi = \int_A \varphi \mathrm{d}P$, and $E_A\Phi = \int_A \Phi\mathrm{d}P$. Let $\varepsilon > 0$ be given and select an $\mathcal{F}$-measurable partition $(A_n^\varepsilon)_{n=1}^{N_\varepsilon}$ of Ω such that $E_{A_n^\varepsilon}\int_0^T ||F_t(\cdot)||\mathrm{d}t < \varepsilon/2^{n+1}$. For every $n = 1,\dots,N_\varepsilon$, there is an $f_n^\varepsilon \in S(F)$ such that $E_{A_n^\varepsilon}\varphi = E_{A_n^\varepsilon}\int_0^T f_n^\varepsilon(t,\cdot)\mathrm{d}t$. Let $f^\varepsilon = \sum_{n=1}^{N_\varepsilon} \mathbb{1}_{A_n^\varepsilon} f_n^\varepsilon$. By the decomposability of $S(F)$, one has $f^\varepsilon \in S(F)$. We have $f^\varepsilon \in S(\mathrm{co}\,F)$ because $S(F) \subset S(\mathrm{co}\,F)$. Taking a sequence $(\varepsilon_k)_{k=1}^\infty$ of positive numbers $\varepsilon_k > 0$ such that $\varepsilon_k \to 0$ as $k \to \infty$, we can select $f \in S(\mathrm{co}\,F)$ and a subsequence, denoted again by $(f^{\varepsilon_k})_{k=1}^\infty$, of $(f^{\varepsilon_k})_{k=1}^\infty$ weakly converging to f in the weak topology of $\mathbb{L}([0,T]\times\Omega,\beta_T\otimes\mathcal{F}_T,\mathbb{R}^d)$, because $S(\mathrm{co}\,F)$ is a sequentially weakly compact subset of $\mathbb{L}([0,T]\times\Omega,\beta_T\otimes\mathcal{F}_T,\mathbb{R}^d)$. For every $A \in \mathcal{F}$ and $k = 1,2,\dots$, there is a subset $\{n_1,\dots,n_p\}$ of $\{1,\dots,N_{\varepsilon_k}\}$ such that $A\cap A_{n_i}^{\varepsilon_k} \neq \emptyset$ for $i = 1,2,\dots,p$ and $A \cap A_r = \emptyset$ for $r \in \{1,2,\dots,N_{\varepsilon_k}\}\setminus\{n_1,\dots,n_p\}$. Therefore,

$$\left|E_A\varphi - E_A\int_0^T f^{\varepsilon_k}(t,\cdot)\mathrm{d}t\right|$$
$$\leq \sum_{n=1}^{N_{\varepsilon_k}}\left|E_{A\cap A_n^{\varepsilon_k}}\varphi - E_{A\cap A_n^{\varepsilon_k}}\int_0^T f_n^{\varepsilon_k}(t,\cdot)\mathrm{d}t\right|$$
$$= \sum_{i=1}^{p}\left|E_{A\cap A_{n_i}^{\varepsilon_k}}\varphi - E_{A\cap A_{n_i}^{\varepsilon_k}}\int_0^T f_n^{\varepsilon_k}(t,\cdot)\mathrm{d}t\right| \leq 2\sum_{i=1}^{p} E_{A_{n_i}^{\varepsilon_k}}\int_0^T ||F_t(\cdot)||\mathrm{d}t \leq \varepsilon_k$$

for every $k = 1,2,\dots$. On the other hand, for every $A \in \mathcal{F}$, we also have

$$\left|E_A\varphi - E_A\int_0^T f(t,\cdot)\mathrm{d}t\right| \leq \left|E_A\varphi - E_A\int_0^T f^{\varepsilon_k}(t,\cdot)\mathrm{d}t\right|$$
$$+\left|E_A\int_0^T f^{\varepsilon_k}(t,\cdot)\mathrm{d}t - E_A\int_0^T f(t,\cdot)\mathrm{d}t\right| \leq \varepsilon_k + \left|E_A\int_0^T f^{\varepsilon_k}(t,\cdot)\mathrm{d}t - E_A\int_0^T f(t,\cdot)\mathrm{d}t\right|$$

for $k = 1,2,\dots$. Hence it follows that $E_A\varphi = E_A\int_0^T f(t,\cdot)\mathrm{d}t$ for every $A \in \mathcal{F}$, because $\varepsilon_k \to 0$ and $|E_A\int_0^T f^{\varepsilon_k}(t,\cdot)\mathrm{d}t - E_A\int_0^T f(t,\cdot)\mathrm{d}t| \to 0$ as $k \to \infty$. Therefore, $\varphi(\omega) = \int_0^T f(t,\omega)\mathrm{d}t$ for a.e. $\omega \in \Omega$, i.e., $\varphi = J(f)$. Then $\varphi \in J[S(\mathrm{co}\,F)]$. □

Corollary 3.1. *For every convex-valued process* $F \in \mathcal{L}(T, \Omega, \mathbb{R}^d)$, *one has* $(\mathcal{A})(\int_0^T F_t \mathrm{d}t)(\omega) = \int_0^T F_t(\omega)\mathrm{d}t$ *for a.e.* $\omega \in \Omega$.

Proof. By the definition of $(\mathcal{A})\int_0^T F_t \mathrm{d}t$ and Corollary 1.1, we have $S_T[(\mathcal{A})\int_0^T F_t \mathrm{d}t] = \mathrm{cl}_{\mathbb{L}}\{J[(S(F))]\} = J[(S(F))]$, because $J[(S(F))]$ is a convex sequentially weakly compact subset of $\mathbb{L}(\Omega, \mathcal{F}_T, \mathbb{R}^d)$. By virtue of Lemma 3.2, one has $S_T[\int_0^T F_t(\cdot)\mathrm{d}t] = J[S(F)]$. Therefore, $S_T[(\mathcal{A})\int_0^T F_t \mathrm{d}t] = S_T[\int_0^T F_t(\cdot)\mathrm{d}t]$, which by Corollary 3.1 of Chap. 2, implies $(\mathcal{A})(\int_0^T F_t \mathrm{d}t)(\omega) = \int_0^T F_t(\omega)\mathrm{d}t$ for a.e. $\omega \in \Omega$. □

Given a measurable and uniformly integrably bounded set-valued mapping $F : [0, T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and a d-dimensional measurable stochastic process $z = (z_t)_{0 \le t \le T}$ on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$, by $S(F \circ z)$ we shall denote the set of all measurable selectors of the set-valued process $F \circ z$ defined by $(F \circ z)_t(\omega) = F(t, z_t(\omega))$ for $(t, \omega) \in [0, T] \times \Omega$. Recall that F is said to be uniformly integrably bounded if there exists $k \in \mathbb{L}([0, T], \mathbb{R}^+)$ such that $\|F(t, x)\| \le k(t)$ for a.e. $t \in [0, T]$ and every $x \in \mathbb{R}^d$.

Lemma 3.3. *Assume that* $F : [0, T] \times \mathbb{R}^d \to Cl(\mathbb{R}^d)$ *is measurable and uniformly integrably bounded,* $(z_t)_{0 \le t \le T}$ *is a* d*-dimensional measurable stochastic process on* $\mathcal{P}_{\mathbb{F}}$, *and* $\mathcal{G}$ *is a sub-*σ*-algebra of* $\mathcal{F}$. *Then*

$$S\left(E\left[\int_0^T F(t, z_t)\mathrm{d}t|\mathcal{G}\right]\right) = \left\{E\left[\int_0^T f(t, \cdot)\mathrm{d}t|\mathcal{G}\right] : f \in S(\mathrm{co}\, F \circ z)\right\}.$$

Proof. By virtue of Lemma 3.2 and the definition of a set-valued conditional expectation, one gets

$$S\left(E\left[\int_0^T F(t, z_t)\mathrm{d}t|\mathcal{G}\right]\right) = \mathrm{cl}_{\mathbb{L}}\left\{E[u|\mathcal{G}] : u \in J[S(\mathrm{co}\, F \circ z)]\right\}$$
$$= \mathrm{cl}_{\mathbb{L}}\left\{E\left[\int_0^T f(t, \cdot)\mathrm{d}t|\mathcal{G}\right] : f \in S(\mathrm{co}\, F \circ z)\right\}.$$

To complete the proof, it remains only to verify that $\mathcal{H} = \{E[\int_0^T f(t, \cdot)\mathrm{d}t|\mathcal{G}] : f \in S(\mathrm{co} F \circ z)\}$ is a closed subset of $\mathbb{L}(\Omega, \mathcal{G}, \mathbb{R}^d)$. This follows immediately from the sequential weak compactness and convexity of $S(\mathrm{co}\, F \circ x)$. Indeed, by the sequential weak compactness and convexity of $S(\mathrm{co}\, F \circ z)$, it follows that $\mathcal{H}$ is a convex sequentially weakly compact subset of $\mathbb{L}(\Omega, \mathcal{G}, \mathbb{R}^d)$. Therefore, it is a closed subset of $\mathbb{L}(\Omega, \mathcal{G}, \mathbb{R}^d)$. □

Theorem 3.1. *Assume that* $F : [0, T] \times \mathbb{R}^d \to Cl(\mathbb{R}^d)$ *is measurable and uniformly integrably bounded and let* $x = (x_t)_{0 \le t \le T}$ *and* $z = (z_t)_{0 \le t \le T}$ *be* d*-dimensional measurable stochastic processes on a filtered probability space* $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ *with a filtration* $\mathbb{F} = (\mathcal{F}_t)_{t \ge 0}$ *satisfying the usual conditions, and let* $E|x_T| < \infty$. *Then*

$$x_s \in E\left[x_t + \int_s^t F(\tau, z_\tau)\mathrm{d}\tau|\mathcal{F}_s\right] \quad a.s. \tag{3.1}$$

for every $0 \le s \le t \le T$ *if and only if there is* $f \in S(\text{co}F \circ z)$ *such that*

$$x_t = E\left[x_T + \int_t^T f(\tau,\cdot)\mathrm{d}\tau|\mathcal{F}_t\right] \quad a.s. \tag{3.2}$$

for every $0 \le t \le T$.

Proof. Suppose there is $f \in S(\text{co}F \circ z)$ such that (3.2) is satisfied. For every $0 \le s \le t \le T$, one has

$$x_s = E\left[x_T + \int_s^T f(\tau,\cdot)\mathrm{d}\tau|\mathcal{F}_s\right] = E\left[\int_s^t f(\tau,\cdot)\mathrm{d}\tau|\mathcal{F}_s\right] + E\left[x_T + \int_t^T f(\tau,\cdot)\mathrm{d}\tau|\mathcal{F}_s\right]$$

and $E[x_t|\mathcal{F}_s] = E\left[x_T + \int_t^T f(\tau,\cdot)\mathrm{d}\tau|\mathcal{F}_s\right]$ a.s. Then $x_s = E\left[x_t + \int_s^t f(\tau,\cdot)\mathrm{d}\tau|\mathcal{F}_s\right]$ a.s. for $0 \le s \le t \le T$. Hence, by Lemma 3.3, it follows that $x_s \in S\left(E\left[x_t + \int_s^t F(\tau, z_\tau)\mathrm{d}\tau|\mathcal{F}_s\right]\right)$ for $0 \le s \le t \le T$. Therefore, (3.1) is satisfied a.s. for $0 \le s \le t \le T$.

Assume that (3.1) is satisfied for every $0 \le s \le t \le T$ a.s. and let $k \in \mathbb{L}([0,T],\mathbb{R}_+)$ be such that $\|F(t,x\| \le k(t)$ for a.e. $t \in [0,T]$ and $x \in \mathbb{R}^d$. For every $0 \le t \le T$, one has $E|x_t| \le E|x_T| + E\int_0^T k(t)\mathrm{d}t < \infty$. Let $\eta > 0$ be fixed and select $\delta > 0$ such that $\delta < T$ and $\sup_{0\le t\le T-\delta}\int_t^{t+\delta} k(\tau)\mathrm{d}\tau < \eta/2$. For fixed $t \in [0, T-\delta]$ and $t \le \tau \le t+\delta$, we have $x_t \in E[x_\tau + \int_t^\tau F(s,z_s)\mathrm{d}s|\mathcal{F}_t]$ a.s. Therefore, for every $A \in \mathcal{F}_t$, we get $E_A(x_t - x_\tau) \in E_A\int_t^\tau F(s,z_s)\mathrm{d}s$. Then

$$|E_A(x_t - x_\tau)| \le E_A\int_t^\tau \|F(s,z_s)\|\mathrm{d}s \le E\int_t^{t+\delta} k(s)\mathrm{d}s < \eta/2$$

for every $0 \le t \le T-\delta$ and $A \in \mathcal{F}_t$. Therefore, $\sup_{t\le\tau\le t+\delta}|E_A(x_t - x_\tau)| \le \eta/2$ for every $A \in \mathcal{F}_t$ and fixed $0 \le t \le T-\delta$.

Let $\tau_0 = 0, \tau_1 = \delta, \ldots, \tau_{N-1} = (N-1)\delta < T \le N\delta$. Immediately from (3.1) and Lemma 3.3, it follows that for every $i = 1,2,\ldots,N-1$, there is $f_i^\eta \in S(\text{co}F \circ z)$ such that

$$E\left|x_{\tau_{i-1}} - E\left[x_{\tau_i} + \int_{\tau_{i-1}}^{\tau_i} f_i^\eta(s,\cdot)\mathrm{d}s|\mathcal{F}_{\tau_{i-1}}\right]\right| = 0.$$

Furthermore, there is $f_N^\eta \in S_{\mathbb{F}}(\text{co}F)$ such that

$$E\left|x_{\tau_{N-1}} - E\left[x_T + \int_{\tau_{N-1}}^{T} f_N^\eta(s,\cdot)\mathrm{d}s|\mathcal{F}_{\tau_{N-1}}\right]\right| = 0.$$

Define $f^\eta(t,\omega) = \sum_{i=1}^{N-1} 1\mathbb{1}_{[\tau_{i-1},\tau_i)}(t) f_i^\eta(t,\omega) + \mathbb{1}_{[\tau_{N-1},T]} f_N^\eta(t,\omega)$ for $(t,\omega) \in [0,T]\times\Omega$. By the decomposability of $S(\text{co}F \circ z)$, we have $f^\eta \in S(\text{co}F \circ z)$. For fixed $t \in [0,T]$, there is $p \in \{1,2,\dots,N-1\}$ or $p=N$ such that $t \in [\tau_{p-1},\tau_p)$ or $t \in [\tau_{N-1},T]$. Let $t \in [\tau_{p-1},\tau_p)$ with $1 \le p \le N-1$. For every $A \in \mathcal{F}_t$, one has

$$\begin{aligned}
&\left|E_A\left(x_t - E\left[x_T + \int_t^T f^\eta(s,\cdot)\mathrm{d}s\mathcal{F}_t\right]\right)\right| \le |E_A(x_t - x_{\tau_p})| + E\left|x_{\tau_p} - E\left[x_{\tau_{p+1}}\right.\right.\\
&\quad \left.\left. + \int_{\tau_p}^{\tau_{p+1}} f^\eta(s,\cdot)\mathrm{d}\tau|\mathcal{F}_{\tau_p}\right]\right| + \ |E_A(E[x_{\tau_{p+1}}|\mathcal{F}_{\tau_p}] - x_{\tau_{p+1}})| + E\left|\int_t^{\tau_p} f^\eta(s,\cdot)\mathrm{d}s\right|\\
&\quad + \left|E_A\left(E\left[\int_{\tau_p}^{\tau_{p+1}} f^\eta(s,\cdot)\mathrm{d}s|\mathcal{F}_{\tau_p}\right] - E\left[\int_{\tau_p}^{\tau_{p+1}} f^\eta(s,\cdot)\mathrm{d}\tau|\mathcal{F}_t\right]\right)\right| + \dots\\
&\quad + E\left|x_{\tau_{N-1}} - E\left[x_T + \int_{\tau_{N-1}}^T f^\eta(s,\cdot)\mathrm{d}\tau|\mathcal{F}_{\tau_{N-1}}\right]\right| + \left|E_A(E[x_{\tau_{N-1}}|\mathcal{F}_{\tau_{N-1}}] - x_{\tau_{N-1}})\right|\\
&\quad + E_A\left(E\left[\int_{\tau_{N-1}}^T f^\eta(s,\cdot)\mathrm{d}s|\mathcal{F}_{\tau_{N-1}}\right] - E\left[\int_{\tau_{N-1}}^T f^\eta(s,\cdot)\mathrm{d}s|\mathcal{F}_t\right]\right)\Bigg|\\
&\le \sup_{t\le\tau\le t+\delta} |E_A(x_t - x_\tau)| + E\int_t^{t+\delta} m(s,\cdot)\mathrm{d}s + \sum_{i=p}^{N-2} E\left|x_{\tau_i} - E\left[x_{\tau_{i+1}} + \int_{\tau_i}^{\tau_{i+1}} f^\eta(s,\cdot)\mathrm{d}s|\mathcal{F}_{\tau_i}\right]\right|\\
&\quad + E\left|x_{\tau_{N-1}} - E\left[x_T + \int_{\tau_{N-1}}^T f^\eta(s,\cdot)\mathrm{d}\tau|\mathcal{F}_{\tau_{N-1}}\right]\right| + \sum_{i=p}^{N-2} |E_A(E[x_{\tau_{i+1}}|\mathcal{F}_{\tau_i}] - x_{\tau_{i+1}})|\\
&\quad + \sum_{i=p}^{N-2}\left|E_A\left(E\left[\int_{\tau_i}^{\tau_{i+1}} f^\eta(s,\cdot)\mathrm{d}s|\mathcal{F}_{\tau_i}\right] - E\left[\int_{\tau_i}^{\tau_{i+1}} f^\eta(s,\cdot)\mathrm{d}s|\mathcal{F}_t\right]\right)\right|\\
&\quad + \left|E_A\left(E\left[\int_{\tau_{N-1}}^T f^\eta(s,\cdot)\mathrm{d}s|\mathcal{F}_{\tau_{N-1}}\right] - E\left[\int_{\tau_{N-1}}^T f^\eta(s,\cdot)\mathrm{d}s|\mathcal{F}_t\right]\right)\right|.
\end{aligned}$$

But $\mathcal{F}_t \subset \mathcal{F}_{\tau_i}$ for $i = p, p+1, \dots, N-1$. Then for $A \in \mathcal{F}_t$, one has

$$\sum_{i=p}^{N-2} |E_A(E[x_{\tau_{i+1}}|\mathcal{F}_{\tau_i}] - x_{\tau_{i+1}})| = 0,$$

$$\sum_{i=p}^{N-2}\left|E_A\left(E\left[\int_{\tau_i}^{\tau_{i+1}} f^\eta(s,\cdot)\mathrm{d}s|\mathcal{F}_{\tau_i}\right] - E\left[\int_{\tau_i}^{\tau_{i+1}} f^\eta(s,\cdot)\mathrm{d}s|\mathcal{F}_t\right]\right)\right| = 0$$

and

$$\left|E_A\left(E\left[\int_{\tau_{N-1}}^T f^\eta(s,\cdot)\mathrm{d}s|\mathcal{F}_{\tau_{N-1}}\right] - E\left[\int_{\tau_{N-1}}^T f^\eta(s,\cdot)\mathrm{d}s|\mathcal{F}_t\right]\right)\right| = 0.$$

With this and the definition of f^η, it follows that

$$\left|E_A\left(x_t - E\left[x_T + \int_t^T f^\eta(s,\cdot)\mathrm{d}s|\mathcal{F}_t\right]\right)\right| \le \eta \tag{3.3}$$

for fixed $0 \le t \le T$ and $A \in \mathcal{F}_t$. Let $(\eta_j)_{j=1}^{\infty}$ be a sequence of positive numbers converging to zero. For every $j = 1, 2, \ldots$, we can select $f^{\eta_j} \in S(\text{co}F \circ z)$ such that (3.3) is satisfied with $\eta = \eta_j$. By the weak sequential compactness of $S(\text{co}F \circ z)$, there are $f \in S(\text{co}F \circ z)$ and a subsequence $(f^{\eta_k})_{k=1}^{\infty}$ of $(f^{\eta_j})_{j=1}^{\infty}$ weakly converging to f in the weak topology of $\mathbb{L}([0,T] \times \Omega, \beta_T \otimes \mathcal{F}, \mathbb{R}^d)$. Then for every $A \in \mathcal{F}_t \subset \mathcal{F}$, one has

$$\lim_{k\to\infty} E_A \int_t^T f^{\eta_k}(s,\cdot)\mathrm{d}s = E_A \int_t^T f(s,\cdot)\mathrm{d}s.$$

On the other hand, for every fixed $t \in [0,T]$ and $A \in \mathcal{F}_t$, we have

$$\begin{aligned}\left|E_A\left(x_t - E\left[x_T + \int_t^T f(s,\cdot)\mathrm{d}s|\mathcal{F}_t\right]\right)\right| &\le \left|E_A\left(x_t - E\left[x_T + \int_t^T f^{\eta_k}(s,\cdot)\mathrm{d}s|\mathcal{F}_t\right]\right)\right| \\ &\quad + \left|E_A\left(E\left[\int_t^T f^{\eta_k}(s,\cdot)\mathrm{d}s|\mathcal{F}_t\right] - E\left[\int_t^T f(s,\cdot)\mathrm{d}s|\mathcal{F}_t\right]\right)\right| \\ &\le \eta_k + \left|E_A\int_t^T f^{\eta_k}(s,\cdot)\mathrm{d}s - E_A\int_t^T f(s,\cdot)\mathrm{d}s\right|\end{aligned}$$

for $k = 1, 2, \ldots$, Therefore, $E_A(x_t - E[x_T + \int_t^T f(s,\cdot)\mathrm{d}s|\mathcal{F}_t]) = 0$ for every $A \in \mathcal{F}_t$ and fixed $0 \le t \le T$. But x_t and $E[x_T + \int_t^T f(s,\cdot)\mathrm{d}s|\mathcal{F}_t]$ are $\mathcal{F}_t$-measurable. Then $x_t = E[x_T + \int_t^T f(s,\cdot)\mathrm{d}s|\mathcal{F}_t]$ a.s. for $0 \le t \le T$. Then there exists $f \in S(\text{co}F \circ z)$ such that (3.2) is satisfied. □

In what follows, by $\mathcal{S}(\mathbb{F}, \mathbb{R}^d)$ we shall denote the set of all d-dimensional $\mathbb{F}$-semimartingales $x = (x_t)_{0\le t\le T}$ on the filtered probability space $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ such that $E[\sup_{0\le t\le T}|x_t|^2] < \infty$. For a measurable process Z on $\mathcal{P}_{\mathbb{F}}$, by $[Z]^{\mathbb{F}}$ we shall denote a "conditional expectation" with respect to a measure $\mu \otimes P$ and an $\mathbb{F}$-optional σ-algebra $\mathcal{O}$, i.e., $[Z]^{\mathbb{F}} = E_{\mu\otimes P}[Z|\mathcal{O}]$, where μ denotes the Lebesgue measure on $[0,T]$.

Corollary 3.2. *If the assumptions of Theorem 3.1 are satisfied, then a process* $x = (x_t)_{0\le t\le T}$ *defined by* $x_t = E\left[x_T + \int_t^T f(\tau,\cdot)\mathrm{d}\tau|\mathcal{F}_t\right]$ *a.s. for* $0 \le t \le T$ *with* $f \in S(\text{co}\,F \circ z)$ *belongs to* $\mathcal{S}(\mathbb{F}, \mathbb{R}^d)$ *and has a semimartingale representation* $x_t = x_0 + M_t + A_t$, *where* $x_0 = E[\xi + \int_0^T f_\tau \mathrm{d}\tau|\mathcal{F}_0]$, $A_t = -\int_0^t [f]_\tau^{\mathbb{F}}\mathrm{d}\tau$ *and* $M_t = E[x_T + \int_0^T f_\tau\mathrm{d}\tau|\mathcal{F}_t] - E[x_T + \int_0^T f_\tau\mathrm{d}\tau|\mathcal{F}_0] - E[\int_0^t (f_\tau - [f]_\tau^{\mathbb{F}})\mathrm{d}\tau|\mathcal{F}_t]$. *A process* x *is continuous if and only if* $(M_t)_{0\le t\le T}$ *is a continuous martingale.*

Proof. It is clear that $x_t = x_0 + M_t + A_t$ a.s. for $0 \le t \le T$, where x_0, M_t, and A_t are as above for every $0 \le t \le T$. To see that $(A_t)_{0\le t\le T}$ is an $\mathbb{F}$-adapted absolutely continuous process and $(M_t)_{0\le t\le T}$ is an $\mathbb{F}$-martingale, let us observe first that $[f]_t^{\mathbb{F}}$ is $\mathcal{F}_t$-measurable for every $f \in S(\text{co}\,F \circ z)$ and $t \in [0,T]$, which implies that also A_t is $\mathcal{F}_t$-measurable for every $f \in S(\text{co}\,F \circ z)$ and $t \in [0,T]$. Furthermore, the process $(A_t)_{0\le t\le T}$ is absolutely continuous, because $|[f]_t^{\mathbb{F}}| \le |f|_t^{\mathbb{F}} \le \|F(t, z_t)\|$

a.s. for a.e. $t \in [0, T]$. To verify that $(M_t)_{0\le t\le T}$ is an $\mathbb{F}$-martingale, let us observe that $E[\int_s^t f_\tau \mathrm{d}\tau|\mathcal{F}_t] = \int_s^t E[f_\tau|\mathcal{F}_t]\mathrm{d}\tau$ a.s. for every $0 \le s \le t \le T$. Indeed, for every $C \in \mathcal{F}_t$ and $0 \le s \le t \le T$, one has

$$\int_C \left\{E\left[\int_s^t f_\tau \mathrm{d}\tau|\mathcal{F}_t\right]\right\} \mathrm{d}P$$
$$= \int_C \left\{\int_s^t f_\tau \mathrm{d}\tau\right\} \mathrm{d}P = \int_s^t \left\{\int_C f_\tau \mathrm{d}P\right\} \mathrm{d}\tau = \int_s^t \left\{\int_C E[f_\tau|\mathcal{F}_t]\,\mathrm{d}P\right\} \mathrm{d}\tau = \int_C \left\{\int_s^t E[f_\tau|\mathcal{F}_t]\,\mathrm{d}\tau\right\} \mathrm{d}P.$$

Then $E[\int_s^t f_\tau \mathrm{d}\tau|\mathcal{F}_t] = \int_s^t E[f_\tau|\mathcal{F}_t]\mathrm{d}\tau$ a.s. for every $0 \le s \le t \le T$. Let $N_t = E[\int_0^t (f_\tau - [f]_\tau^{\mathbb{F}})\mathrm{d}\tau|\mathcal{F}_t]$ a.s. for $0 \le s \le t \le T$. It is clear that $(M_t)_{0\le t\le T}$ is an $\mathbb{F}$-martingale if and only if the process $(N_t)_{0\le t\le T}$ is an $\mathbb{F}$-martingale. We have $E|N_t| < \infty$ for every $0 \le t \le T$. Furthermore, for every $0 \le s < t \le T$, one has

$$\int_C \left\{E\left[\int_s^t f_\tau \mathrm{d}\tau|\mathcal{F}_t\right]\right\} \mathrm{d}P$$
$$= \int_C \left\{\int_s^t f_\tau \mathrm{d}\tau\right\} \mathrm{d}P = \int_s^t \left\{\int_C f_\tau \mathrm{d}P\right\} \mathrm{d}\tau = \int_s^t \left\{\int_C E[f_\tau|\mathcal{F}_t]\,\mathrm{d}P\right\} \mathrm{d}\tau$$
$$= \int_C \left\{\int_s^t E[f_\tau|\mathcal{F}_t]\,\mathrm{d}\tau\right\} \mathrm{d}P.E[N_t - N_s|\mathcal{F}_s]$$
$$= E\left[\left(E\left[\int_0^t (f_\tau - [f]_\tau^{\mathbb{F}})\mathrm{d}\tau|\mathcal{F}_t\right] - E\left[\int_0^s (f_\tau - [f]_\tau^{\mathbb{F}})\mathrm{d}\tau|\mathcal{F}_s\right]\right)\Big|\mathcal{F}_s\right]$$
$$= E\left[\int_0^t E[(f_\tau - [f]_\tau^{\mathbb{F}})|\mathcal{F}_t]\Big|\mathcal{F}_s\right] - E\left[\int_0^s E[(f_\tau - [f]_\tau^{\mathbb{F}})|\mathcal{F}_s]\Big|\mathcal{F}_s\right]$$
$$= \int_0^t E[(f_\tau - [f]_\tau^{\mathbb{F}})|\mathcal{F}_s]\mathrm{d}\tau - \int_0^s E[(f_\tau - [f]_\tau^{\mathbb{F}})|\mathcal{F}_s]\mathrm{d}\tau = \int_s^t E[(f_\tau - [f]_\tau^{\mathbb{F}})|\mathcal{F}_s]\mathrm{d}\tau.$$

But for every $C \in \mathcal{F}_s$, one has $(s, t] \times C \in \mathcal{O}$. Therefore, for every $C \in \mathcal{F}_s$, one gets

$$\int_C \left[\int_s^t E[(f_\tau - [f]_\tau^{\mathbb{F}})|\mathcal{F}_s]\mathrm{d}\tau\right] \mathrm{d}P = \int\int_{(s,t]\times C} f_\tau \mathrm{d}\tau \mathrm{d}P - \int\int_{(s,t]\times C} [f]_\tau^{\mathbb{F}} \mathrm{d}\tau \mathrm{d}P$$
$$= \int\int_{(s,t]\times C} f_\tau \mathrm{d}\tau \mathrm{d}P - \int\int_{(s,t]\times C} f_\tau \mathrm{d}\tau \mathrm{d}P = 0.$$

Hence it follows that $\int_s^t E[(f_\tau - [f]_\tau^{\mathbb{F}})|\mathcal{F}_s]\mathrm{d}\tau = 0$ a.s. for every $0 \le s < t \le T$, which implies that $E[N_t - N_s|\mathcal{F}_s] = 0$ a.s. for every $0 \le s < t \le T$. Finally, by the equality $x_t = x_0 + M_t + A_t$ and the continuity of the process $(A_t)_{0\le t\le T}$, it follows that the process x is continuous if and only if $(M_t)_{0\le t\le T}$ is a continuous martingale. □

4 Selection Properties of Set-Valued Integrals Depending on Random Parameters

Let $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ be a complete filtered probability space with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ satisfying the usual conditions and let $G = (G_t)_{0\leq t}$ be a measurable set-valued stochastic process on $\mathcal{P}_{\mathbb{F}}$ with values in $\mathbb{R}^{r\times m}$. It is said to be diagonally convex if for every $t \in [0, T]$ and a.e. $\omega \in \Omega$, the set $D(G)_t(\omega) = \{v \cdot v^* : v \in G_t(\omega)\}$ is a convex subset of $\mathbb{R}^{r\times r}$, where v^* denotes the transpose of v. In what follows, the set-valued process $(D(G)_t)_{0\leq t\leq T}$ will be denoted by $D(G)$.

Remark 4.1. It can be proved that if the process $G = (G_t)_{t\in T}$ on $\mathcal{P}_{\mathbb{F}}$ with values in $\mathbb{R}^{1\times m}$ takes convex values, than it is diagonally convex. In the general case, convexity and diagonal convexity are not related to each other. □

Corollary 4.1. *If $G = (G_t)_{0\leq t\leq T}$ is a measurable square integrably bounded set-valued stochastic process on $\mathcal{P}_{\mathbb{F}}$ with closed values in $\mathbb{R}^{r\times m}$, then $D(G)$ is measurable and integrably bounded.*

Proof. It is enough to observe that $D(G) = l(G)$ and that a mapping $l : \mathbb{R}^{r\times m} \to \mathbb{R}^{r\times r}$ defined by $l(u) = u\cdot u^*$ for $u \in \mathbb{R}^{r\times m}$ is continuous and that $|l(u)| \leq \|u\|^2$ for every $u \in \mathbb{R}^{r\times m}$. □

Lemma 4.1. *If $G = (G_t)_{0\leq t\leq T}$ is a measurable square integrably bounded set-valued stochastic process on $\mathcal{P}_{\mathbb{F}}$ with closed values in $\mathbb{R}^{r\times m}$, then $S(D(G)) \neq \emptyset$, and for every $\sigma \in S(D(G))$, there exists $g \in S(G)$ such that $\sigma = g \cdot g^*$.*

Proof. The result follows immediately from Corollary 4.1 and Theorem 2.5 of Chap. 2 applied to the function $l(u) = u \cdot u^*$ and the set-valued mapping $\Gamma : [0, T] \times \Omega \to \mathrm{Cl}(\mathbb{R}^{r\times m})$ defined by $\Gamma(t, \omega) = G_t(\omega)$ for $(t, \omega) \in [0, T] \times \Omega$. □

Denote by $\mathcal{C}_1 = C_b(\mathbb{R}^r, \mathbb{R})$, $\mathcal{C}_r = C_b(\mathbb{R}^r, \mathbb{R}^r)$, and $\mathcal{C}_{r\times r} = C_b(\mathbb{R}^r, \mathbb{R}^{r\times r})$ Banach spaces of continuous bounded functions defined on $\mathbb{R}^r$ with values in $\mathbb{R}$, $\mathbb{R}^r$, and $\mathbb{R}^{r\times r}$, respectively. Define on $\mathcal{C}_r\times\mathbb{R}^r\times\mathbb{R}^r$ and $\mathcal{C}_{r\times r}\times\mathbb{R}^{r\times r}\times\mathbb{R}^r$ mappings Φ and Ψ by setting $\Phi(\varphi, u)(z) = \sum_{i=1}^r \varphi_i(z)u_i$ and $\Psi(\psi, v)(z) = \sum_{i=1}^r \sum_{j=1}^r \psi_{ij}(z)v_{ij}$ for $\varphi \in \mathcal{C}_r$, $\psi \in \mathcal{C}_{r\times r}$, $u \in \mathbb{R}^r$, $v \in \mathbb{R}^{r\times r}$, and $z \in \mathbb{R}^r$ with $\varphi = (\varphi_1, \dots, \varphi_r)$, $\psi = (\psi_{ij})_{r\times r}$, $u = (u_1, \dots, u_r)$, and $v = (v_{ij})_{r\times r}$. For given $A \subset \mathbb{R}^r$, $B \subset \mathbb{R}^{r\times r}$, $\varphi \in \mathcal{C}_r$, $\psi \in \mathcal{C}_{r\times r}$, and $z \in \mathbb{R}^r$, by $\Phi(\varphi, A)(z)$ and $\Psi(\psi, B)(z)$ we denote sets $\Phi(\{\varphi(z)\} \times A)$ and $\Psi(\{\psi(z)\} \times B)$, respectively.

In what follows, we shall restrict the functional parameters φ and ψ to the set $K_k = \{z \in \mathbb{R}^r : |z| \leq k\}$ for $k = 1, 2, \dots$ and consider mappings Φ and Ψ on the spaces $\mathcal{C}_r^k \times \mathbb{R}^r \times \mathbb{R}^r$ and $\mathcal{C}_{r\times r}^k \times \mathbb{R}^{r\times r} \times \mathbb{R}^r$, where $\mathcal{C}_r^k = C_b(K_k, \mathbb{R}^r)$ and $\mathcal{C}_{r\times r}^k = C_b(K_k, \mathbb{R}^{r\times r})$, respectively. We shall also consider mappings Φ and Ψ restricted to the sets $\{\varphi(h) : h \in C_b^2(\mathbb{R}^r, \mathbb{R})\} \times \mathbb{R}^r \times \mathbb{R}^r$ and $\{\psi(h) : h \in C_b^2(\mathbb{R}^r, \mathbb{R})\} \times \mathbb{R}^{r\times r} \times \mathbb{R}^{r\times r}$, respectively, where $\varphi(h) = (h'_{x_1}, \dots, h'_{x_r})$ and $\psi(h) = (h''_{x_i x_j})_{r\times r}$ for $h \in C_b^2(\mathbb{R}^r, \mathbb{R})$. For simplicity, the spaces $C([0, T], \mathbb{R}^r)$ and $C_b^2(\mathbb{R}^r)$ will be denoted by C_T and C_b^2, respectively.

For given measurable and uniformly square integrably bounded set-valued mappings $F : [0,T] \times \mathbb{R}^r \to Cl(\mathbb{R}^r)$ and $G : [0,T] \times \mathbb{R}^r \to Cl(\mathbb{R}^{r\times m})$, $l \in \mathcal{C}_1$, and r-dimensional $\mathbb{F}$-nonanticipative stochastic process $z = (z_t)_{0\le t\le T}$ on a filtered probability space $\mathcal{P}_{\mathbb{F}}$, the integrals $E\int_s^t [l(z_s)\Phi(\varphi, F(\tau, z_\tau))(z_\tau)]\mathrm{d}\tau$ and $E\int_s^t [l(z_s)\Psi(\psi, D(G(\tau, z_\tau)))(z_\tau)\mathrm{d}\tau$ are understood as set-valued Aumann integrals of $\Sigma_{\mathbb{F}}$-measurable set-valued multifunctions $\mathbb{1}_{[s,t]}l(x_s)\Phi(\varphi, F \circ z)(z)$ and $\mathbb{1}_{[s,t]}l(z_s)\Psi(\psi, D(G\circ z))(z)$, respectively, with respect to the product measure $dt\times P$ on the σ-algebra $\Sigma_{\mathbb{F}}$, i.e., as set-valued integrals of the form

$$\int_s^t [l(z_s)\Phi(\varphi, F(\tau, z_\tau))(z_\tau)]\,\mathrm{d}\tau$$
$$= \left\{\int_\Omega \int_0^T \mathbb{1}_{[s,t]}(\tau)l(z_s)\alpha_\tau \mathrm{d}\tau\mathrm{d}P : \alpha \in S_{\mathbb{F}}\left[\Phi(\varphi, F \circ z)(z)\right]\right\}$$

and

$$E\int_s^t [l(z_s)\Psi(\psi, D(G(\tau, z_\tau)))(z_\tau)]\,\mathrm{d}\tau$$
$$= \left\{\int_\Omega \int_0^T \mathbb{1}_{[s,t]}(\tau)l(z_s)\beta_\tau \mathrm{d}\tau\mathrm{d}P : \beta \in S_{\mathbb{F}}\left[\Psi(\psi, D(G \circ z))(z)\right]\right\}$$

for every $0 \le s < t \le T$.

Remark 4.2. If F, G, l, φ, and ψ are as above with convex-valued multifunctions F and $D(G)$, then for every r-dimensional $\mathbb{F}$-nonanticipative process $z = (z_t)_{0\le t\le T}$, the above-defined integrals are compact, convex subsets of $\mathbb{R}^r$.

Proof. Let us observe that for F and G with the properties mentioned above, their subtrajectory integrals $S_{\mathbb{F}}[\Phi(\varphi, F \circ z)(z)]$ and $S_{\mathbb{F}}[\Psi(\psi, D(G \circ z))(z)]$ are convex and sequentially weakly compact subsets of $\mathbb{L}^2([0,T]\times\Omega, \Sigma_{\mathbb{F}}, \mathbb{R}^r)$, and $\mathbb{L}^2([0,T]\times\Omega, \Sigma_{\mathbb{F}}, \mathbb{R}^{r\times r})$, respectively. Hence it follows that the above integrals are compact, convex subsets of $\mathbb{R}^r$. □

Remark 4.3. If F, G, l, φ, and ψ are as above, then for every r-dimensional $\mathbb{F}$-nonanticipative process $z = (z_t)_{0\le t\le T}$, one has

$$\sigma(p, E\int_s^t l(z_s)\Phi(\varphi, F(\tau, z_\tau))(z_\tau)\mathrm{d}\tau) = E\int_s^t \sigma(p, l(z_s)\Phi(\varphi, F(\tau, z_\tau))(z_\tau))\mathrm{d}\tau$$

and

$$\sigma(p, E\int_s^t l(z_s)\Psi(\psi, D(G(\tau, z_\tau)))(z_\tau))\mathrm{d}\tau) = E\int_s^t \sigma(p, l(z_s)\Psi(\psi, D(G(\tau, z_\tau)))(z_\tau))\mathrm{d}\tau$$

for every $0 \le s < t \le T$ and $p \in \mathbb{R}$.

Proof. For the first above-defined integral, we have

$$\sigma(p, E\int_s^t l(z_s)\Phi(\varphi, F(\tau, z_\tau))(z_\tau)\mathrm{d}\tau)$$
$$= \sup\left\{\left\langle p, \int_\Omega \int_s^t l(z_s)\alpha_\tau \mathrm{d}\tau \mathrm{d}P\right\rangle : \alpha \in S_{\mathbb{F}}\left[\Phi(\varphi, F\circ z)(z)\right]\right\}.$$

Hence, immediately from Corollary 3.2 of Chap. 2 applied to a $\Sigma_{\mathbb{F}}$-measurable multifunction, it follows that

$$\sup\left\{\left\langle p, \int_\Omega \int_s^t l(z_s)\alpha_\tau \mathrm{d}\tau \mathrm{d}P\right\rangle : \alpha \in S_{\mathbb{F}}\left[\Phi(\varphi, F\circ z)(z)\right]\right\}$$
$$= \int_\Omega \int_s^t \sup\{\langle p, x\rangle : x \in l(z_s)\Phi(\varphi, F(\tau, z_\tau))(z_\tau)\mathrm{d}\tau\mathrm{d}P\} = E\int_s^t \sigma(p, l(z_s)\Phi(\varphi, F(\tau, z_\tau))(z_\tau))\mathrm{d}\tau$$

for every $p \in \mathbb{R}$. Hence the first required equality follows. In a similar way, we can verify that the second equality is also satisfied. □

Lemma 4.2. *Assume that $\lambda : [0,T]\times\mathbb{R}^r \to \mathbb{R}$ is measurable and uniformly integrably bounded such that $\lambda(t,\cdot)$ is continuous for fixed $t \in [0,T]$. For every r-dimensional continuous processes $x = (x_t)_{0\le t\le T}$ and $\tilde{x} = (\tilde{x}_t)_{0\le t\le T}$ defined on probability spaces $(\Omega, \mathcal{F}, P)$ and $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$, respectively, such that $Px^{-1} = P\tilde{x}^{-1}$, one has*

$$E\int_0^T \lambda(t, x_t)\mathrm{d}t = \tilde{E}\int_0^T \lambda(t, \tilde{x}_t)\mathrm{d}t,$$

where E and $\tilde{E}$ denote the mean value operators with respect to the probability measures P and $\tilde{P}$, respectively.

Proof. Let $I : \mathcal{C}_r \to \mathbb{R}$ be defined by $I(z) = \int_0^T \lambda(t, z(t))\mathrm{d}t$ for $z \in \mathcal{C}_r$. It is clear that I is continuous on $\mathcal{C}_r$. Hence, by the properties of the processes x and $\tilde{x}$, it follows $E[I(x)] = \tilde{E}[I(\tilde{x})]$. Then $E\int_0^T \lambda(t, x_t)\mathrm{d}t = \tilde{E}\int_0^T \lambda(t, \tilde{x}_t)\mathrm{d}t$. □

Lemma 4.3. *Assume that $F : [0,T]\times\mathbb{R}^r \to Cl(\mathbb{R}^r)$ and $G : [0,T]\times\mathbb{R}^r \to Cl(\mathbb{R}^{r\times m})$ are measurable and uniformly square integrably bounded such that $F(t,\cdot)$ and $G(t,\cdot)$ are continuous for fixed $t \in [0,T]$. Let F and $D(G)$ be convex-valued, and let $x = (x_t)_{0\le t\le T}$ and $\tilde{x} = (\tilde{x}_t)_{0\le t\le T}$ be r-dimensional continuous $\mathbb{F}$- and $\tilde{\mathbb{F}}$-nonanticipative processes on filtered probability spaces $(\Omega, \mathcal{F}, \mathbb{F}, P)$ and $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{F}}, \tilde{P})$, respectively, such that $Px^{-1} = P\tilde{x}^{-1}$. For every $l \in \mathcal{C}_1$, $\varphi \in \mathcal{C}_r$, $\psi \in \mathcal{C}_{r\times r}$, and $0 \le s < t \le T$, one has*

$$E\int_s^t l(x_s)\Phi(\varphi, F(\tau, x_\tau))(x_\tau)\mathrm{d}\tau = \tilde{E}\int_s^t l(\tilde{x}_s)\Phi(\varphi, F(\tau, \tilde{x}_\tau))(\tilde{x}_\tau)\mathrm{d}\tau$$

and

$$E\int_s^t l(x_s)\Psi(\psi, D(G(\tau, x_\tau)))(x_\tau)\mathrm{d}\tau = \tilde{E}\int_s^t l(\tilde{x}_s)\Psi(\psi, D(G(\tau, \tilde{x}_\tau)))(\tilde{x}_\tau)\mathrm{d}\tau.$$

Proof. Let $\lambda_p(t,u) = \sigma(p, l(u)\Phi(\varphi, F(t,u)(u))$ for $t \in [0,T]$, $p \in \mathbb{R}$, and $u \in \mathbb{R}^r$. By the definition of $\Phi(\varphi, F(t,u)(u))$, the properties of F, Remark 2.3 of Chap. 2, and Remark 2.4 of Chap. 2, it follows that the function λ_p satisfies the assumptions of Lemma 4.2 for every $p \in \mathbb{R}$. Then $E\int_s^t \lambda_p(t,x_t)\mathrm{d}t = \tilde{E}\int_s^t \lambda_p(t,\tilde{x}_t)\mathrm{d}t$ for every $p \in \mathbb{R}$. By virtue of Remark 4.3, for every $p \in \mathbb{R}$, one has

$$\sigma\left(p, E\int_s^t l(x_s)\Phi(\varphi, F(\tau, x_\tau))(x_\tau)\mathrm{d}\tau\right) = E\int_s^t \lambda_p(\tau, x_\tau)\mathrm{d}\tau$$

and

$$\sigma\left(p, \tilde{E}\int_s^t l(\tilde{x}_s)\Phi(\varphi, F(\tau, \tilde{x}_\tau))(\tilde{x}_\tau)\mathrm{d}\tau\right) = \tilde{E}\int_s^t \lambda_p(\tau, \tilde{x}_\tau)\mathrm{d}\tau.$$

Then

$$\sigma\left(p, E\int_s^t l(x_s)\Phi(\varphi, F(\tau, x_\tau))(x_\tau)\mathrm{d}\tau\right) = \sigma\left(p, \tilde{E}\int_s^t l(\tilde{x}_s)\Phi(\varphi, F(\tau, \tilde{x}_\tau))(\tilde{x}_\tau)\mathrm{d}\tau\right)$$

for every $p \in \mathbb{R}$, which by virtue of Remark 4.2, implies the first equality. In a similar way, we can verify that the second equality is satisfied. □

Let $F : [0,T] \times \mathbb{R}^r \to Cl(\mathbb{R}^r)$ and $G : [0,T] \times \mathbb{R}^r \to Cl(\mathbb{R}^{r\times m})$ be measurable and uniformly square integrably bounded and let $x = (x_t)_{t\in[0,T]}$ be an r-dimensional $\mathbb{F}$-nonanticipative continuous stochastic process on a filtered probability space $\mathcal{P}_{\mathbb{F}}$. Similarly as above, by $S_{\mathbb{F}}(F \circ x)$ and $S_{\mathbb{F}}(G \circ x)$, we denote the sets of all $\mathbb{F}$-nonanticipative selectors of the set-valued stochastic processes $(F(t,x_t))_{0\le t\le T}$ and $(G(t,x_t))_{0\le t\le T}$, respectively. It is clear that for F and G as given above, one has $S_{\mathbb{F}}(F \circ x) \subset \mathbb{L}^2([0,T]\times\Omega, \Sigma_{\mathbb{F}}, \mathbb{R}^r)$ and $S_{\mathbb{F}}(G \circ x) \subset \mathbb{L}^2([0,T]\times\Omega, \Sigma_{\mathbb{F}}, \mathbb{R}^{r\times m})$.

For every $(f,g) \in S_{\mathbb{F}}(F \circ x) \times S_{\mathbb{F}}(G \circ x)$, let us define the semielliptic partial differential operator $\mathbb{L}^x_{fg}$ by setting

$$\mathbb{L}^x_{fg}(\varphi, \psi)_t = \sum_{i=1}^n \varphi_i(x_t) f_t^i + \frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n \psi_{ij}(x_t)\sigma_t^{ij}$$

for $t \in [0,T]$, $\varphi = (\varphi_i)_{1\times r} \in \mathcal{C}_r$, $\psi = (\psi_{ij})_{r\times r} \in \mathcal{C}_{r\times r}$, where $\sigma = g \cdot g^*$. For every $h \in C_b^2(\mathbb{R}^r)$ and $t \in [0,T]$, by $(\mathbb{L}^x_{fg}h)_t$ we shall denote the random variable $\mathbb{L}^x_{fg}(\varphi(h), \psi(h))_t$, where $\varphi(h)$ and $\psi(h)$ are as above. We shall show that if the assumptions of Lemma 4.3 are satisfied, then for every $f \in S_{\mathbb{F}}(F \circ x)$ and $g \in S_{\mathbb{F}}(G \circ x)$, there exist $\tilde{f} \in S_{\tilde{\mathbb{F}}}(\mathrm{co}F \circ \tilde{x})$ and $\tilde{g} \in S_{\tilde{\mathbb{F}}}(G \circ \tilde{x})$ such that

$$E\int_s^t l(x_s)\mathbb{L}^x_{fg}(\varphi, \psi)_\tau \mathrm{d}\tau = \tilde{E}\int_s^t l(\tilde{x}_s)\mathbb{L}^{\tilde{x}}_{\tilde{f}\tilde{g}}(\varphi, \psi)_\tau \mathrm{d}\tau$$

for every $0 \le s < t \le T$, $l \in \mathcal{C}_1$, $\varphi \in \mathcal{C}_r$, and $\psi \in \mathcal{C}_{r\times r}$. To get such a result, we begin with the following lemmas.

Lemma 4.4. *Assume that $F : [0,T] \times \mathbb{R}^r \to Cl(\mathbb{R}^r)$ and $G : [0,T] \times \mathbb{R}^d \to Cl(\mathbb{R}^{r\times m})$ are measurable and uniformly square integrably bounded such that $F(t,\cdot)$ and $G(t,\cdot)$ are continuous for fixed $t \in [0,T]$. Let G be diagonally convex and $x = (x_t)_{0\le t\le T}$ a continuous r-dimensional $\mathbb{F}$-adapted process on $\mathcal{P}_{\mathbb{F}}$. Assume that $\tilde{x} = (\tilde{x}_t)_{0\le t\le T}$ is a continuous r-dimensional $\tilde{\mathbb{F}}$-adapted process on $\tilde{\mathcal{P}}_{\tilde{\mathbb{F}}} = (\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{F}}, \tilde{P})$ with $\tilde{\mathbb{F}} = (\tilde{\mathcal{F}}_t)_{0\le t\le T}$ such that $Px^{-1} = P\tilde{x}^{-1}$. Then for every $f \in S_{\mathbb{F}}(\mathrm{co}\, F \circ x)$, $g \in S_{\mathbb{F}}(G \circ x)$, $l \in \mathcal{C}_1$, $\varphi \in \mathcal{C}_r$, and $\psi \in \mathcal{C}_{r\times r}$, there are $\tilde{\mathbb{F}}$-nonanticipative processes $(\tilde{\alpha}_t(l,\varphi))_{0\le t\le T}$ and $(\tilde{\beta}_t(l,\psi))_{0\le t\le T}$ on $\tilde{\mathcal{P}}_{\tilde{\mathbb{F}}}$ such that*

(i)

$$\tilde{\alpha}_t(l,\varphi) \in \Phi(\varphi, \mathrm{co}\, F(t,\tilde{x}_t))(\tilde{x}_t), \quad \text{a.e.} \quad \text{on} \quad [0,T]\times\Omega,$$

(ii)

$$\tilde{\beta}_t(l,\psi) \in \Psi(\psi, D(G(t,\tilde{x}_t)))(\tilde{x}_t) \quad \text{a.e.} \quad \text{on} \quad [0,T]\times\Omega$$

and

(iii)

$$E\int_s^t l(x_s)\mathbb{L}^x_{fg}(\varphi,\psi)_\tau \mathrm{d}\tau = \tilde{E}\int_s^t l(\tilde{x}_s)[\tilde{\alpha}_\tau(l,\varphi) + \frac{1}{2}\tilde{\beta}_\tau(l,\psi)]\mathrm{d}\tau$$

for every $0 \le s < t \le T$.

Proof. Let $f \in S_{\mathbb{F}}(\mathrm{co}\, F \circ x)$, $g \in S_{\mathbb{F}}(G \circ x)$, $l \in \mathcal{C}_1$, $\varphi \in \mathcal{C}_r$, and $\psi \in \mathcal{C}_{r\times r}$ be given and let $(\alpha_t)_{0\le\le T}$ and $(\beta_t)_{0\le\le T}$ be defined by

$$\alpha_t = \sum_{i=1}^r \varphi_i(x_t) f_t^i \quad \text{and} \quad \beta_t = \sum_{i=1}^r\sum_{j=1}^r \psi_{ij}(x_t)\sigma_t^{ij} \quad \text{for} \quad t \in [0,T].$$

We have $\alpha_t \in \Phi(\varphi, \mathrm{co}\, F(t,x_t))(x_t)$ and $\beta_t \in \Psi(\psi, D(G(t,x_t)))(x_t)$ a.e. on $[0,T]\times\Omega$, which by virtue of Lemma 4.3, implies that

$$E\int_s^t l(x_s)\alpha_\tau \mathrm{d}\tau) \in \tilde{E}\int_s^t l(\tilde{x}_s)\Phi(\varphi, co\, F(\tau,\tilde{x}_\tau))(\tilde{x}_\tau))\mathrm{d}\tau$$

and

$$E\int_s^t l(x_s)\beta_\tau \mathrm{d}\tau \in \tilde{E}\int_s^t l(\tilde{x}_s)\Psi(\psi, D(G(\tau,\tilde{x}_\tau)))(\tilde{x}_\tau))\mathrm{d}\tau$$

for every $0 \le s \le t \le T$. Let $L > 0$ be such that $|l(x)| \le L$ for every $x \in \mathbb{R}^r$. By the definition of Φ and the properties of φ and F, it follows that there exists $m \in \mathbb{L}^2([0,T]\mathcal{F}_T, \mathbb{R}^+)$ such that $\|\Phi(\varphi, \mathrm{co}\, F(t,x))(x)\| \le m(t)$ for every $x \in \mathbb{R}^r$ and a.e. $0 \le t \le T$.

Let $\varepsilon > 0$ and select $\delta > 0$ such that $\sup_{0\le t\le T}\int_t^{t+\delta} m(\tau)\mathrm{d}\tau \le \varepsilon/4L$. Put $\tau_0 = 0$ and $\tau_k = k\delta$ for $k = 1,\ldots,N$, where N is such that $(N-1)\delta < T \le N\delta$.

From the above inclusions and the definition of the set-valued integral $\tilde{E}\int_s^t l(\tilde{x}_s)$ $\Phi(\varphi, \text{co}\, F(\tau, \tilde{x}_\tau))\ (\tilde{x}_\tau))\text{d}\tau$, it follows that for every $k = 1, \ldots, N$, there exists $(\tilde{\alpha}_t^k)_{0\le t\le T} \in S_{\tilde{\mathbb{F}}}(\Phi(\varphi, \text{co}\, F\ (\cdot, \tilde{x})(\tilde{x}))$ such that $E[\int_{\tau_{k-1}}^{\tau_k} l(x_\tau)\alpha_\tau \text{d}\tau] = \tilde{E}[\int_{\tau_{k-1}}^{\tau_k} l(\tilde{x}_\tau)$ $\tilde{\alpha}_\tau^k \text{d}\tau]$. Define $\tilde{\alpha}^\varepsilon = \mathbb{1}_{\{0\}}\tilde{\alpha}_0^1 + \mathbb{1}_{(\tau,\tau_1]}\tilde{\alpha}^1 + \ldots + \mathbb{1}_{(\tau_{N-1},T]}\tilde{\alpha}^N$. We get $\tilde{\alpha}^\varepsilon \in S_{\tilde{\mathbb{F}}}(\Phi(\varphi, \text{co}\, F(\cdot, \tilde{x}))(\tilde{x}))$, because $S_{\tilde{\mathbb{F}}}(\Phi(\varphi, \text{co}\, F(\cdot, \tilde{x}))(\tilde{x}))$ is decomposable. Let $x^\varepsilon = \mathbb{1}_{[0,\tau_1]}x_0 + \mathbb{1}_{(\tau_1,\tau_2]}x_{\tau_1} + \ldots + \mathbb{1}_{(\tau_{N-1},T]}x_{\tau_{N-1}}$ and $\tilde{x}^\varepsilon = \mathbb{1}_{[0,\tau_1]}\tilde{x}_0 + \mathbb{1}_{(\tau_1,\tau_2]}\tilde{x}_{\tau_1} + \ldots + \mathbb{1}_{(\tau_{N-1},T]}\tilde{x}_{\tau_{N-1}}$. By the continuity of the mapping l and processes x and $\tilde{x}$, we have $\sup_{0\le t\le T} E|l(x_t) - l(x_t^\varepsilon)| \to 0$ and $\sup_{0\le t\le T} \tilde{E}|l(\tilde{x}_t) - l(\tilde{x}_t^\varepsilon)| \to 0$ as $\varepsilon \to 0$. Thus for every $\eta > 0$, there exists $\varepsilon_\eta > 0$ such that for every $\varepsilon < \varepsilon_\eta$, one has $\sup_{0\le t\le T} E|l(x_t) - l(x_t^\varepsilon)| \le \eta/2M$ and $\sup_{0\le t\le T} \tilde{E}|l(\tilde{x}_t) - l(\tilde{x}_t^\varepsilon)| \le \eta/2M$, where $M = \int_0^T m(t)\text{d}t$. We shall show that $E\int_s^t l(x_s)\alpha_\tau \text{d}\tau = \tilde{E}\int_s^t l(\tilde{x}_s)\tilde{\alpha}_\tau \text{d}\tau$ for $0 \le s < t \le T$.

For every fixed $0 \le s \le t \le T$, there are positive integers $1 \le r < l \le N$ such that $s \in (\tau_{r-1}, \tau_r]$ and $t \in (\tau_l, \tau_{l-1}]$ or $s, t \in (\tau_{r-1}, \tau_r]$ or $s, t \in (\tau_{l-1}, \tau_l]$. In the last two cases, we have

$$\left|E\int_s^t l(x_s)\alpha_\tau \text{d}\tau - \tilde{E}\int_s^t l(\tilde{x}_s)\tilde{\alpha}_\tau^\varepsilon\right| \le 2L\int_s^t m_\tau \text{d}\tau \le \varepsilon/2 < \varepsilon.$$

If $s \in (\tau_{r-1}, \tau_r]$ and $t \in (\tau_{l-1}, \tau_l]$, we obtain

$$\begin{aligned}
&\left|E\int_s^t l(x_s)\alpha_\tau \text{d}\tau - \tilde{E}\int_s^t l(\tilde{x}_s)\tilde{\alpha}_\tau^\varepsilon \text{d}\tau\right| \le \left|E\int_s^t l(x_s)\alpha_\tau \text{d}\tau - E\int_s^t l(x_s^\varepsilon)\alpha_\tau \text{d}\tau\right| \\
&\quad + \left|E\int_s^t l(x_s^\varepsilon)\alpha_\tau \text{d}\tau - \tilde{E}\int_s^t l(\tilde{x}_s^\varepsilon)\tilde{\alpha}_\tau^\varepsilon \text{d}\tau\right| + \left|\tilde{E}\int_s^t l(\tilde{x}_s^\varepsilon)\tilde{\alpha}_\tau^\varepsilon \text{d}\tau - \tilde{E}\int_s^t l(\tilde{x}_s)\tilde{\alpha}_\tau^\varepsilon \text{d}\tau\right| \\
&\le M\sup_{0\le t\le T} E|l(x_t) - l(x_t^\varepsilon)| + M\sup_{0\le t\le T} \tilde{E}|l(\tilde{x}_t) - l(\tilde{x}_t^\varepsilon)| \\
&\quad + \left|E\int_s^t l(x_s^\varepsilon)\alpha_\tau \text{d}\tau - \tilde{E}\int_s^t l(\tilde{x}_s^\varepsilon)\tilde{\alpha}_\tau^\varepsilon \text{d}\tau\right| \le \eta + \left|E\int_s^t l(x_s^\varepsilon)\alpha_\tau \text{d}\tau - \tilde{E}\int_s^t l(\tilde{x}_s^\varepsilon)\tilde{\alpha}_\tau^\varepsilon \text{d}\tau\right|.
\end{aligned}$$

But

$$\begin{aligned}
&\left|E\int_s^t l(x_s^\varepsilon)\alpha_\tau \text{d}\tau - \tilde{E}\int_s^t l(\tilde{x}_s^\varepsilon)\tilde{\alpha}_\tau^\varepsilon \text{d}\tau\right| \le \left|E\int_s^{\tau_r} l(x_s^\varepsilon)\alpha_\tau \text{d}\tau\right. \\
&\quad \left.-\tilde{E}\int_s^{\tau_r} l(\tilde{x}_s^\varepsilon)\tilde{\alpha}_\tau^\varepsilon \text{d}\tau\right| + \sum_{i=r+1}^{l-1}\left|E\int_{\tau_{i-1}}^{\tau_i} l(x_{\tau_{i-1}})\alpha_\tau \text{d}\tau - \tilde{E}\int_{\tau_{i-1}}^{\tau_i} l(\tilde{x}_{\tau_{i-1}})\tilde{\alpha}_\tau^i \text{d}\tau\right| \\
&\quad + \left|E\int_{\tau_{l-1}}^t l(x_\tau)\alpha_\tau \text{d}\tau - \tilde{E}\int_{\tau_{l-1}}^t l(\tilde{x}_\tau)\tilde{\alpha}_\tau^l \text{d}\tau\right| \le 2L\int_s^{\tau_r} m(t)\text{d}t + 2L\int_{\tau_{l-1}}^t m(t)\text{d}t \le \varepsilon.
\end{aligned}$$

Therefore, for every fixed $s \in (\tau_{r-1}, \tau_r]$, $t \in (\tau_{l-1}, \tau_l]$, and $\varepsilon \in (0, \varepsilon_\eta)$, one gets $|E\int_s^t l(x_s)\alpha_\tau \text{d}\tau - \tilde{E}\int_s^t l(\tilde{x}_s)\tilde{\alpha}_\tau^\varepsilon \text{d}\tau| \le \eta + \varepsilon$. By the sequential weak compactness of $S_{\tilde{\mathbb{F}}}(\Phi(\varphi, \mathrm{co}\, F(\cdot, \tilde{x}))(\tilde{x})) \subset \mathbb{L}^2([0,T] \times \tilde{\Omega}, \Sigma_{\tilde{\mathbb{F}}}, \mathbb{R}^r)$, for every sequence $(\varepsilon_n)_{n=1}^{\infty}$ of positive numbers with $\varepsilon_n \to 0$, we can select a subsequence, say $(\varepsilon_k)_{k=1}^{\infty}$, such that $(\tilde{\alpha}^{\varepsilon_k})_{k=1}^{\infty}$ converges weakly to an $\tilde{\alpha} \in S_{\tilde{\mathbb{F}}}(\Phi(\varphi, \mathrm{co}\, F(\cdot, \tilde{x})))(\tilde{x}))$ as $k \to \infty$. Therefore, $|E \int_s^t l(x_s)\alpha_\tau \mathrm{d}\tau - \tilde{E} \int_s^t l(\tilde{x}_s)\tilde{\alpha}_\tau \mathrm{d}\tau| \le \eta$ for every $\eta > 0$, which implies that $E \int_s^t l(x_s)\alpha_\tau \mathrm{d}\tau = \tilde{E} \int_s^t l(\tilde{x}_s)\tilde{\alpha}_\tau \mathrm{d}\tau$ for $0 \le s < t \le T$. Taking $\tilde{\alpha}(l, \varphi) = \tilde{\alpha}$, we obtain condition (i). In a similar way, we can establish the existence of $\tilde{\beta}(l, \psi) = \tilde{\beta} \in S_{\tilde{\mathbb{F}}}(\Psi(\psi, D(G(\cdot, \tilde{x})))(\tilde{x})$ such that $E \int_s^t l(x_s)\beta_\tau \mathrm{d}\tau = \tilde{E} \int_s^t l(\tilde{x}_s)\tilde{\beta}_\tau \mathrm{d}\tau$ for every $0 \le s < t \le T$. Thus (ii) is satisfied. By the definition of $\mathbb{L}_{fg}^x(\varphi, \psi)$ and conditions (i) and (ii), it follows that (iii) is also satisfied. □

Lemma 4.5. *Suppose the assumptions of Lemma 4.4 are satisfied and let* $\tau_k = \inf\{t \in [0,T] : x_t \notin K_k\} \wedge T$ *and* $\tilde{\tau}_k = \inf\{t \in [0,T] : \tilde{x} \notin K_k\} \wedge T$, *where* $K_k = \{z \in \mathbb{R}^d : |z| \le k\}$ *for* $k = 1, 2, \ldots$. *For every* $f \in S_{\mathbb{F}}(co\, F \circ x)$, $g \in S_{\mathbb{F}}((G \circ x)$, $l \in \mathcal{C}_1$, $\varphi \in \mathcal{C}_r$, $\psi \in \mathcal{C}_{r\times r}$, *and* $k = 1, 2, \ldots$, *there are* $\tilde{\mathbb{F}}$-*nonanticipative processes* $(\tilde{\alpha}_t^k(l, \varphi)_{0 \le t \le T}$ *and* $(\tilde{\beta}_t^k(l, \psi)_{0 \le t \le T}$ *such that for every* $k \ge 1$, *one has*

(i) $\tilde{\alpha}_t^k(l, \varphi) \in \Phi(\varphi, \mathrm{co}\, F(t, \tilde{x}_{t\wedge\tilde{\tau}_k}))(\tilde{x}_{t\wedge\tilde{\tau}_k}))$, *a.e. on* $[0,T] \times \Omega$;
(ii) $\tilde{\beta}_t^k(l, \psi) \in \Psi(\psi, D(G(t, \tilde{x}_{t\wedge\tilde{\tau}_k})))(\tilde{x}_{t\wedge\tilde{\tau}_k}))$, *a.e. on* $[0,T] \times \Omega$;
(iii) $E \int_{s\wedge\tau_k}^{t\wedge\tau_k} l(x_{s\wedge\tau_k})\mathbb{L}_{fg}^x(\varphi, \psi)\mathrm{d}\tau = \tilde{E} \int_{s\wedge\tilde{\tau}_k}^{t\wedge\tilde{\tau}_k} l(\tilde{x}_{s\wedge\tilde{\tau}_k})[\tilde{\alpha}_\tau^k(l, \varphi) + \frac{1}{2}\tilde{\beta}_\tau^k\,(l, \psi)]\mathrm{d}\tau$ *for every* $0 \le s < t \le T$;
(iv) $\tilde{\alpha}_t^k(l, \varphi)$ *and* $\tilde{\beta}_t^k(l, \psi)$ *are continuous with respect to* (l, φ) *and* (l, ψ) *on* $\mathcal{C}_1^k \times \mathcal{C}_r^k$ *and* $\mathcal{C}_1^k \times \mathcal{C}_{r\times r}^k$, *respectively, a.s. for fixed* $t \in [0,T]$.

Proof. Let us observe that $\mathcal{C}_1^k, \mathcal{C}_r^k$ and $\mathcal{C}_{r\times r}^k$ are separable metric spaces for $k = 1, 2, \ldots$. Denote countable dense subsets each by $\mathcal{D}_1^k, \mathcal{D}_r^k$ and $\mathcal{D}_{r\times r}^k$, respectively, and assume that $\mathcal{D}_1^k = \{l_1, l_2, \ldots\}$, $\mathcal{D}_r^k = \{\varphi_1, \varphi_2, \ldots\}$ and $\mathcal{D}_{r\times r}^k = \{\psi_1, \psi_2, \ldots\}$. Similarly as in Lemma 4.4, we can show that for every fixed $k = 1, 2, \ldots$ and $i = 1, 2, \ldots$, there are $\tilde{\mathbb{F}}$-nonanticipative processes $(\tilde{\alpha}_t^i)_{0 \le t \le T}$ and $(\tilde{\beta}_t^i)_{0 \le t \le T}$ such that

(i′) $\tilde{\alpha}_t^i \in \Phi(\varphi_i, \mathrm{co}\, F(t, \tilde{x}_{t\wedge\tilde{\tau}_k}))(\tilde{x}_{t\wedge\tilde{\tau}_k})$, a.e. on $[0,T] \times \Omega$;
(ii′) $\tilde{\beta}_t^i \in \Psi(\psi_i, D(G(t, \tilde{x}_{t\wedge\tilde{\tau}_k})))(\tilde{x}_{t\wedge\tilde{\tau}_k})$ a.e. on $[0,T] \times \Omega$;
(iii′) $E \int_{s\wedge\tau_k}^{s\wedge\tau_k} l_i(x_{s\wedge\tau_k})\mathbb{L}_{fg}^x(\varphi_i, \psi_i)_\tau \mathrm{d}\tau = \tilde{E} \int_{s\wedge\tilde{\tau}_k}^{s\wedge\tilde{\tau}_k} l_i(\tilde{x}_{s\wedge\tilde{\tau}_k})[\tilde{\alpha}_\tau^i + \frac{1}{2}\tilde{\beta}_\tau^i]\mathrm{d}\tau$;

for $k \ge 1$ and every $0 \le s < t \le T$. Define on $\mathcal{D}_1^k \times \mathcal{D}_r^k$ and $\mathcal{D}_1^k \times \mathcal{D}_{r\times r}^k$ multifunctions $\Phi_t^i(l, \varphi)$ and $\Psi_t^i(l, \psi)$ by setting

$$\Phi_t^i(l, \varphi) = \begin{cases} \Phi(\varphi, \mathrm{co}\, F(t, \tilde{x}_{t\wedge\tilde{\tau}_k}))(\tilde{x}_{t\wedge\tilde{\tau}_k}) & \text{for } (l, \varphi) \ne (l_i, \varphi_i), \\ \{\tilde{\alpha}_t^i\} & \text{for } (l, \varphi) = (l_i, \varphi_i), \end{cases}$$

and

$$\Psi_t^i(l,\psi) = \begin{cases} \Psi(\psi, D(G(t,\tilde{x}_{t\wedge\tilde{\tau}_k})))(\tilde{x}_{t\wedge\tilde{\tau}_k}) & \text{for } (l,\psi) \neq (l_i,\psi_i), \\ \{\tilde{\beta}_t^i\} & \text{for } (l,\varphi) = (l_i,\varphi_i), \end{cases}$$

for $i = 1, 2, \ldots$. It is easy to see that $\Phi_t^i(l,\varphi)$ and $\Psi_t^i(l,\psi)$ are closed and convex-valued. Furthermore, similarly as in Remark 2.1 of Chap. 2, it can be verified that the set-valued mappings $\Phi_t^i : \mathcal{C}_1^k \times \mathcal{C}_r^k \ni (l,\varphi) \to \Phi_t^i(l,\varphi) \subset \mathbb{R}$ and $\Psi_t^i : \mathcal{C}_1^k \times \mathcal{C}_{r\times r}^k \ni (l,\psi) \to \Psi_t^i(l,\psi) \subset \mathbb{R}$ are l.s.c., $\tilde{P}$-a.s. for every fixed $t \in [0,T]$. It is clear that stochastic processes $(\Phi_t^i(l,\varphi))_{0\le t\le T}$ and $(\Psi_t^i(l,\psi))_{0\le t\le T}$ are $\tilde{\mathbb{F}}$-nonanticipative for every fixed l, φ and ψ. Therefore, by Theorem 2.7 of Chap. 2, there are mappings $\gamma^i : [0,T] \times \tilde{\Omega} \times \mathcal{C}_1^k \times \mathcal{C}_r^k \ni (t,\tilde{\omega},l,\varphi) \to \gamma_t^i(l,\varphi)(\tilde{\omega}) \in \mathbb{R}$ and $\lambda^i : [0,T] \times \tilde{\Omega} \times \mathcal{C}_1^k \times \mathcal{C}_{r\times r}^k \ni (t,\tilde{\omega},l,\psi) \to \lambda_t^i(l,\psi)(\tilde{\omega}) \in \mathbb{R}$, $\Sigma_{\tilde{\mathbb{F}}}$-measurable on $[0,T] \times \tilde{\Omega}$ and continuous with respect to (l,φ) and (l,ψ), respectively, such that $\gamma_t^i(l,\varphi) \in \Phi_t^i(l,\varphi)$ and $\lambda_t^i(l,\psi) \in \Psi_t^i(l,\psi)$ a.e. on $[0,T] \times \tilde{\Omega}$ for $(l,\varphi) \in \mathcal{C}_1^k \times \mathcal{C}_r^k$ and $(l,\psi) \in \mathcal{C}_1^k \times \mathcal{C}_{r\times r}^k$, respectively.

Let $(U_i^k)_{i=1}^\infty$ and $(V_i^k)_{i=1}^\infty$ be countable open coverings for $\mathcal{C}_1^k \times \mathcal{C}_r^k$ and $\mathcal{C}_1^k \times \mathcal{C}_{r\times r}^k$, respectively, such that $(l_i,\varphi_i) \in U_i$ and $(l_i,\psi_i) \in V_i^k$ for $i = 1, 2, \ldots$. Select continuous locally finite partitions of unity $(p_i^k)_{i=1}^\infty$ and $(q_i^k)_{i=1}^\infty$ subordinate to $(U_i^k)_{i=1}^\infty$ and $(V_i^k)_{i=1}^\infty$, respectively, and define $\tilde{\alpha}_t^k(l,\varphi)$ and $\tilde{\beta}_t^k(l,\varphi)$ by setting

$$\tilde{\alpha}_t^k(l,\varphi)(\tilde{\omega}) = \sum_{i=1}^\infty p_i^k(l,\varphi) \cdot \gamma^i(t,\tilde{\omega},l,\varphi)$$

and

$$\tilde{\beta}_t^k(l,\psi)(\tilde{\omega}) = \sum_{i=1}^\infty q_i^k(l,\psi) \cdot \lambda^i(t,\tilde{\omega},l,\psi)$$

for $l \in \mathcal{C}_1^k$, $\varphi \in \mathcal{C}_r^k$, $\psi \in \mathcal{C}_{r\times r}^k$ and $(t,\tilde{\omega}) \in [0,T] \times \tilde{\Omega}$. It is clear that

$$\tilde{\alpha}_t^k(l,\varphi) \in \Phi(\varphi, \operatorname{co} F(t,\tilde{x}_{t\wedge\tilde{\tau}_k}))(\tilde{x}_{t\wedge\tilde{\tau}_k}) \qquad \text{on} \qquad [0,T] \times \tilde{\Omega}$$

and

$$\tilde{\beta}_t^k(l,\psi) \in \Psi(\psi, D(G(t,\tilde{x}_{t\wedge\tilde{\tau}_k})))(\tilde{x}_{t\wedge\tilde{\tau}_k}) \qquad \text{on} \qquad [0,T] \times \tilde{\Omega}$$

for $t \in [0,T]$, because $(p_i^k)_{i=1}^\infty$ and $(q_i^k)_{i=1}^\infty$ are locally finite, and the multifunctions $\Phi(\varphi, \operatorname{co} F(t,z))(z)$ and $\Psi(\psi, D(G(t,z)))(z)$ are convex-valued. Immediately from the above definitions, it follows that $\tilde{\alpha}_t^k(\cdot,\cdot)(\tilde{\omega})$ and $\tilde{\beta}_t^k(\cdot,\cdot)(\tilde{\omega})$ are continuous on $\mathcal{C}_1^k \times \mathcal{C}_r^k$ and $\mathcal{C}_1^k \times \mathcal{C}_{r\times r}^k$, respectively, for a.e. fixed $\tilde{\omega} \in \tilde{\Omega}$ and for fixed $t \in [0,T]$. Thus (iv) has been proved.

Finally, by the properties of the above-defined functions $\tilde{\alpha}_\tau^k(l,\varphi)$ and $\tilde{\beta}_\tau^k(l,\psi)$, one gets

$$\mathcal{H}_{st}^k(l,\varphi,\psi) = E\int_{s\wedge\tau_k}^{t\wedge\tau_k} l(x_{s\wedge\tau_k})\mathbb{L}_{fg}^x(\varphi,\psi)_\tau \mathrm{d}\tau - \tilde{E}\int_{s\wedge\tau_k}^{t\wedge\tilde{\tau}_k} l(\tilde{x}_{t\wedge\tilde{\tau}_k})[\tilde{\alpha}_\tau^k(l,\varphi)$$
$$+\frac{1}{2}\tilde{\beta}_\tau^k(l,\psi)]\mathrm{d}\tau = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} p_i(l,\varphi)q_j(l,\psi)\Bigg[E\int_{s\wedge\tau_k}^{t\wedge\tau_k} l(x_{s\wedge\tau_k})\mathbb{L}_{fg}^x(\varphi,\psi)_\tau \mathrm{d}\tau$$
$$-\tilde{E}\int_{s\wedge\tilde{\tau}_k}^{t\wedge\tilde{\tau}_k} l(\tilde{x}_{t\wedge\tilde{\tau}_k})[\gamma^i(\tau,\cdot,l,\varphi)+\frac{1}{2}\lambda^j(\tau,\cdot,l,\psi)]\mathrm{d}\tau\Bigg]$$

for $k \geq 1$, $0 \leq s \leq t \leq T$, $l \in \mathcal{C}_r^k$, $\varphi \in \mathcal{C}_r^k$, and $\psi \in \mathcal{C}_{r\times r}^k$. Hence, by the properties of γ^i and λ^i and (iii)$'$, it follows that for every $k \geq 1$, one has $\mathcal{H}_{st}^k(l_i,\varphi_i,\psi_i) = 0$ for every $i = 1,2,\ldots$ and $0 \leq s < t \leq T$. But $\mathcal{H}_{st}^k$ is continuous on $\mathcal{C}_1^k \times \mathcal{C}_r^k \times \mathcal{C}_{r\times r}^k$ and is equal to zero on $\mathcal{D}_1^k \times \mathcal{D}_r^k \times \mathcal{D}_{r\times r}^k$ for every $k \geq 1$. Hence by the density of the set $\mathcal{D}_1^k \times \mathcal{D}_r^k \times \mathcal{D}_{r\times r}^k$ in $\mathcal{C}_1^k \times \mathcal{C}_r^k \times \mathcal{C}_{r\times r}^k$, we obtain $\mathcal{H}_{st}^k(l,\varphi,\psi) = 0$ for every $k \geq 1$, $l \in \mathcal{C}_1^k$, $\varphi \in \mathcal{C}_r^k$, and $\psi \in \mathcal{C}_{r\times r}^k$ and every $0 \leq s \leq t \leq T$. Thus (iii) is satisfied. □

Lemma 4.6. *Assume that the assumptions and notation of Lemma 4.4 are satisfied. For every* $f \in S_{\mathbb{F}}(\mathrm{co}F \circ x)$, $g \in S_{\mathbb{F}}(G \circ x)$, *and* $k = 1,2,\ldots$, *there are* $\tilde{f}^k \in S_{\widetilde{\mathbb{F}}}(\mathrm{co}\,F \circ (\widetilde{x} \circ \widetilde{\tau}_k))$ *and* $\tilde{g}^k \in S_{\widetilde{\mathbb{F}}}(G \circ (\widetilde{x} \circ \widetilde{\tau}_k))$ *such that for* $k = 1,2,\ldots$, *we have*

$$E\int_{s\wedge\tau_k}^{t\wedge\tau_k} l(x_{s\wedge\tau_k})\mathbb{L}_{fg}^x(\varphi,\psi)_\tau \mathrm{d}\tau = \tilde{E}\int_{s\wedge\tilde{\tau}_k}^{t\wedge\tilde{\tau}_k} l(\tilde{x}_{s\wedge\tilde{\tau}_k})\mathbb{L}_{\tilde{f}^k\tilde{g}^k}^{\tilde{x}}(\varphi,\psi)_\tau \mathrm{d}\tau \tag{4.1}$$

for every $0 \leq s < t \leq T$, $l \in \mathcal{C}_1^k$, $\varphi \in \mathcal{C}_r^k$, *and* $\psi \in \mathcal{C}_{r\times r}^k$, *where* $(\widetilde{x} \circ \widetilde{\tau}_k)_t = \widetilde{x}_{t\wedge\widetilde{\tau}_k}$.

Proof. Let $(\tilde{\alpha}_t^k(l,\varphi))_{0\leq t\leq T}$ and $(\tilde{\beta}_t^k(l,\psi))_{0\leq t\leq T}$ be, for every $k \geq 1$, as in Lemma 4.5, and let us define multifunctions $\mathcal{K}^k$ and Q^k by setting

$$\mathcal{K}_t^k(\tilde{\omega}) = \mathrm{co}\,F(t,\widetilde{x}_{t\wedge\widetilde{\tau}_k}) \cap \{u \in \mathbb{R}^r : \sup_{(l,\varphi)\in\mathcal{C}_1^k\times\mathcal{C}_r^k} \mathrm{dist}(\tilde{\alpha}_t^k(l,\varphi)(\tilde{\omega}),\Phi(\varphi,u)(\tilde{x}_{t\wedge\tilde{\tau}_k}(\tilde{\omega}))) = 0\}$$

and

$$Q_t^k(\tilde{\omega}) = D(G(t,\widetilde{x}_{t\wedge\widetilde{\tau}_k})) \cap \{v \in \mathbb{R}^{r\times r} : \sup_{(l,\psi)\in\mathcal{C}_1^k\times\mathcal{C}_{r\times r}^k} \mathrm{dist}(\tilde{\beta}_t^k(l,\psi)(\tilde{\omega}),\Psi(\psi,v)(\tilde{x}_{t\wedge\tilde{\tau}_k}(\tilde{\omega})))=0\}$$

for $t \in [0,T]$ and $\widetilde{\omega} \in \widetilde{\Omega}$. It is clear that $\mathcal{K}_t^k(\widetilde{\omega}) \in \mathrm{Cl}(\mathbb{R}^r)$ and $Q_t^k(\widetilde{\omega}) \in \mathrm{Cl}(\mathbb{R}^{r\times r})$ for $0 \leq t \leq T$ and $\widetilde{\omega} \in \widetilde{\Omega}$. By the separability of the metric spaces $\mathcal{C}_1^k \times \mathcal{C}_r^k$ and $\mathcal{C}_1^k \times \mathcal{C}_{r\times r}^k$, we have

$$\mathcal{K}_t^k(\tilde{\omega}) = \text{co}\, F(t, \widetilde{x}_{t\wedge\widetilde{\tau}_k}) \cap \{u \in \mathbb{R}^r : \sup_{(l,\varphi)\in\mathcal{D}_1^k\times\mathcal{D}_r^k} \text{dist}(\tilde{\alpha}_t^k(l,\varphi)(\tilde{\omega}), \Phi(\varphi,u)(\tilde{x}_{t\wedge\tilde{\tau}_k}(\tilde{\omega}))) = 0\}$$

and

$$\mathcal{Q}_t^k(\tilde{\omega}) = D(G(t, \widetilde{x}_{t\wedge\widetilde{\tau}_k})) \cap \{v \in \mathbb{R}^{r\times r} : \sup_{(l,\psi)\in\mathcal{D}_1^k\times\mathcal{D}_{r\times r}^k} \text{dist}(\tilde{\beta}_t^k(l,\psi)(\tilde{\omega}), \Psi(\psi,v)(\tilde{x}_{t\wedge\tilde{\tau}_k}(\tilde{\omega}))) = 0\}$$

for $(t,\tilde{\omega}) \in [0,T]\times\tilde{\Omega}$. By continuity of the functions $\text{dist}(\tilde{\alpha}_t^k(\cdot,\cdot)(\tilde{\omega}), \Phi(\cdot,u),(\tilde{x}_{t\wedge\tilde{\tau}_k}(\tilde{\omega})))$, and $\text{dist}(\tilde{\beta}_t^k(\cdot,\cdot)(\tilde{\omega}), \Psi(\cdot,v)(\tilde{x}_{t\wedge\tilde{\tau}_k}(\tilde{\omega})))$, for fixed $(t,\tilde{\omega}) \in [0,T]\times\tilde{\Omega}$, $u\in\mathbb{R}^r$, and $v\in\mathbb{R}^{r\times r}$, it follows that the mappings

$$[0,T]\times\tilde{\Omega} \ni (t,\tilde{\omega}) \to \text{dist}(\tilde{\alpha}_t^k(l,\varphi)(\tilde{\omega}), \Phi(\varphi,u)(\tilde{x}_{t\wedge\tilde{\tau}_k}(\tilde{\omega}))) \in \mathbb{R}$$

and

$$[0,T]\times\tilde{\Omega} \ni (t,\tilde{\omega}) \to \text{dist}(\tilde{\beta}_t^k(l,\psi)(\tilde{\omega}), \Psi(\psi,v)(\tilde{x}_{t\wedge\tilde{\tau}_k}(\tilde{\omega}))) \in \mathbb{R}$$

are $\Sigma_{\tilde{\mathbb{F}}}$-measurable, i.e., $\tilde{\mathbb{F}}$-nonanticipative for fixed $l \in \mathcal{D}_1^k$, $\varphi\in\mathcal{D}_r^k$, $\psi\in\mathcal{D}_{r\times r}^k$, $u\in\mathbb{R}^r$, and $v\in\mathbb{R}^{r\times r}$. Then, by the countability of $\mathcal{D}_1^k\times\mathcal{D}_r^k$ and $\mathcal{D}_1^k\times\mathcal{D}_{r\times r}^k$, the functions

$$[0,T]\times\tilde{\Omega} \ni (t,\tilde{\omega}) \to \sup_{(l,\varphi)\in\mathcal{D}_1^k\times\mathcal{D}_r^k} \text{dist}(\tilde{\alpha}_t^k(l,\varphi)(\tilde{\omega}), \Phi(\varphi,u)(\tilde{x}_{t\wedge\tilde{\tau}_k}(\tilde{\omega}))) \in \mathbb{R}$$

and

$$[0,T]\times\tilde{\Omega} \ni (t,\tilde{\omega}) \to \sup_{(l,\psi)\in\mathcal{D}_1^k\times\mathcal{D}_{r\times r}^k} \text{dist}(\tilde{\beta}_t^k(l,\psi)(\tilde{\omega}), \Psi(\psi,v)(\tilde{x}_{t\wedge\tilde{\tau}_k}(\tilde{\omega}))) \in \mathbb{R}$$

are also $\Sigma_{\tilde{\mathbb{F}}}$-measurable for fixed $u\in\mathbb{R}^d$ and $v\in\mathbb{R}^{n\times n}$. Hence it follows that processes $(\mathcal{K}_t^k)_{0\le t\le T}$ and $(\mathcal{Q}_t^k)_{0\le t\le T}$ are $\tilde{\mathbb{F}}$-nonanticipative. Therefore, by virtue of the Kuratowski and Ryll –Nardzewski measurable selection theorem, for every $k\ge 1$, there are $\tilde{\mathbb{F}}$-nonanticipative selectors $\tilde{f}^k = (\tilde{f}_t^k)_{0\le t\le T}$ and $\tilde{\sigma}^k = (\tilde{\sigma}_t^k)_{0\le t\le T}$ of $(\mathcal{K}_t^k)_{0\le t\le T}$ and $(\mathcal{Q}_t^k)_{0\le t\le T}$, respectively. By the definitions of $\mathcal{K}_t^k(\tilde{\omega})$ and $\mathcal{Q}_t^k(\tilde{\omega})$, it follows that $\tilde{f}^k \in S_{\widetilde{\mathbb{F}}}(\text{co}\, F\circ(\widetilde{x}\circ\widetilde{\tau}_k))$ and $\tilde{\sigma}^k \in S_{\widetilde{\mathbb{F}}}(D(G\circ(\widetilde{x}\circ\widetilde{\tau}_k)))$ are such that

$$\sup_{(l,\varphi)\in\mathcal{C}_1^k\times\mathcal{C}_r^k} \text{dist}(\tilde{\alpha}_t^k(l,\varphi)(\tilde{\omega}), \Phi(\varphi,\tilde{f}_t^k(\tilde{\omega}))(\tilde{x}_{t\wedge\tilde{\tau}_k}(\tilde{\omega}))) = 0$$

and

$$\sup_{(l,\psi)\in\mathcal{C}_1^k\times\mathcal{C}_{r\times r}^k} \text{dist}(\tilde{\beta}_t^k(l,\psi)(\tilde{\omega}), \Psi(\psi,\tilde{\sigma}_t^k(\tilde{\omega}))(\tilde{x}_{t\wedge\tilde{\tau}_k}(\tilde{\omega}))) = 0$$

a.e. on $[0,T]\times\tilde{\Omega}$. By virtue of Lemma 4.1, for every $k\ge 1$, there exists $\tilde{g}^k \in S_{\widetilde{\mathbb{F}}}(G\circ(\widetilde{x}\circ\widetilde{\tau}_k))$ such that $\tilde{\sigma}^k = \tilde{g}^k\cdot(\tilde{g}^k)^*$. Hence, by the properties of $(\tilde{\alpha}_t^k(l,\varphi))_{0\le t\le T}$ and $(\tilde{\beta}_t^k(l,\psi))_{0\le t\le T}$ and the definition of $\mathbb{L}_{fg}^x(\varphi,\psi)_t$, it follows that (4.1) is satisfied.

□

Theorem 4.1. *Let $F : [0,T] \times \mathbb{R}^r \to Cl(\mathbb{R}^r)$ and $G : [0,T] \times \mathbb{R}^r \to Cl(\mathbb{R}^{r\times m})$ be measurable and uniformly square integrably bounded, and let G be diagonally convex. Assume that $F(t,\cdot)$ and $G(t,\cdot)$ are continuous for fixed $t \in [0,T]$ and let $(x_t)_{0\le t\le T}$ and $\tilde{x} = (\tilde{x}_t)_{0\le t\le T}$ be continuous r-dimensional $\mathbb{F}$- and $\tilde{\mathbb{F}}$-adapted processes on filtered probability spaces $(\Omega, \mathcal{F}, \mathbb{F}, P)$ and $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{F}}, \tilde{\mathcal{P}})$, respectively, such that $Px^{-1} = P\tilde{x}^{-1}$. For every $f \in S_{\mathbb{F}}(\mathrm{co}\, F \circ x)$ and $g \in S_{\mathbb{F}}(G \circ x)$, there are $\widetilde{f} \in S_{\widetilde{\mathbb{F}}}(\mathrm{co}\, F \circ \widetilde{x})$ and $\widetilde{g} \in S_{\widetilde{\mathbb{F}}}(G \circ \widetilde{x})$ such that*

$$E\int_s^t l(x_s)\mathbb{L}^x_{fg}(\varphi,\psi)_\tau \mathrm{d}\tau = \tilde{E}\int_s^t l(\tilde{x}_s)\mathbb{L}^{\tilde{x}}_{\tilde{f}\tilde{g}}(\varphi,\psi)_\tau \mathrm{d}\tau \tag{4.2}$$

for every $0 \le s < t \le T$, $l \in \mathcal{C}_1$, $\varphi \in \mathcal{C}_r$, and $\psi \in \mathcal{C}_{r\times r}$.

Proof. Let $(\tau_k)_{k=1}^\infty$ and $(\tilde{\tau}_k)_{k=1}^\infty$ be as in Lemma 4.5. We have $0 < \tau_1 < \tau_2, \ldots \le T$, $0 < \tilde{\tau}_1 < \tilde{\tau}_2, \ldots \le T$, $\lim_{k\to\infty} \tau_k = T$, and $\lim_{k\to\infty}\tilde{\tau}_k = T$ with $(P.1)$ and $(\tilde{P}.1)$, respectively. Denote by l_k, φ_k and ψ_k the restrictions of $l \in \mathcal{C}_1$, $\varphi \in \mathcal{C}_r$, and $\psi \in \mathcal{C}_{r\times r}$ to the set $K_k = \{x \in \mathbb{R}^r : |x| \le k\}$ for $k = 1, 2, \ldots$, respectively. For every $k \ge 1$, we have $l_k(x_{s\wedge\tau_k}) = l(x_{s\wedge\tau_k})$, $l_k(\tilde{x}_{s\wedge\tilde{\tau}_k}) = l(\tilde{x}_{s\wedge\tilde{\tau}_k})$,

$$\int_{s\wedge\tau_k}^{t\wedge\tau_k} \mathbb{L}^x_{fg}(\varphi_k,\psi_k)_\tau \mathrm{d}\tau = \int_{s\wedge\tau_k}^{t\wedge\tau_k} \mathbb{L}^x_{fg}(\varphi,\psi)_\tau \mathrm{d}\tau,$$

and

$$\int_{s\wedge\tilde{\tau}_k}^{t\wedge\tilde{\tau}_k} \mathbb{L}^{\tilde{x}}_{\tilde{f}^k\tilde{g}^k}(\varphi_k,\psi_k)_\tau \mathrm{d}\tau = \int_{s\wedge\tilde{\tau}_k}^{t\wedge\tilde{\tau}_k} \mathbb{L}^{\tilde{x}}_{\tilde{f}^k\tilde{g}^k}(\varphi,\psi)_\tau \mathrm{d}\tau$$

with $(P.1)$ and $(\tilde{P}.1)$, respectively, where $S_{\widetilde{\mathbb{F}}}(\mathrm{co}\, F \circ (\widetilde{x} \circ \widetilde{\tau}_k))$ and $S_{\widetilde{\mathbb{F}}}(G \circ (\widetilde{x} \circ \widetilde{\tau}_k))$ are such that the conditions of Lemma 4.6 are satisfied for $k = 1, 2, \ldots$. Put

$$\tilde{f} = \mathbb{1}_{\{0\}}\tilde{f}_0^1 + \mathbb{1}_{(0,\tilde{\tau}_1]}\tilde{f}^1 + \mathbb{1}_{(\tilde{\tau}_1,\tilde{\tau}_2]}\tilde{f}^2 + \cdots \quad \text{and} \quad \tilde{g} = \mathbb{1}_{\{0\}}\tilde{g}_0^1 + \mathbb{1}_{(0,\tilde{\tau}_1]}\tilde{g}^1 + \mathbb{1}_{(\tilde{\tau}_1,\tilde{\tau}_2]}\tilde{g}^2 + \cdots.$$

Let us observe that for every $k \ge 1$ and $t \in (\widetilde{\tau}_{k-1}, \widetilde{\tau}_k]$, one has $\widetilde{f}_t = \widetilde{f}_t^k \in \mathrm{co}\, F(t, \widetilde{x}_{t\wedge\widetilde{\tau}_k})$ and $\widetilde{g}_t = \widetilde{g}_t^k \in G(t, \widetilde{x}_{t\wedge\widetilde{\tau}_k})$. Then $\widetilde{f} \in S_{\widetilde{\mathbb{F}}}(\mathrm{co}\, F \circ \widetilde{x})$ and $\widetilde{g} \in S_{\widetilde{\mathbb{F}}}(G \circ \widetilde{x})$. Furthermore, for every $k \ge 1$, one has

$$E\int_{s\wedge\tau_k}^{t\wedge\tau_k} l(x_{s\wedge\tau_k})\mathbb{L}^x_{fg}(\varphi,\psi)_\tau \mathrm{d}\tau = \tilde{E}\int_{s\wedge\tilde{\tau}_k}^{t\wedge\tilde{\tau}_k} l(\tilde{x}_{s\wedge\tilde{\tau}_k})\mathbb{L}^{\tilde{x}}_{\tilde{f}\tilde{g}}(\varphi,\psi)_\tau \mathrm{d}\tau$$

for $0 \le s < t \le T$, $l \in \mathcal{C}_1$, $\varphi \in \mathcal{C}_r$, and $\psi \in \mathcal{C}_{r\times r}$. Passing to the limit $k \to \infty$, we obtain

$$E\int_s^t l(x_s)\mathbb{L}^x_{fg}(\varphi,\psi)_\tau \mathrm{d}\tau = \tilde{E}\int_s^t l(\tilde{x}_s)\mathbb{L}^{\tilde{x}}_{\tilde{f}\tilde{g}}(\varphi,\psi)_\tau \mathrm{d}\tau$$

for $0 \le s < t \le T$, $l \in \mathcal{C}_1$, $\varphi \in \mathcal{C}_r$, and $\psi \in \mathcal{C}_{r\times r}$. □

Theorem 4.2. *Assume that the assumptions of Theorem 4.1 are satisfied. For every* $f \in S_{\mathbb{F}}(\text{co}\, F \circ x)$ *and* $g \in S_{\mathbb{F}}(G \circ x)$, *there are* $\widetilde{f} \in S_{\widetilde{\mathbb{F}}}(\text{co}\, F \circ \widetilde{x})$ *and* $\widetilde{g} \in S_{\widetilde{\mathbb{F}}}(G \circ \widetilde{x})$ *such that*

$$E \int_s^t l(x_s)(\mathbb{L}^x_{fg} h)_\tau \mathrm{d}\tau = \tilde{E} \int_s^t l(\tilde{x}_s)(\mathbb{L}^{\tilde{x}}_{\tilde{f}\tilde{g}} h)_\tau \mathrm{d}\tau \tag{4.3}$$

for every $0 \le s < t \le T$, $l \in \mathcal{C}_1$, *and* $h \in C_b^2(\mathbb{R}^r)$.

Proof. The proof follows immediately from Theorem 4.1. Indeed, for every $f \in S_{\mathbb{F}}(\text{co}\, F \circ x)$ and $g \in (G \circ x)$, there are $\widetilde{f} \in S_{\widetilde{\mathbb{F}}}(\text{co}\, F \circ \widetilde{x})$ and $\widetilde{g} \in S_{\widetilde{\mathbb{F}}}(G \circ \widetilde{x})$ such that (4.1) is satisfied for every $0 \le s < t \le T$, $l \in \mathcal{C}_1$, $\varphi \in \mathcal{C}_r$, and $\psi \in \mathcal{C}_{r\times r}$. In particular, (4.1) is also satisfied for $0 \le s < t \le T$, $\varphi(h) \in \mathcal{C}_r$, and $\psi(h) \in \mathcal{C}_{r\times r}$ for every $l \in \mathcal{C}_1$ and $h \in C_b^2(\mathbb{R}^r)$. But for every $h \in C_b^2(\mathbb{R}^r)$, we have

$$\mathbb{L}^x_{fg}(\varphi(h), \psi(h))_t = (\mathbb{L}^x_{fg} h)_t \quad \text{and} \quad \mathbb{L}^{\tilde{x}}_{\tilde{f}\tilde{g}}(\varphi(h), \psi(h))_t = (\mathbb{L}^{\tilde{x}}_{\tilde{f}\tilde{g}} h)_t$$

for $t \in [0, T]$. Thus (4.2) is satisfied for every $h \in C_b^2(\mathbb{R}^r)$. □

5 Notes and Remarks

The results dealing with stochastic integrals are based on Kisielewicz [55] and [62]. In particular, the first definitions and properties of set-valued stochastic integrals, called in this book set-valued functional stochastic integrals, were introduced by Hiai [38] and Kisielewicz [51]. Later, they were extended in [55] to a more general case by considering set-valued stochastic functional integrals with respect to Poisson measures. Some further generalizations of the results contained in [51] and [55] are given by Michta [76, 77] and Motyl [81, 84]. The first results dealing with set-valued stochastic integrals, defined as certain set-valued random variables, are due to Bocşan [22]. Unfortunately, such set-valued stochastic integrals are not applicable to stochastic inclusions. The set-valued stochastic integrals introduced by Hiai [38] and Kisielewicz [51, 55] are understood as certain subsets of square integrable random variables. This is a natural approach, because the original Itô integral is defined as a square integrable random variable. But the multivalued analytic methods require that one define set-valued stochastic integrals to be certain set-valued random variables. Therefore, the question of the existence of set-valued random variables having subtrajectory integrals equal to given set-valued stochastic integrals has been considered by many authors. Unfortunately, there is no simple solution of this problem, because the set-valued stochastic functional integrals defined in [38] and [51] are not decomposable subsets of the space $\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{R}^n)$. The first results dealing with this problem were given by Kim and Kim [48]. The paper [48] is written by two authors: B.K. Kim and J.H. Kim. Unfortunately, the definition given in [48] is not correct, because the authors assumed that the set-valued integrals defined in [51] are decomposable subsets of $\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{R}^n)$. Later, Jung and Kim [46] corrected their definition of set-valued stochastic integrals by considering in the above procedure a set-valued random variable defined by the closed decomposable hull of the set-valued stochastic integrals defined in [51]. Unfortunately, the proofs of the properties such set-valued integrals presented in [46] are not quite correct. Some remarks on the proofs presented in [46] are given in the author's paper [62], where proofs of some of the properties of set-valued integrals defined in [46] are also given. Probably, a correct theory of set-valued stochastic integrals needs a new idea. Perhaps it has to begin with the definition of set-valued stochastic integrals for some simple set-valued stochastic processes and then extend the results obtained to a more general case. The definition of set-valued stochastic integrals presented in Sect. 2 is taken from Jung and Kim in [46]. Extensions of the above definition of set-valued stochastic integrals to the case of multiprocesses with values in Banach spaces are given in Zhang and in Li et al. [98]. The main problem with applications of such integrals in the theory of stochastic differential inclusions is a lack of conditions for their square integrable boundedness. Results of this chapter come from the author's papers [55, 56] and [59]. In particular, the selection theorems contained in Sect. 1 and Sect. 3 were proved in [55] and [59], respectively. The results contained in Sect. 4 come from [56].

Chapter 4
Stochastic Differential Inclusions

This chapter is devoted to the theory of stochastic differential inclusions. The main results deal with stochastic functional inclusions defined by set-valued functional stochastic integrals. Subsequent sections discuss properties of stochastic and backward stochastic differential inclusions.

1 Stochastic Functional Inclusions

Throughout this section, by $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ we shall denote a complete filtered probability space and assume that $F : [0,T]\times\mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : [0,T]\times\mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ satisfy the following conditions $(\mathcal{H})$:

(i) F and G are measurable,
(ii) F and G are uniformly square integrably bounded.

For set-valued mappings F and G as given above, by stochastic functional inclusions $SFI(F,G)$, $SFI(\overline{F},G)$, and $\overline{SFI}(F,G)$ we mean relations of the form

$$x_t - x_s \in J_{st}[S_{\mathbb{F}}(F\circ x)] + \mathcal{J}_{st}[S_{\mathbb{F}}(G\circ x)],$$

$$x_t - x_s \in \mathrm{cl}_{\mathbb{L}}\{J_{st}[S_{\mathbb{F}}(F\circ x)]\} + \mathcal{J}_{st}[S_{\mathbb{F}}(G\circ x)]\},$$

and

$$x_t - x_s \in \mathrm{cl}_{\mathbb{L}}\{J_{st}[S_{\mathbb{F}}(F\circ x)] + \mathcal{J}_{st}[S_{\mathbb{F}}(G\circ x)]\},$$

respectively, which have to be satisfied for every $0 \le s \le t \le T$ by a system $(\mathcal{P}_{\mathbb{F}}, X, B)$ consisting of a complete filtered probability space $\mathcal{P}_{\mathbb{F}}$ with a filtration $\mathbb{F} = (\mathcal{F}_t)_{0\le t\le T}$ satisfying the usual conditions, an d-dimensional $\mathbb{F}$-adapted continuous stochastic process $X = (X_t)_{0\le t\le T}$, and an m-dimensional $\mathbb{F}$-Brownian motion $B = (B_t)_{0\le t\le T}$ defined on $\mathcal{P}_{\mathbb{F}}$. Such systems $(\mathcal{P}_{\mathbb{F}}, X, B)$ are said to be weak solutions of $SFI(F,G)$, $SFI(\overline{F},G)$, and $\overline{SFI}(F,G)$, respectively. If

M. Kisielewicz, *Stochastic Differential Inclusions and Applications*,
Springer Optimization and Its Applications 80, DOI 10.1007/978-1-4614-6756-4_4,

μ is a given probability measure on $\beta(\mathbb{R}^d)$, then a system $(\mathcal{P}_{\mathbb{F}}, X, B)$ is said to be a weak solution of the initial value problems $SFI(F,G,\mu)$, $SFI(\overline{F},G,\mu)$, or $\overline{SFI}(F,G,\mu)$, respectively, if it satisfies condition (1)–(1), respectively, and $PX_0^{-1} = \mu$. The set of all weak solutions of $SFI(F,G,\mu)$, $SFI(\overline{F},G,\mu)$, and $\overline{SFI}(F,G,\mu)$ (equivalence classes $[(\mathcal{P}_{\mathbb{F}}, X, B)]$ consisting of all systems $(\mathcal{P}'_{\mathbb{F}}, X', B,)$ satisfying (1)–(1), respectively and such that $PX_0^{-1} = P(X'_0)^{-1} = \mu$ and $PX^{-1} = P(X')^{-1}$) are denoted by $\mathcal{X}_\mu(F,G)$, $\mathcal{X}_\mu(\overline{F},G)$, and $\overline{\mathcal{X}}_\mu(F,G)$,, respectively. By $\mathcal{X}^0_\mu(F,G)$ we denote the set of all $[(\mathcal{P}_{\mathbb{F}}, X, B)] \in \mathcal{X}_\mu(F,G)$ with a separable filtered probability space $\mathcal{P}_{\mathbb{F}}$.

Remark 1.1. We can also consider initial value problems for $SFI(F,G)$, $SFI(\overline{F},G)$, and $\overline{SFI}(F,G)$ with an initial condition $x_s = x$ a.s. for a fixed $0 \le s \le T$ and $x \in \mathbb{R}^d$. The sets of all weak solutions for such initial value problems are denoted by $\mathcal{X}_{s,x}(F,G)$, $\mathcal{X}_{s,x}(\overline{F},G)$, and $\overline{\mathcal{X}}_{s,x}(F,G)$, respectively. □

Remark 1.2. The following inclusions follow immediately from Lemma 1.6 of Chap. 3: $\mathcal{X}_\mu(F,G) \subset \mathcal{X}_\mu(\overline{F},G) \subset \overline{\mathcal{X}}_\mu(F,G) \subset \mathcal{X}_\mu(\mathrm{co}\,F, \mathrm{co}\,G)$ for all measurable set-valued functions $F : [0,T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : [0,T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ and probability measure μ on $\beta(\mathbb{R}^d)$. □

Remark 1.3. In what follows, we shall identify weak solutions (equivalence classes $[(\mathcal{P}_{\mathbb{F}}, X, B)]$) of $SFI(F,G)$, $SFI(\overline{F},G)$, and $\overline{SFI}(F,G)$, respectively, with pairs (X,B) of stochastic processes X and B defined on $\mathcal{P}_{\mathbb{F}}$ or with a process X. □

If apart from the set-valued mappings F and G, we are also given a filtered probability space $\mathcal{P}_{\mathbb{F}}$ and an m-dimensional $\mathbb{F}$-Brownian motion on $\mathcal{P}_{\mathbb{F}}$, then a continuous $\mathbb{F}$-adapted process X such that $(\mathcal{P}_{\mathbb{F}}, X, B)$ satisfies (1)–(1), respectively, is called a strong solution for $SFI(F,G)$, $SFI(\overline{F},G)$, and $\overline{SFI}(F,G)$, respectively. For a given $\mathcal{F}_0$-measurable random variable $\xi : \Omega \to \mathbb{R}^d$, the sets of all strong solutions of the above stochastic functional inclusions corresponding to a filtered probability space $\mathcal{P}_{\mathbb{F}}$ and an m-dimensional $\mathbb{F}$-Brownian motion B satisfying an initial condition $X_0 = \xi$ a.s. will be denoted by $\mathcal{S}_\xi(F,G,B,\mathcal{P}_{\mathbb{F}})$, $\mathcal{S}_\xi(\overline{F},G,B,\mathcal{P}_{\mathbb{F}})$, and $\overline{\mathcal{S}}_\xi(F,G,B,\mathcal{P}_{\mathbb{F}})$, respectively. Immediately from Lemma 1.6 of Chap. 3, it follows that $\mathcal{S}_\xi(F,G,B,\mathcal{P}_{\mathbb{F}}) \subset \mathcal{S}_\xi(\overline{F},G,B,\mathcal{P}_{\mathbb{F}}) \subset \overline{\mathcal{S}}_\xi(F,G,B,\mathcal{P}_{\mathbb{F}}) \subset \mathcal{S}_\xi(\overline{\mathrm{co}}\,F, \overline{\mathrm{co}}\,G, B, \mathcal{P}_{\mathbb{F}}) \subset \mathcal{S}(\mathbb{F}, \mathbb{R}^d)$, where $\mathcal{S}(\mathbb{F},\mathbb{R}^d)$ denotes the Banach space of all d-dimensional $\mathbb{F}$-semimartingales $(X_t)_{0\le t\le T}$ on $\mathcal{P}_{\mathbb{F}}$ such that $E[\sup_{0\le t\le T} |X_t|^2] < \infty$. If $\mathcal{P}_{\mathbb{F}}$ is separable, then by virtue of Lemma 1.7 of Chap. 3, one has $\mathcal{S}_\xi(\overline{F},G,B,\mathcal{P}_{\mathbb{F}}) = \mathcal{S}_\xi(\mathrm{co}\,F,G,B,\mathcal{P}_{\mathbb{F}})$.

In what follows, norms of $\mathbb{R}^r$, $\mathbb{L}^2(\Omega,\mathcal{F},\mathbb{R}^r)$, and $\mathbb{L}^2([0,T]\times\Omega, \Sigma_{\mathbb{F}}, \mathbb{R}^r)$ with $r = d$ and $r = d\times m$ will be denoted by $|\cdot|$. It will be clear from the context which of the above normed space is considered.

Theorem 1.1. *Let $B = (B_t)_{0\le t\le T}$ be an m-dimensional $\mathbb{F}$-Brownian motion on $\mathcal{P}_{\mathbb{F}}$, and $\xi : \Omega \to \mathbb{R}^d$ an $\mathcal{F}_0$-measurable random variable. If F and G satisfy conditions $(\mathcal{H})$ and are such that $F(t,\cdot)$ and $G(t,\cdot)$ are Lipschitz continuous with* *a Lipschitz function* $k \in \mathbb{L}^2([0,T],\mathbb{R}^+)$ *such that* $K(\sqrt{T}+1) < 1$, *where* $K = (\int_0^T k^2(t)\mathrm{d}t)^{1/2}$, *then* $\mathcal{S}_\xi(F,G,B,\mathcal{P}_\mathbb{F}) \neq \emptyset$.

Proof. Let $\mathcal{X} = \mathbb{L}^2([0,T]\times\Omega,\Sigma_\mathbb{F},\mathbb{R}^d)\times\mathbb{L}^2([0,T]\times\Omega,\Sigma_\mathbb{F},\mathbb{R}^{d\times m})$ and put $X_t^{fg} = \xi + \int_0^t f_\tau \mathrm{d}\tau + \int_0^t g_\tau \mathrm{d}B_\tau$ a.s. for $0 \le t \le T$ and $(f,g)\in\mathcal{X}$. It is clear that $X^{fg} = (X_t^{fg})_{0\le t\le T} \in \mathcal{S}(\mathbb{F},\mathbb{R}^d)$. Define on $\mathcal{X}$ a set-valued mapping Q by setting $Q(f,g) = S_\mathbb{F}(F\circ X^{fg})\times S_\mathbb{F}(G\circ X^{fg})$ for every $(f,g)\in\mathcal{X}$. It is clear that for every $(f,g)\in\mathcal{X}$, we have $Q(f,g)\in \mathrm{Cl}(\mathcal{X})$.

Let $\lambda(A\times C, B\times D) = \max\{H(A,B),H(C,D)\}$, for $A,B\in \mathrm{Cl}(\mathbb{L}^2([0,T]\times\Omega,\Sigma_\mathbb{F},\mathbb{R}^d)$ and $C,D\in\mathrm{Cl}(\mathbb{L}^2([0,T]\times\Omega,\Sigma_\mathbb{F},\mathbb{R}^{d\times m})$, where for simplicity, H denotes the Hausdorff metric on $\mathrm{Cl}(\mathbb{L}^2([0,T]\times\Omega,\Sigma_\mathbb{F},\mathbb{R}^d)$ and $\mathrm{Cl}(\mathbb{L}^2([0,T]\times\Omega,\Sigma_\mathbb{F},\mathbb{R}^{d\times m})$. It is clear that λ is a metric on $\mathrm{Cl}(\mathcal{X})$. By virtue of Lemma 3.7 of Chap. 2, we have $H(S_\mathbb{F}(F\circ X^{fg}),S_\mathbb{F}(F\circ X^{f'g'}))\le K\|X^{fg}-X^{f'g'}\|_c$ and $H(S_\mathbb{F}(G\circ X^{fg}),S_\mathbb{F}(G\circ X^{f'g'}))\le K\|X^{fg}-X^{f'g'}\|_c$ for every $(f,g),(f',g')\in\mathcal{X}$, where $\|\cdot\|_c$ denotes the norm of $\mathcal{S}(\mathbb{F},\mathbb{R}^d)$ defined by $\|x\|_c^2 = E[\sup_{0\le t\le T}|x_t|^2]$ for $x = (x_t)_{0\le t\le T}\in\mathcal{S}(\mathbb{F},\mathbb{R}^d)$. But

$$
\begin{aligned}
\|X^{fg}-X^{f'g'}\|_c &= \left(E\sup_{0\le t\le T}\left|\int_0^t (f_\tau - f'_\tau)\mathrm{d}\tau + \int_0^t (g_\tau - g'_\tau)\mathrm{d}B_\tau\right|^2\right)^{1/2}\\
&\le \left(E\sup_{0\le t\le T}\left|\int_0^t (f_\tau - f'_\tau)\mathrm{d}\tau\right|^2\right)^{1/2}\\
&\quad + \left(E\sup_{0\le t\le T}\left|\int_0^t (g_\tau - g'_\tau)\mathrm{d}B_\tau\right|^2\right)^{1/2}\\
&\le \sqrt{T}\left(E\sup_{0\le t\le T}\int_0^t |f_\tau - f'_\tau|^2\mathrm{d}\tau\right)^{1/2}\\
&\quad + \left(E\sup_{0\le t\le T}\int_0^t |g_\tau - g'_\tau|^2\mathrm{d}\tau^2\right)^{1/2}\\
&= \sqrt{T}\,|f-f'| + |g-g'| \le (\sqrt{T}+1)\,\|(f,g)-(f',g')\|,
\end{aligned}
$$

where $\|\cdot\|$ denotes the norm on $\mathcal{X}$. Therefore,

$$
\lambda(Q(f,g),Q(f',g')) \le K(\sqrt{T}+1)\,\|(f,g)-(f',g')\|
$$

for every $(f,g),(f',g')\in\mathcal{X}$, which by th Covitz–Nadler fixed-point theorem, implies the existence of $(f,g)\in\mathcal{X}$ such that $(f,g)\in Q(f,g)$. Hence it follows that $\int_s^t f_\tau \mathrm{d}\tau + \int_s^t g_\tau \mathrm{d}B_\tau \in J_{st}[S_{\mathbb{F}}(F \circ X^{fg})] + \mathcal{J}_{st}[S_{\mathbb{F}}(G \circ X^{fg})]$ for every $0 \le s \le t \le T$. This, by the definition of X^{fg}, implies that $X^{fg} \in \mathcal{S}_\xi(F, G, B, \mathcal{P}_{\mathbb{F}})$.

□

Remark 1.4. By an appropriate changing the norm (see Remark 1.1 of Chap. 7) of the space $\mathcal{X}$, the result of Theorem 1.1 can be obtained for every $T > 0$ and $k \in \mathbb{L}^2([0,T], \mathbb{R}^+)$ without the constraint $K(\sqrt{T}+1) < 1$. □

Let us denote by $\Lambda_\xi(F, G, B, \mathcal{P}_{\mathbb{F}})$ the set of all fixed points of the set-valued mapping Q defined in the proof of Theorem 1.1.

Theorem 1.2. *If the assumptions of Theorem 1.1 are satisfied, then*

(i) $\Lambda_\xi(F, G, B, \mathcal{P}_{\mathbb{F}})$ *is a closed subset of* $\mathcal{X}$*;*
(ii) $\mathcal{S}_\xi(\mathrm{co}\, F, \mathrm{co}\, G, B, \mathcal{P}_{\mathbb{F}}) \neq \emptyset$ *if and only if* $\Lambda_\xi(\mathrm{co}\, F, \mathrm{co}\, G, B, \mathcal{P}_{\mathbb{F}}) \neq \emptyset$*;*
(iii) $\mathcal{S}_\xi(\mathrm{co}\, F, \mathrm{co}\, G, B, \mathcal{P}_{\mathbb{F}})$ *is a closed subset of* $\mathcal{S}(\mathbb{F}, \mathbb{R}^d)$*;*
(iv) for every $x \in \overline{\mathcal{S}}_\xi(F, G, B, \mathcal{P}_{\mathbb{F}})$ *and* $\varepsilon > 0$*, there exists* $x^\varepsilon \in \mathcal{S}(\mathbb{F}, \mathbb{R}^d)$ *such that* $\sup_{0 \le t \le t}(E|x - x^\varepsilon|^2)^{1/2} \le \varepsilon$ *and* $\mathrm{dist}(x_t - x_s, J_{st}[S_{\mathbb{F}}(F \circ x)] + \mathcal{J}_{st}[S_{\mathbb{F}}(G \circ x)]) \le \varepsilon$*;*
(v) $\mathcal{X}_\mu(F, G) \neq \emptyset$ *for every probability measure* μ *on* $\beta(\mathbb{R}^d)$.

Proof. (i) The closedness of $\Lambda_\xi(F, G, B, \mathcal{P}_{\mathbb{F}})$ follows immediately from the properties of the set-valued mappings $\mathcal{X} \ni (f, g) \to S_{\mathbb{F}}(F \circ X^{fg}) \subset \mathbb{L}^2([0,T] \times \Omega, \Sigma_{\mathbb{F}}, \mathbb{R}^d)$ and $\mathcal{X} \ni (f, g) \to S_{\mathbb{F}}(G \circ X^{fg}) \subset \mathbb{L}^2([0,T] \times \Omega, \Sigma_{\mathbb{F}}, \mathbb{R}^{d \times m})$. Indeed, if $\{(f^n, g^n)\}_{n=1}^\infty$ is a sequence of $\Lambda_\xi(F, G, B, \mathcal{P}_{\mathbb{F}})$ converging to (f, g), then $\mathrm{dist}(f, S_{\mathbb{F}}(F \circ X^{fg})) = 0$, because

$$\mathrm{dist}(f, S_{\mathbb{F}}(F \circ X^{fg})) \le |f - f^n| + \mathrm{dist}(f^n, S_{\mathbb{F}}(F \circ X^{f^n g^n})) + H(S_{\mathbb{F}}(F \circ X^{fg}), S_{\mathbb{F}}(F \circ X^{f^n g^n})),$$

and by virtue of Lemma 3.7 of Chap. 2, for every $n \ge 1$ one has

$$H(S_{\mathbb{F}}(F \circ X^{fg}), S_{\mathbb{F}}(F \circ X^{f^n g^n}) \le K(\sqrt{T}+1)\|(f, g) - (f^n, g^n)\|.$$

In a similar way, we also get $\mathrm{dist}(g, S_{\mathbb{F}}(G \circ x^{fg})) = 0$. Hence, by the closedness of $S_{\mathbb{F}}(F \circ x^{fg})$ and $S_{\mathbb{F}}(G \circ x^{fg})$, it follows that $(f, g) \in Q(f, g)$. Then $(f, g) \in \Lambda_\xi(F, G, B, \mathcal{P}_{\mathbb{F}})$.

(ii) The implication $\Lambda_\xi(\mathrm{co}\, F, \mathrm{co}\, G, B, \mathcal{P}_{\mathbb{F}}) \neq \emptyset \Rightarrow \mathcal{S}_\xi(\mathrm{co}\, F, \mathrm{co}\, G, B, \mathcal{P}_{\mathbb{F}}) \neq \emptyset$ follows immediately from the proof of Theorem 1.1. The converse implication follows immediately from Theorem 1.5 of Chap. 3.

(iii) Let $(u^n)_{n=1}^\infty$ be a sequence of $\mathcal{S}_\xi(\mathrm{co}\, F, \mathrm{co}\, G, B, \mathcal{P}_{\mathbb{F}})$ converging to $u \in \mathcal{S}(\mathbb{F}, \mathbb{R}^d)$. By Theorem 1.5 of Chap. 3, there exists a sequence $\{(f^n, g^n)\}_{n=1}^\infty$ of $S_{\mathbb{F}}(\mathrm{co}\, F \circ u^n) \times S_{\mathbb{F}}(\mathrm{co}\, G \circ u^n)$ such that $u_t^n = \xi + J_{0t}(f^n) + \mathcal{J}_{0t}(g^n)$ for $n \ge 1$ and $t \in [0,T]$. By Remark 3.1 of Chap. 2, there is a subsequence $\{(f^{n_k}, g^{n_k})\}_{k=1}^\infty$ of $\{(f^n, g^n)\}_{n=1}^\infty$ weakly converging to (f, g), which implies that $J_{0t}(f^{n_k}) + \mathcal{J}_{0t}(g^{n_k}) \to J_{0t}(f) + \mathcal{J}_{0t}(g)$ for every $t \in [0,T]$ in the weak topology of $\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{R}^d)$ as $k \to \infty$. But for every $t \in [0, T]$, a sequence $(u_t^{n_k})_{k=1}^{\infty}$ also converges weakly in $\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{R}^d)$ to u_t. Therefore, $u_t = \xi + J_{0t}(f) + \mathcal{J}_{0t}(g)$ for every $t \in [0, T]$. Then $u \in \mathcal{S}_\xi(\text{co}\, F, \text{co}\, G, B, \mathcal{P}_\mathbb{F})$.

(iv) For every $x \in \overline{\mathcal{S}}_\xi(F, G, B, \mathcal{P}_\mathbb{F})$ and $\varepsilon > 0$, there exists $x^\varepsilon \in \mathcal{S}_\xi(F, G, B, \mathcal{P}_\mathbb{F})$ such that $\sup_{0 \le t \le t}(E|x - x^\varepsilon|^2)^{1/2} \le \varepsilon/[2 + L(\sqrt{T} + 1)]$, where $L = (\int_0^T k^2(t)\mathrm{d}t)^{1/2}$. Similarly as in the proof of Lemma 3.7 of Chap. 2 (see Lemma 1.3 of Chap. 5), it follows that set-valued mappings $\mathcal{S}(\mathbb{F}, \mathbb{R}^d) \ni x \to J_{st}[S_\mathbb{F}(F \circ x)] \subset \mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{R}^d)$ and $\mathcal{S}(\mathbb{F}, \mathbb{R}^d) \ni x \to J_{st}[S_\mathbb{F}(G \circ x)] \subset \mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{R}^d)$ are Lipschitz continuous with Lipschitz constants $\sqrt{T}L$ and L, respectively. Therefore,

$$
\begin{aligned}
&\text{dist}(x_t - x_s, J_{st}[S_\mathbb{F}(F \circ x)] + \mathcal{J}_{st}[S_\mathbb{F}(G \circ x)]) \\
&\quad \le |(x_t - x_s) - (x_t^\varepsilon - x_s^\varepsilon)| \\
&\qquad + \text{dist}(x_t^\varepsilon - x_s^\varepsilon, J_{st}[S_\mathbb{F}(F \circ x^\varepsilon)] + \mathcal{J}_{st}[S_\mathbb{F}(G \circ x^\varepsilon)]) \\
&\qquad + H(J_{st}[S_\mathbb{F}(F \circ x^\varepsilon)], J_{st}[S_\mathbb{F}(F \circ x)]) \\
&\qquad + H(\mathcal{J}_{st}[S_\mathbb{F}(G \circ x^\varepsilon)], \mathcal{J}_{st}[S_\mathbb{F}(G \circ x)]) \\
&\quad \le [2 + L(\sqrt{T} + 1)]\|x - x^\varepsilon\|_c \le \varepsilon.
\end{aligned}
$$

(v) If μ is a given probability measure on $\beta(\mathbb{R}^d)$, then taking an $\mathcal{F}_0$-measurable random variable ξ such that $P\xi^{-1} = \mu$, we obtain the existence of a strong solution X for $SFD(F, G)$ such that $PX_0^{-1} = \mu$, which implies that $\mathcal{X}_\mu(F, G) \neq \emptyset$, because $(\mathcal{P}_\mathbb{F}, X, B) \in \mathcal{X}_\mu(F, G)$. □

We associate now with $SFI(F, G)$ and its weak solution $(\mathcal{P}_\mathbb{F}, x, B)$ a set-valued partial differential operator $\mathbb{L}_{FG}^x$ defined on the space $C_b^2(\mathbb{R}^d)$ of all real-valued continuous bounded functions $h : \mathbb{R}^d \to \mathbb{R}$ having continuous bounded partial derivatives h'_{x_i} and $h''_{x_i x_j}$ for $i, j = 1, 2, \ldots$. Assume that $F : [0, T] \times \mathbb{R}^d \to \text{Cl}(\mathbb{R}^d)$ and $G : [0, T] \times \mathbb{R}^d \to \text{Cl}(\mathbb{R}^{d \times m})$ are measurable and uniformly square integrably bounded such that $F(t, \cdot)$ and $G(t, \cdot)$ are continuous for fixed $t \in [0, T]$. Let G be diagonally convex and $x = (x_t)_{0 \le t \le T}$ a d-dimensional continuous process on a filtered probability space $\mathcal{P}_\mathbb{F} = (\Omega, \mathcal{F}, \mathbb{F}, P)$. For every $(f, g) \in S_\mathbb{F}(\text{co}\, F \circ x) \times S_\mathbb{F}(G \circ x)$, we define a linear operator $\mathbb{L}_{fg}^x : C_b^2(\mathbb{R}^d) \to \mathbb{L}^2([0, T] \times \Omega, \mathbb{R}^d)$ by setting

$$
(\mathbb{L}_{fg}^x h)_t = \sum_{i=1}^n h'_{x_i}(x_t) f_t^i + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n h''_{x_i x_j}(x_t) \sigma_t^{ij}
$$

a.s. for $0 \le t \le T$ and $h \in C_b^2(\mathbb{R}^d)$, where $f_t = (f_t^1, \ldots, f_t^n)$, and $\sigma = g \cdot g^* = (\sigma^{ij})_{n \times m}$. For a process x as given above and sets $A \subset \mathbb{L}^2([0, T] \times \Omega, \Sigma_\mathbb{F}, \mathbb{R}^d)$ and $B \subset \mathbb{L}^2([0, T] \times \Omega, \Sigma_\mathbb{F}, \mathbb{R}^{d \times m})$, by $\mathbb{L}_{AB}^x$ we denote a family $\{\mathbb{L}_{fg}^x : (f, g) \in A \times B\}$.

We say that $\mathbb{L}^x_{fg} \in \mathbb{L}^x_{AB}$ generates on $C^2_b(\mathbb{R}^d)$ a continuous local $\mathbb{F}$-martingale if the process $[(\varphi^x_{fg}h)_t]_{0\le t\le T}$ defined by

$$(\varphi^x_{fg}h)_t = h(x_t) - h(x_0) - \int_0^t (\mathbb{L}^x_{fg}h)_\tau \mathrm{d}\tau \quad \text{with} \quad (P.1) \tag{1.1}$$

for $t \in [0,T]$ is for every $h \in C^2_b(\mathbb{R}^d)$ a continuous local $\mathbb{F}$-martingale on $\mathcal{P}_{\mathbb{F}}$. The family of all $\mathbb{L}^x_{fg} \in \mathbb{L}^x_{AB}$ generating on $C^2_b(\mathbb{R}^d)$ a family of continuous local $\mathbb{F}$-martingales is denoted by $\mathcal{M}^x_{AB}$. In what follows, for the set-valued mappings $F : [0,T]\times\mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : [0,T]\times\mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ as given above, the families $\mathbb{L}^x_{S_{\mathbb{F}}(\mathrm{co}\,F\circ x)S_{\mathbb{F}}(G\circ x)}$ and $\mathcal{M}^x_{S_{\mathbb{F}}(\mathrm{co}\,F\circ x)S_{\mathbb{F}}(G\circ x)}(C^2_b)$ will be denoted by $\mathbb{L}^x_{FG}$ and $\mathcal{M}^x_{FG}$, respectively.

Lemma 1.1. *Assume that $F : [0,T]\times\mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : [0,T]\times\mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ are measurable and uniformly square integrably bounded such that $F(t,\cdot)$ and $G(t,\cdot)$ are continuous for fixed $t \in [0,T]$. Let G be diagonally convex, and let $x = (x_t)_{0\le t\le T}$ and $\tilde{x} = (\tilde{x}_t)_{0\le t\le T}$ be d-dimensional continuous $\mathbb{F}$- and $\tilde{\mathbb{F}}$-adapted processes on $\mathcal{P}_{\mathbb{F}} = (\Omega,\mathcal{F},\mathbb{F},P)$ and $(\tilde{\Omega},\tilde{\mathcal{F}},\tilde{\mathbb{F}},\tilde{P})$, respectively, such that $Px^{-1} = P\tilde{x}^{-1}$. Then $\mathcal{M}^x_{FG} \ne \emptyset$ if and only if $\mathcal{M}^{\tilde{x}}_{FG} \ne \emptyset$.*

Proof. Let $\mathcal{M}^x_{FG} \ne \emptyset$. There exist $f \in S_{\mathbb{F}}(\mathrm{co}\,F\circ x)$ and $g \in S_{\mathbb{F}}(G\circ x)$ such that for every $h \in C^2_b(\mathbb{R}^d)$, the process $[(\varphi^x_h)_t]_{0\le t\le T}$ defined by (1.1) is a continuous local $\mathbb{F}$-martingale on $\mathcal{P}_{\mathbb{F}}$. Therefore, there exists a sequence $(r_k)_{k=1}^\infty$ of $\mathbb{F}$-stopping times on $\mathcal{P}_{\mathbb{F}}$ such that $r_{k-1} \le r_k$ for $k = 1,2,\dots$ with $r_0 = 0$, $\lim_{k\to\infty} r_k = +\infty$ with $(P.1)$ and such that for every $k = 1,2,\dots$, the process $[(\varphi^x_h)_{t\wedge r_k}]_{0\le t\le T}$ is a continuous square integrable $\mathbb{F}$-martingale on $\mathcal{P}_{\mathbb{F}}$. In particular, it follows that for every $0 \le s < t \le T$, one has $E[(\varphi^x_h)_{t\wedge r_k}|\mathcal{F}_s] = (\varphi^x_h)_{s\wedge r_k}$ with $(P.1)$. Thus for every $0 \le s < t \le T$ and $h \in C^2_b(\mathbb{R}^d)$, we have $E\{[(\varphi^x_h)_{t\wedge r_k}) - (\varphi^x_h)_{s\wedge r_k}]|\mathcal{F}_s\} = 0$ with $(P.1)$. Let $l \in \mathcal{C}_1$. By the continuity of $l \in \mathcal{C}_1$ and the $\mathcal{F}_s$-measurability of x_s, the random variable $l(x_s)$ is also $\mathcal{F}_s$-measurable. Therefore, $E\{(l(x_s)[(\varphi^x_h)_{t\wedge r_k}) - (\varphi^x_h)_{s\wedge r_k}]|\mathcal{F}_s\} = 0$ with $(P.1)$ for every $0 \le s < t \le T$, which, in particular, implies that $E(l(x_s)[(\varphi^x_h)_{t\wedge r_k}) - (\varphi^x_h)_{s\wedge r_k}]) = 0$. Thus in the limit $k \to \infty$, we obtain $E\{(l(x_s)[(\varphi^x_h)_t - (\varphi^x_h)_s]) = 0$. Then

$$E\left(l(x_s)[(h(x_t) - h(x_s)]\right) = E\left(l(x_s)\int_s^t (\mathbb{L}^x_{fg}h)_\tau \mathrm{d}\tau\right)$$

for every $0 \le s < t \le T$, $l \in \mathcal{C}_1$, and $h \in C^2_b(\mathbb{R}^d)$. By virtue of Theorem 4.2 of Chap. 3, there exist $\tilde{f} \in S_{\tilde{\mathbb{F}}}(\mathrm{co}F\circ\tilde{x})$ and $\tilde{g} \in S_{\tilde{\mathbb{F}}}(G\circ\tilde{x})$ such that

$$E\int_s^t l(x_s)(\mathbb{L}^x_{fg}h)_\tau \mathrm{d}\tau = \tilde{E}\int_s^t l(\tilde{x}_s)(\mathbb{L}^{\tilde{x}}_{\tilde{f}\tilde{g}}h)_\tau \mathrm{d}\tau$$

for every $0 \le s < t \le T$, $l \in \mathcal{C}_1$, and $h \in C^2_b(\mathbb{R}^r)$. But

$$E\int_s^t l(x_s)(\mathbb{L}^x_{fg}h)_\tau \mathrm{d}\tau = E\left[l(x_s)\int_s^t (\mathbb{L}^x_{fg}h)_\tau \mathrm{d}\tau\right],$$

$$\tilde{E}\int_s^t l(\tilde{x}_s)(\mathbb{L}^{\tilde{x}}_{\tilde{f}\tilde{g}}h)_\tau \mathrm{d}\tau = \tilde{E}\left[l(\tilde{x}_s)\int_s^t (\mathbb{L}^{\tilde{x}}_{\tilde{f}\tilde{g}}h)_\tau \mathrm{d}\tau\right] \quad \text{and}$$

$$E\{l(x_s)[h(x_t)-h(x_s)]\} = \tilde{E}\{l(\tilde{x}_s)[h(\tilde{x}_t)-h(\tilde{x}_s)]\}$$

for every $0 \le s < t \le T$, because $l \in \mathcal{C}_1$ and $h \in C_b^2(\mathbb{R}^d)$ are continuous and $Px^{-1} = P\tilde{x}^{-1}$. Therefore,

$$\tilde{E}\,\{l(\tilde{x}_s)[h(\tilde{x}_t)-h(\tilde{x}_s)]\} = \tilde{E}\left\{l(x_s)\int_s^t (\mathbb{L}^{\tilde{x}}_{\tilde{f}\tilde{g}}h)_\tau \mathrm{d}\tau\right\}$$

for $0 \le s < t \le T$, $l \in \mathcal{C}_1$, and $h \in C_b^2(\mathbb{R}^d)$. Then $\tilde{E}\{l(\tilde{x}_s)[(\varphi_h^{\tilde{x}})_t - (\varphi_h^{\tilde{x}})_s]\} = 0$, which, in particular, implies that $\tilde{E}[l(\tilde{x}_s)\cdot E\{[(\varphi_n^{\tilde{x}})_t - (\varphi_n^{\tilde{x}})_s]|\tilde{\mathcal{F}}_s\}] = 0$ for $0 \le s < t \le T$, $l \in C_1$, and $h \in C_b^2(\mathbb{R}^d)$. By the monotone class theorem, it follows that the above equality is also true for every measurable bounded function $l : \mathbb{R}^d \to \mathbb{R}$. Taking in particular l such that $l(\tilde{x}_s) = \tilde{E}\{[(\varphi_h^{\tilde{x}})_t - (\varphi_h^{\tilde{x}})_s]|\tilde{\mathcal{F}}_s\}$ with $(\tilde{P}.1)$, we get $\tilde{E}|\tilde{E}\{[(\varphi_n^{\tilde{x}})_t - (\varphi_n^{\tilde{x}})_s]|\tilde{\mathcal{F}}_s\}|^2 = 0$ for $0 \le s < t \le T$ and $h \in C_b^2(\mathbb{R}^d)$. Therefore, $\tilde{E}\{[(\varphi_n^{\tilde{x}})_t - (\varphi_n^{\tilde{x}})_s]|\tilde{\mathcal{F}}_s\} = 0$ with $(\tilde{P}.1)$ for every $0 \le s < t \le T$ and $h \in C_b^2(\mathbb{R}^d)$. Then $\mathbb{L}^{\tilde{x}}_{\tilde{f}\tilde{g}} \in \mathcal{M}^{\tilde{x}}_{FG}(C_b^2)$. In a similar way, we can verify that $\mathcal{M}^{\tilde{x}}_{FG} \ne \emptyset$ implies that $\mathcal{M}^x_{FG} \ne \emptyset$. □

Lemma 1.2. *Assume that $F : [0,T]\times\mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : [0,T]\times\mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ are measurable and uniformly square integrably bounded such that $F(t,\cdot)$ and $G(t,\cdot)$ are continuous for fixed $t \in [0,T]$. Let G be diagonally convex and let $(x_t)_{0\le t\le T}$ and $(x_t^k)_{0\le t\le T}$ be d-dimensional continuous stochastic processes on $(\Omega,\mathcal{F},\mathbb{F},P)$ for every $k = 1,2,\ldots$ such that $\lim_{k\to\infty} P(\{\sup_{0\le t\le T}|x_t - x^k| > \varepsilon\}) = 0$ for every $\varepsilon > 0$ and $\mathcal{M}^{x^k}_{FG} \ne \emptyset$ for every $k = 1,2,\ldots$. Then $\mathcal{M}^x_{FG} \ne \emptyset$.*

Proof. Let $f^k \in S_{\mathbb{F}}(\mathrm{co}F\circ x^k)$ and $g^k \in S_{\mathbb{F}}(G\circ x^k)$ be such that $\mathbb{L}^{x^k}_{f_k g_k} \in \mathcal{M}^{x^k}_{FG}$ for every $k = 1,2,\ldots$. Let $(x^{k_r})_{r=1}^\infty$ be a subsequence of $(x^k)_{k=1}^\infty$ such that $\lim_{r\to\infty}\sup_{0\le t\le T}|x_t - x_t^{k_r}| = 0$ with $(P.1)$. By the uniform square integrably boundedness of $F\circ x^k$, it follows that the sequence $(f^{k_r})_{r=1}^\infty$ is weakly compact. Then there exist a d-dimensional $\mathbb{F}$-nonanticipative process f and a subsequence, still denoted by $(f^{k_r})_{r=1}^\infty$, of $(f^{k_r})_{r=1}^\infty$ weakly converging to f. For every $A \in \Sigma_{\mathbb{F}}$ and $k = 1,2,\ldots$, one has

$$\begin{aligned}
&\mathrm{dist}\left(\int_A f_t(\omega)\mathrm{d}t\mathrm{d}P, \int_A \mathrm{co}F(t,x_t(\omega))\mathrm{d}t\mathrm{d}P\right)\\
&\le \left|\int_A f_t(\omega)\mathrm{d}t\mathrm{d}P - \int_A f_t^{k_r}\mathrm{d}t\mathrm{d}P\right|\\
&\quad + h\left(\int_A \mathrm{co}F(t,x_t^{k_r}(\omega))\mathrm{d}t\mathrm{d}P, \int_A \mathrm{co}F(t,x_t(\omega))\mathrm{d}t\mathrm{d}P)\right).
\end{aligned}$$

Then $\int_A f_t(\omega)\mathrm{d}t\mathrm{d}P \in \int_A \mathrm{co}F(t,x_t(\omega))\mathrm{d}t\mathrm{d}P$ for every $A \in \Sigma_{\mathbb{F}}$, which implies that $f \in S(\mathrm{co}\,F \circ x)$. Hence, by the properties of the set-valued mapping $\Phi(\varphi,\cdot)$ defined in Sect. 4 of Chap. 3, it follows that

$$\lim_{r\to\infty} E\left(l(x_s^{k_r})\int_s^t \Phi(\varphi(h), f_\tau^{k_r})(x_\tau^{k_r})\mathrm{d}\tau\right) = E\left(l(x_s)\int_s^t \Phi(\varphi(h), f_\tau)(x_\tau)\mathrm{d}\tau\right)$$

for every $0 \le s < t \le T$, $l \in \mathcal{C}^1$, and $h \in C_b^2(\mathbb{R}^d)$. In a similar way, we can verify the existence of $g \in S_{\mathbb{F}}(G \circ x)$ such that

$$\lim_{r\to\infty} E\left(l(x_s^{k_r})\int_s^t \Psi(\psi(h), \sigma_\tau^{k_r})(x_\tau^{k_r})\mathrm{d}\tau\right) = E\left(l(x_s)\int_s^t \Psi(\psi(h), \sigma_\tau)(x_\tau)\mathrm{d}\tau\right)$$

for every $0 \le s < t \le T$, $l \in \mathcal{C}^1$, and $h \in C_b^2(\mathbb{R}^d)$, where $\Psi(\psi,\cdot)$ is defined in Sect. 4 of Chap. 3, $\sigma^{k_r} = g^{k_r}\cdot(g^{k_r})^*$, and $\sigma = g\cdot g^*$. By the definitions of $\mathbb{L}_{fg}^x$ and mappings $\Phi(\varphi,\cdot)$ and $\Psi(\psi,\cdot)$, it follows that

$$\lim_{r\to\infty} E\left(l(x_s^{k_r})\int_s^t (\mathbb{L}_{f^{k_r}g^{k_r}}^{x^{k_r}} h)_\tau\mathrm{d}\tau\right) = E\left(l(x_s)\int_s^t (\mathbb{L}_{fg}^x h)_\tau\right)\mathrm{d}\tau,$$

for every $0 \le s < t \le T$, $l \in \mathcal{C}^1$, and $h \in C_b^2(\mathbb{R}^d)$. But $\mathbb{L}_{f^kg^k}^{x^k} \in \mathcal{M}_{FG}^{x^k}$ for $k = 1,2,\ldots$. Then

$$E\left(l(x_s^{k_r})[h(x_t^{k_r}) - h(x_s^{k_r})]\right) = E\left(l(x_s^{k_r})\int_s^t (\mathbb{L}_{f^{kr}g^{kr}}^{x^{kr}} h)_\tau\mathrm{d}\tau\right)$$

for every $0 \le s < t \le T$, $k = 1,2,\ldots,$ $l \in \mathcal{C}_1$, and $h \in C_b^2(\mathbb{R}^d)$. Passing to the limit as $r \to \infty$, we obtain $E\{l(x_s)[(\varphi_h^x)_t - (\varphi_h^x)_s]\} = 0$ for $0 \le s < t \le T$, $l \in \mathcal{C}_1$, and $h \in C_b^2(\mathbb{R}^d)$. Similarly as in the proof of Lemma 1.1, it follows that $\mathbb{L}_{fg}^x \in \mathcal{M}_{FG}^x$. Then $\mathcal{M}_{FG}^x \neq \emptyset$. □

Remark 1.5. In a similar way, it can be verified that by the assumptions of Lemma 1.2, without the continuity of $F(t,\cdot)$ and $G(t,\cdot)$ for fixed $t \in [0,T]$ the, nonemptiness of $\mathcal{M}_{FG}^{x^k}$ for every $k = 1,2,\ldots$ implies that $\mathcal{M}_{FG}^x \neq \emptyset$. □

Lemma 1.3. *Assume that $F : [0,T]\times\mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : [0,T]\times\mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ are measurable and uniformly square integrably bounded such that $F(t,\cdot)$ and $G(t,\cdot)$ are continuous for fixed $t \in [0,T]$. Let G be diagonally convex and let $(x_t^k)_{0\le t\le T}$ be for every $k = 1,2,\ldots$, a d-dimensional continuous $\mathbb{F}^k$-adapted stochastic process on $(\Omega^k,\mathcal{F}^k,\mathbb{F}^k,P^k)$ such that $\mathcal{M}_{FG}^{x^k} \neq \emptyset$ for every $k = 1,2,\ldots$. Let $\tilde{x}^k = (\tilde{x}_t^k)_{0\le t\le T}$ and $\tilde{x} = (\tilde{x}_t)_{0\le t\le T}$ be for $k = 1,2,\ldots$, continuous d-dimensional $\tilde{\mathbb{F}}$-adapted processes on $(\tilde{\Omega},\tilde{\mathcal{F}},\tilde{\mathbb{F}},\tilde{P})$ such that $P(\tilde{x}^k)^{-1} = P(x^k)^{-1}$ for $k = 1,2,\ldots$ and $\lim_{k\to\infty}\tilde{P}(\{\sup_{0\le t\le T}|\tilde{x}_t - \tilde{x}^k| > \varepsilon\}) = 0$ for every $\varepsilon > 0$. Then $\mathcal{M}_{FG}^{\tilde{x}} \neq \emptyset$.*

Proof. By virtue of Lemma 1.1, one has $\mathcal{M}_{FG}^{\tilde{x}^k} \neq \emptyset$ for every $k = 1, 2, \ldots$, which by Lemma 1.2, implies that $\mathcal{M}_{FG}^{\tilde{x}} \neq \emptyset$. □

Lemma 1.4. *Let* $F : [0, T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ *and* $G : [0, T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ *be measurable and uniformly square integrably bounded. If* $(x_t, B_t)_{0\le t\le T}$ *is a weak solution of* $SFI(\mathrm{co}\, F, G)$ *on a complete probability space* $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ *with a filtration* $\mathbb{F} = (\mathcal{F}_t)_{0\le t\le T}$*, then there is a sequence* $(x^k)_{k=1}^{\infty}$ *of Itô processes* $x^k = (x_t^k)_{0\le t\le T}$ *of the form* $x_t^k = x_0 + \int_0^t f_\tau^k \mathrm{d}\tau + \int_0^t g_\tau^k \mathrm{d}B_\tau$ *a.s. for* $t \in [0, T]$ *with* $f^k \in S_{\mathbb{F}}(\mathrm{co} F \circ x))$ *and* $g^k \in S_{\mathbb{F}}(G \circ x))$ *such that* $\lim_{k\to\infty} P(\{\sup_{0\le t\le T} |x_t - x_t^k| > \varepsilon\}) = 0$ *for every* $\varepsilon > 0$.

Proof. By virtue of Theorem 1.4 of Chap. 3, there are sequences $(f^k)_{k=1}^{\infty}$ and $(g^k)_{k=1}^{\infty}$ of $S_{\mathbb{F}}(\mathrm{co}\, F \circ x)$ and $S_{\mathbb{F}}(G \circ x)$, respectively, such that $\sup_{0\le t\le T} E|x_t - x_t^k|^2 \to 0$ as $k \to \infty$, where $x_t^k = x_0 + \int_0^t f_\tau^k \mathrm{d}\tau + \int_0^t g_\tau^k \mathrm{d}B_\tau$ with $(P.1)$ for $t \in [0, T]$ and $k = 1, 2, \ldots$. By Theorem 3.4 of Chap. 1, we can assume that $(x_t)_{0\le t\le T}$ and $(x_t^k)_{0\le t\le T}$ are continuous for $k \ge 1$ because for $\alpha = 2r$, and $\beta = r$ with $r \ge 1$, there is a positive number M such that $E|x_t - x_s|^\alpha \le M|t - s|^{1+\beta}$ and $E|x_t^k - x_s^k|^\alpha \le M|t - s|^{1+\beta}$ for every $0 \le s < t \le T$ and $k = 1, 2, \ldots$. For every $\varepsilon > 0$, $0 \le s < t \le T$, and $k = 1, 2, \ldots$, we have

$$P(\{|x_t - x_t^k| > \varepsilon\}) \le \frac{1}{\varepsilon^\alpha} E|x_t - x_t^k|^\alpha, \quad P(\{|x_t - x_s| > \varepsilon\}) \le \frac{1}{\varepsilon^\alpha} E|x_t - x_s|^\alpha$$

and

$$P(\{|x_t^k - x_s^k| > \varepsilon\}) \le \frac{1}{\varepsilon^\alpha} E|x_t^k - x_s^k|^\alpha.$$

Then for every $m = 1, 2, \ldots$, there is a positive integer k_m such that

$$\max\,[P(\{|x_{i/2^m} - x_{i/2^m}^k| > 1/2^{ma}\}),$$

$$P(\{|x_{(i+1)/2^m} - x_{i/2^m}| > 1/2^{ma}\}),$$

$$P(\{|x_{(i+1)/2^m}^k - x_{i/2^m}^k| > 1/2^{ma}\})] \le M\frac{2^{m\alpha}}{2^{m(1+\beta)}}$$

for $k \ge k_m$ and $0 \le i \le 2^m T - 1$, where $a > 0$ is such that $a < \beta/\alpha$.

Hence in particular, it follows that

$$\max\left[P\left(\left\{\max_{0\le i\le 2^m T-1} |x_{i/2^m} - x_{i/2^m}^k| > i/2^{ma}\right\}\right),\right.$$

$$P\left(\left\{\max_{0\le i\le 2^m T-1} |x_{(i+1)/2^m} - x_{i/2^m}^k| > i/2^{ma}\right\}\right),$$

$$\left.P\left(\left\{\max_{0\le i\le 2^m T-1} |x_{(i+1)/2^m}^k - x_{i/2^m}^k| > i/2^{ma}\right\}\right)\right]$$

$$\le MT2^{-m(\beta - a\alpha)}$$

for $k \geq k_m$ and $m = 1, 2, \ldots$. For $\varepsilon > 0$ and $\delta > 0$ select $\nu = \nu(\varepsilon, \delta)$ such that $(1 + 2/(2^a - 1))/2^{\nu a} \leq \varepsilon$ and $\sum_{m=\nu}^{\infty} 2^{-m(\beta - a\alpha)} \leq \frac{\delta}{3MT}$. For every $m \geq \nu$ and $k \geq k_m$, one gets

$$P\left(\bigcup_{m=\nu}^{\infty}\left\{\max_{0\leq i\leq 2^m T-1}|x_{i/2^m} - x_{i/2^m}^k| > 1/2^{ma}\right\}\right)$$

$$\leq \delta/\varepsilon,\ P\left(\bigcup_{m=\nu}^{\infty}\left\{\max_{0\leq i\leq 2^m T-1}|x_{(i+1)/2^m} - x_{i/2^m}| > 1/2^{ma}\right\}\right) \leq \delta/\varepsilon$$

$$\text{and } P\left(\bigcup_{m=\nu}^{\infty}\left\{\max_{0\leq i\leq 2^m T-1}|x_{(i+1)/2^m}^k - x_{i/2^m}^k| > 1/2^{ma}\right\}\right) \leq \delta\varepsilon.$$

Let

$$\Omega_\nu^{1,k} = \bigcup_{m=\nu}^{\infty}\left\{\max_{0\leq i\leq 2^m T-1}|x_{i/2} - x_{i/2^m}^k| > 1/2^{ma}\right\},$$

$$\Omega_\nu^{2} = \bigcup_{m=\nu}^{\infty}\left\{\max_{0\leq i\leq 2^m T-1}|x_{(i+1)/2^m} - x_{i/2^m}| > 1/2^{ma}\right\}$$

$$\text{and } \Omega_\nu^{3,k} = \bigcup_{m=\nu}^{\infty}\left\{\max_{0\leq i\leq 2^m T-1}|x_{(i+1)/2^m}^k - x_{i/2^m}^k| > 1/2^{ma}\right\}$$

for $k \geq k_\nu$. Taking $\Omega_\nu^k = \Omega_\nu^{1,k} \cup \Omega_\nu^2 \cup \Omega_\nu^{3,k}$, one obtains $P(\Omega_\nu^k) \leq \delta$ for every $k \geq k_\nu$. By the definition of Ω_ν^k, for every $\omega \notin \Omega_\nu^k$, $k \geq k_\nu$, and $0 \leq i \leq 2^\nu T - 1$, we get

$$|x_{i/2^\nu} - x_{i/2^\nu}^k| \leq \frac{1}{2^{\nu a}},\ |x_{(i+1)/2^\nu} - x_{i/2^\nu}| \leq \frac{1}{2^{\nu a}} \text{ and } |x_{(i+1)/2^\nu}^k - x_{i/2^\nu}^k| \leq \frac{1}{2^{\nu a}}.$$

Let D_T be the set of dyadic numbers of $[0, T]$. For every $t \in D_T \cap [i/2^\nu, (i+1)/2^\nu]$, one has $t = i/2^\nu + \sum_{i=1}^{j} \alpha_l/2^{\nu+1}$ with $\alpha_l \in \{0, 1\}$ for $l = 1, 2, \ldots, j$. For every $k \geq k_\nu$, $\omega \notin \Omega_\nu^k$ and i fixed above, we get

$$\begin{aligned}
|x_t - x_t^k| &\leq |x_t - x_{i/2^\nu}| + |x_{i/2^\nu} - x_{i/2^\nu}^k| + |x_{i/2^\nu}^k - x_t^k| \\
&\leq \sum_{r=1}^{j} |x_{i/2^\nu + \sum_{l=1}^{r}\alpha_l/2^{\nu+l}} - x_{i/2^\nu + \sum_{l=1}^{r-1}\alpha_l/2^{\nu+l}}| + |x_{i/2^\nu} - x_{i/2^\nu}^k| \\
&\quad + \sum_{r=1}^{j} |x_{i/2^\nu + \sum_{l=1}^{r}\alpha_l/2^{\nu+l}} - x_{i/2^\nu + \sum_{l=1}^{r}\alpha_l/2^{\nu+l}}| \leq 2\sum_{r=1}^{j} 1/2^{(\nu+r)a} + \frac{1}{2^{\nu a}}
\end{aligned}$$

$$\leq 2\sum_{r=1}^{\infty} 1/2^{(\nu+r)a} + \frac{1}{2^{\nu a}} = \frac{2}{(2^a-1)2^{\nu a}} + \frac{1}{2^{\nu a}}$$
$$= (1 + 2/(2^a-1)2^{\nu a}) \leq \varepsilon.$$

But D_T is dense in $[0,T]$, and $(x_t)_{0\leq t\leq T}$ and $(x_t^k)_{0\leq t\leq T}$ are continuous. Then for every $k \geq k_\nu$ and $\omega \notin \Omega_\nu^k$, one obtains $|x_t(\omega) - x_t^k(\omega)| \leq \varepsilon$ for $t \in [0,T]$, which implies that

$$P(\{\max_{0\leq t\leq T} |x_t - x_t^k| > \varepsilon\}) \leq P(\Omega_\nu^k) < \delta$$

for every $k \geq k_\nu$. Thus for every $\varepsilon > 0$ and $\delta > 0$, there is $k_\nu = k_{\nu(\varepsilon,\delta)}$ such that

$$P\left(\left\{\sup_{0\leq t\leq T} |x_t - x_t^k| > \varepsilon\right\}\right) \leq \delta$$

for $k \geq k_\nu$, i.e., $\lim_{k\to\infty} P(\{\sup_{0\leq t\leq T} |x_t - x_t^k| > \varepsilon\}) = 0$ for every $\varepsilon > 0$. □

Theorem 1.3. *Let $F : [0,T]\times\mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : [0,T]\times\mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ be measurable and uniformly square integrably bounded and let G be diagonally convex. For every probability measure μ on $\beta(\mathbb{R}^d)$, the problem $SFI(\mathrm{co}\,F, G, \mu)$ possesses at least one weak solution with an initial distribution μ if and only if there exist a filtered probability space $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ with a filtration $\mathbb{F} = (\mathcal{F}_t)_{0\leq t\leq T}$ and a d-dimensional continuous $\mathbb{F}$-adaptive stochastic process $x = (x_t)_{0\leq t\leq T}$ on $\mathcal{P}_{\mathbb{F}}$ such that $Px_0^{-1} = \mu$ and $\mathcal{M}_{FG}^x \neq \emptyset$.*

Proof. ($\Rightarrow$) Let $(\mathcal{P}_{\mathbb{F}}, x, B)$ be a weak solution of $SFI(\mathrm{co}\,F, G, \mu)$ with $x = (x_t)_{0\leq t\leq T}$. By virtue of Lemma 1.4, there exist sequences $(f^k)_{k=1}^{\infty}$ and $(g^k)_{k=1}^{\infty}$ of $S_{\mathbb{F}}(\mathrm{co}\,F\circ x))$ and $S_{\mathbb{F}}(G\circ x)$, respectively, such that the sequence $(x^k)_{k=1}^{\infty}$ of continuous $\mathbb{F}$-adapted processes $x^k = (x_t^k)_{0\leq t\leq T}$ defined by $x_t^k = x_0 + \int_0^t f_\tau^k + \int_0^t g_\tau^k \mathrm{d}B_\tau$ a.s. for $0\leq t\leq T$ is such that $\lim_{k\to\infty} P(\{\sup_{0\leq t\leq T} |x_t - x_t^k| > \varepsilon\}) = 0$ for every $\varepsilon > 0$. By Itô's formula, for every $h \in C_b^2(\mathbb{R}^d)$ and $k = 1, 2\ldots$ one obtains

$$h(x_t^k) - h(x_0^k) - \int_0^t (\mathbb{L}_{f^k g^k}^{x^k} h)_\tau \mathrm{d}\tau = \sum_{i=1}^n\sum_{j=1}^n \int_0^t h'_{x_i}(x_\tau^k)(g^k)_\tau^{ij}\,\mathrm{d}B_\tau^j$$

with $(P.1)$ for $t\in[0,T]$, where $B_t = (B_t^1,\ldots,B_t^m)^*$ and $g_t^k = [(g^k)_t^{ij}]_{d\times m}$ for $0\leq t\leq T$. By the definition of $[\varphi_{f^k g^k}^{x^k} h]_t$, the above equality can be written in the form

$$[\varphi_{f^k g^k}^{x^k} h]_t = \sum_{i=1}^n\sum_{j=1}^n \int_0^t h'_{x_i}(x_\tau^k)(g^k)_\tau^{ij}\,\mathrm{d}B_\tau^j$$

with $(P.1)$ for $t \in [0,T]$. Hence, by the properties of Itô integrals, it follows that $[(\varphi^{x^k}_{f^k g^k} h)_t]_{0\le t\le T}$ is a continuous local $\mathbb{F}$-martingale on $\mathcal{P}_{\mathbb{F}}$ for every $k = 1,2,\dots$ and $h \in C_b^2(\mathbb{R}^d)$. Therefore, $\mathcal{M}^{x^k}_{FG} \neq \emptyset$ for $k = 1,2,\dots$, which by Remark 1.5 implies that $\mathcal{M}^x_{FG} \neq \emptyset$.

($\Leftarrow$) Let $\mathcal{P}_{\mathbb{F}} = (\Omega,\mathcal{F},\mathbb{F},P)$ be a filtered probability space with a filtration $\mathbb{F} = (\mathcal{F}_t)_{0\le t\le T}$ and $(x_t)_{0\le t\le T}$ a d-dimensional continuous $\mathbb{F}$-adapted process on $\mathcal{P}_{\mathbb{F}}$ such that $x_0^{-1} = \mu$ and $\mathcal{M}^x_{FG} \neq \emptyset$. Then there exist $f \in S_{\mathbb{F}}(\mathrm{co}F \circ x)$ and $g \in S_{\mathbb{F}}(G \circ x)$ such that $\mathbb{L}^x_{fg} \in \mathcal{M}^x_{FG}$. Let $(\tau_k)_{k=1}^\infty$ be a sequence of stopping times $\tau_k = \inf\{t \in [0,T] : x_t \notin K_k\}$, where $K_k = \{x \in \mathbb{R}^d : |x| \le k\}$ for $k = 1,2,\dots$. Select now, in particular, $h_i \in C_b^2(\mathbb{R}^d)$ such that $h_i(x) = x_i$ for $x \in K_k$, where $x = (x^1,\dots,x^n)$. For such $h_i \in C_b^2(\mathbb{R}^d)$, we have

$$\int_0^{t\wedge\tau_k} (\mathbb{L}^x_{fg} h_i)_\tau \mathrm{d}\tau = \int_0^{t\wedge\tau_k} f^i_\tau \mathrm{d}\tau \ \text{ and hence } \ (\varphi^x_{h_i})_{t\wedge\tau_k} = x^i_{t\wedge\tau_l} - x^i_0 - \int_0^{t\wedge\tau_k} f^i_\tau \mathrm{d}\tau$$

a.s. for $k \ge 1$ and $i = 1,2,\dots,d$ and $t \in [0,T]$. But $\mathbb{L}^x_{fg} \in \mathcal{M}^x_{FG}(C_b^2)$. Then $[(\varphi^x_{h_i})_{t\wedge\tau_l}]_{0\le t\le T}$ is for every $i = 1,\dots,d$ and $k = 1,2,\dots$ a continuous local $\mathbb{F}$-martingale on $\mathcal{P}_{\mathbb{F}}$. Let $M^i_t = (\varphi^x_{h_i})_t$ for $i = 1,\dots,d$ and $t \in [0,T]$. Taking, in particular, $h_{ij} \in C_b^2(\mathbb{R}^d)$ such that $h_{ij}(x) = x^i x^j$ for $x \in K_k$ and $i,j = 1,2,\dots,d$, we obtain a family $(M^{ij}_t)_{0\le t\le T}$ for $i,j = 1,\dots,d$ of continuous local $\mathbb{F}$-martingales on $\mathcal{P}_{\mathbb{F}}$ such that

$$M^{ij}_t = x^i_t x^j_t - x^i_0 x^j_0 - \int_0^t [x^i_\tau f^j_\tau + x^j_\tau f^i_\tau x_\tau) + \sigma^{ij}_\tau]\mathrm{d}\tau$$

a.s. for $i,j = 1,2,\dots,n$ and $t \in [0,T]$, where $\sigma = g \cdot g^*$. Let $\sigma = (\sigma^{ij})_{d\times d}$. Similarly as in the proof of Theorem 9.1 of Chap. 1, it follows that

$$\langle M^i, M^j\rangle_t = \int_0^t \sigma^{ij}_\tau \mathrm{d}\tau$$

a.s. for $i,j = 1,2,\dots,d$ and $t \in [0,T]$, which similarly as in the proof of Theorem 9.1 of Chap. 1, implies that there exist a standard extension $\hat{\mathcal{P}}_{\hat{\mathbb{F}}} = (\hat\Omega,\hat{\mathcal{F}},\hat{\mathbb{F}},\hat P)$ of $(\Omega,\mathcal{F},\mathbb{F},P)$ and an m-dimensional $\hat{\mathbb{F}}$-Brownian motion $\hat B = (\hat B_t)_{0\le t\le T}$ on $(\hat\Omega,\hat{\mathcal{F}},\hat{\mathbb{F}},\hat P)$ such that

$$M^i_t = \sum_{j=1}^m \int_0^t \hat g^{ij}_\tau \mathrm{d}\hat B^j_\tau$$

$\hat P$-a.s. for $i = 1,2\dots,d$ and $t \in [0,T]$, with $\hat g_t(\hat\omega) = g_t(\pi(\hat\omega))$ for $\hat\omega \in \hat\Omega$, where $\pi : \hat\Omega \to \Omega$ is the $(\hat{\mathcal{F}},\mathcal{F})$-measurable mapping described in the definition of the extension of $\mathcal{P}_{\mathbb{F}} = (\Omega,\mathcal{F},\mathbb{F},P)$ because a standard extension $\hat{\mathcal{P}}_{\hat{\mathbb{F}}}$ of $\mathcal{P}_{\mathbb{F}}$ is also an extension of it. Let $\hat x_t(\hat\omega) = x_t(\pi(\hat\omega))$ for $\hat\omega \in \hat\Omega$. For every $A \in \beta_T$, we get $(P\hat x_0^{-1})(A) = \hat P[\hat x_0^{-1}(A)] = \hat P[(x\circ\pi)^{-1}(A)] = (\hat P \circ \pi^{-1})[(x_0^{-1}(A)] = P[x_0^{-1}(A)] = (Px_0^{-1})(A) = \mu(A)$, which implies that $P\hat x_0^{-1} = \mu$. By the definition of M^i_t, it follows that

$$\hat{x}_t^i = \hat{x}_0^i + \int_0^t \hat{f}_\tau^i \mathrm{d}\tau + \sum_{j=1}^m \int_0^t \hat{g}^{ij}(\tau, \hat{x}_\tau)\mathrm{d}\hat{B}_\tau^j$$

$\hat{P}$-a.s. for $i = 1, 2, \ldots, d$ and $t \in [0, T]$, where $\hat{f}_t(\hat{\omega}) = f_t(\pi(\hat{\omega}))$ for $\hat{\omega} \in \hat{\Omega}$. Then

$$\hat{x}_t = \hat{x}_0 + \int_0^t \hat{f}_\tau \mathrm{d}\tau + \int_0^t \hat{g}_\tau \mathrm{d}\hat{B}_\tau$$

$\hat{P}$-a.s. for $0 \le t \le T$. Therefore, $\hat{x}_t - \hat{x}_s \in J_{st}[S_{\hat{\mathbb{F}}}(\mathrm{co}\, F \circ \hat{x})] + \mathcal{J}_{st}[S_{\hat{\mathbb{F}}}(G \circ \hat{x})]$ for every $0 \le s < t \le T$ and $P\hat{x}_0^{-1} = \mu$. Thus $(\hat{\mathcal{P}}_{\hat{\mathbb{F}}}, \hat{x}, \hat{B})$ is a weak solution of $SFI(\mathrm{co}F, G, \mu)$. □

Theorem 1.4. *Let $F : [0, T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : [0, T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d \times m})$ be measurable and uniformly square integrably bounded, and let G be diagonally convex. For every probability measure μ on $\beta(\mathbb{R}^n)$, the problem $SFI(\mathrm{co}\, F, G, \mu)$ possesses a weak solution $(\mathcal{P}_{\mathbb{F}}, x, B)$ with a separable filtered probability space $\mathcal{P}_{\mathbb{F}}$ if and only if there exist a separable filtered probability space $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ with a filtration $\mathbb{F} = (\mathcal{F}_t)_{0 \le t \le T}$ and a d-dimensional continuous $\mathbb{F}$-adaptive stochastic process $x = (x_t)_{0 \le t \le T}$ on $\mathcal{P}_{\mathbb{F}}$ such that $Px_0^{-1} = \mu$ and $\mathcal{M}_{FG}^x \neq \emptyset$.*

Proof. Similarly as of the proof of Theorem 1.3, we can verify that if $(\mathcal{P}_{\mathbb{F}}, x, B)$ is a weak solution of $SFI(\mathrm{co}\, F, G, \mu)$ with a separable filtered probability space $\mathcal{P}_{\mathbb{F}}$, then $\mathcal{M}_{FG}^x \neq \emptyset$. Let $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ be a separable filtered probability space with a filtration $\mathbb{F} = (\mathcal{F}_t)_{0 \le t \le T}$, and $(x_t)_{0 \le t \le T}$ a d-dimensional continuous $\mathbb{F}$-adapted process on $\mathcal{P}_{\mathbb{F}}$ such that $\mathcal{M}_{FG}^x \neq \emptyset$. Then there exist $f \in S_{\mathbb{F}}(\mathrm{co}F \circ x)$ and $g \in S_{\mathbb{F}}(G \circ x)$ such that $\mathbb{L}_{fg}^x \in \mathcal{M}_{FG}^x$. Similarly as in the proof of Theorem 1.3, we can define a local $\mathbb{F}$-martingale $(M_t^i)_{0 \le t \le T}$, on $\mathcal{P}_{\mathbb{F}}$ such that $\langle M^i, M^j \rangle_t = \int_0^t \sigma_\tau^{ij} \mathrm{d}\tau$ with $(P.1)$ for $i, j = 1, \ldots, d$ and $t \in [0, T]$. Therefore, by virtue of Theorem 8.2 of Chap. 1 and Remark 8.2 of Chap. 1, there exist a standard separable extension $\hat{\mathcal{P}}_{\hat{\mathbb{F}}} = (\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{F}}, \hat{P})$ of $(\Omega, \mathcal{F}, \mathbb{F}, P)$ and an $\hat{\mathbb{F}}$-Brownian motion $\hat{B} = (\hat{B}_t^1, \ldots, \hat{B}_t^m)_{0 \le t \le T}$ on $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{F}}, \hat{P})$ such that

$$M_t^i = \sum_{j=1}^m \int_0^t \hat{g}_\tau^{ij} \mathrm{d}\hat{B}^j{}_\tau$$

$\hat{P}$-a.s. for $i = 1, 2, \ldots, d$ and $t \in [0, T]$, where $\hat{x}$ and $\hat{g}$ denote extensions of x and g on $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{F}}, \hat{P})$ defined in the usual way. It is clear that $P\hat{x}_0^{-1} = \mu$. Hence it follows that

$$\hat{x}_t^i = \hat{x}_0^i + \int_0^t \hat{f}_\tau^i \mathrm{d}\tau + \sum_{j=1}^m \int_0^t \hat{g}_\tau^{ij} \mathrm{d}\hat{B}_\tau^j$$

$\hat{P}$-a.s. for $i = 1, 2, \ldots, d$ and $t \in [0, T]$, where $\hat{f}$ denotes an extension of f on $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{F}}, \hat{P})$. Then

$$\hat{x}_t = \hat{x}_0 + \int_0^t \hat{f}_\tau \mathrm{d}\tau + \int_0^t \hat{g}_\tau \mathrm{d}\hat{B}_\tau$$

$\hat{P}$-a.s. for $0 \le t \le T$ with $P\hat{x}_0^{-1} = \mu$. Therefore, $(\hat{\mathcal{P}}_{\hat{\mathbb{F}}}, \hat{x}, \hat{B})$ is a weak solution of $SFI(co\, F, G, \mu)$ with a separable filtered probability space $\hat{\mathcal{P}}_{\hat{\mathbb{F}}}$. □

It follows immediately from Theorem 1.2 that if F and G satisfy the assumptions of Theorem 1.1, then $\mathcal{X}_\mu(F, G) \neq \emptyset$ for every probability measure μ on $\beta(\mathbb{R}^d)$. We shall show that if F and G are convex-valued and G is diagonally convex, then for nonemptiness of $\mathcal{X}_\mu(F, G)$, it is enough to assume that $F(t, \cdot)$ and $G(t, \cdot)$ are continuous instead of Lipschitz continuous.

Theorem 1.5. *Let $F : [0, T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : [0, T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ be measurable, uniformly square integrably bounded, and convex-valued such that $F(t, \cdot)$ and $G(t, \cdot)$ are continuous for a.e. fixed $t \in [0, T]$. If G is diagonally convex, then $\mathcal{X}_\mu(F, G) \neq \emptyset$ for every probability measure μ on $\beta(\mathbb{R}^d)$.*

Proof. Let $\mathcal{P}_\mathbb{F} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ be a complete filtered probability space with a filtration $\mathbb{F} = (\mathcal{F}_t)_{0\le t\le T}$ such that there exists an m-dimensional $\mathbb{F}$-Brownian motion $(B_t)_{0\le t\le T}$ on $\mathcal{P}_\mathbb{F}$. Assume that x_0 is an $\mathcal{F}_0$-measurable random variable such that $Px_0^{-1} = \mu$. By virtue of Lemma 3.8 of Chap. 2, there exist $\beta_T \otimes \beta(\mathbb{R}^d)$-measurable selectors f and g of F and G, respectively, such that $\int_0^t f(\tau, \cdot)\mathrm{d}\tau$ and $\int_0^t g(\tau\, \cdot)\mathrm{d}\tau$ are continuous on $\mathbb{R}^d$ for every $t \in [0, T]$. Define for every $k = 1, 2, \ldots$ a continuous process $(x_t^k)_{0\le t\le T}$ by setting

$$x_t^k = \begin{cases} x_0 \ \text{ a.s. for } \ -\frac{T}{k} \le t \le 0, \\ x_0 + \int_0^t f(\tau, x^k_{\tau-\frac{T}{k}})\mathrm{d}\tau + \int_0^t g(\tau, x^k_{\tau-\frac{T}{k}})\mathrm{d}B_\tau \\ \text{a.s. for } \ t \in [0, T]. \end{cases} \tag{1.2}$$

It is clear that x^k is continuous and $\mathbb{F}$-adapted for every $k = 1, 2, \ldots,$. it follows immediately from (1.2) that $P(\{|x_0^k| > N\}) = P(\{|x_0| > N\})$ for every $k \ge 1$ and $N \ge 1$. Then $\lim_{N\to\infty} \sup_{k\ge 1} P(\{|x_0^k| > N\}) = \lim_{N\to\infty} P(\{|x_0| > N\}) = 0$. For every λ and $k \ge 1$, we get

$$P\left(\{|x_t^k - x_s^k| > \lambda\}\right) \le P\left(\left\{\left|\int_s^t f(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}\tau\right| > \lambda\right\}\right)$$
$$+P\left(\left\{\left|\int_s^t g(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}B_\tau\right| > \lambda\right\}\right).$$

By Chebyshev's inequality, it follows that

$$P\left(\left\{\left|\int_s^t f(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}\tau\right| > \lambda\right\}\right) \le \frac{1}{\lambda^4} E\left[\left|\int_s^t f(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}\tau\right|^4\right]$$
$$\le \frac{T^2}{\lambda^4}\left(\int_s^t K^2(t)\mathrm{d}t\right)^2,$$

where $K \in \mathbb{L}^2([0,T], \mathbb{R}^+)$ is such that $\max(\|F(t,x)\|, \|G(t,x)\|) \le K(t)$ for a.e. $t \in [0,T]$ and every $x \in \mathbb{R}^d$. Similarly, we obtain

$$P\left(\left\{\left|\int_s^t g(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}B_\tau\right| > \lambda\right\}\right) \le \frac{1}{\lambda^4} E\left[\left|\int_s^t g(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}B_\tau\right|^4\right].$$

By the definition of $\int_s^t g(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}B_\tau$, one has

$$\left|\int_s^t g(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}B_\tau\right| = \max_{1\le i\le d}\left|\sum_{j=1}^m \int_s^t g^{ij}(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}B^j_\tau\right|$$
$$\le \max_{1\le i\le d}\sum_{j=1}^m \left|\int_s^t g^{ij}(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}B^j_\tau\right|$$
$$\le \sum_{j=1}^m \left[\max_{1\le i\le d, 1\mathrm{j}\le m}\left|\int_s^t g^{ij}(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}B^j_\tau\right|\right]$$
$$= m\cdot \max_{1\le i\le d, 1\le j\le m}\left|\int_s^t g^{ij}(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}B^j_\tau\right|.$$

Then

$$\left|\int_s^t g(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}B_\tau\right|^4 \le m^4\cdot \max_{1\le i\le d, 1\le j\le m}\left|\int_s^t g^{ij}(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}B^j_\tau\right|^4.$$

By Itô's formula, we obtain

$$E\left[\left|\int_s^{t\wedge\tau_N} g^{ij}(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}B^j_\tau\right|^4\right]$$
$$= 6\,E\left[\int_s^{t\wedge\tau_N}\left(\left|\int_s^\tau g^{ij}(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}B^j_\tau\right|^2\cdot\left|g^{ij}(\tau, x^k_{\tau-\frac{1}{k}})\right|^2\right)\mathrm{d}\tau\right]$$
$$\le 6\,E\left[\int_s^t\left(\left|\int_s^\tau g^{ij}(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}B^j_\tau\right|^2\cdot K^2(\tau)\right)\mathrm{d}\tau\right]$$

$$= 6\int_s^t \left[K^2(\tau)\cdot E\left|\int_s^\tau g^{ij}(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}B^j_\tau\right|^2\right]\mathrm{d}\tau \le 6\left(\int_s^t K^2(t)\mathrm{d}t\right)^2,$$

for every $1 \le i \le d$ and $1 \le j \le m$, where

$$\tau_N = \inf\left\{t > 0 : \sup_{s\le\tau\le t}\left|\int_s^\tau g(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}B_\tau\right| \ge N\right\} \wedge T.$$

Then

$$E\left[\left|\int_s^{t\wedge\tau_N} g(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}B_\tau\right|^4\right] \le 6m^4\left(\int_s^t K^2(t)\mathrm{d}t\right)^2$$

for every $N \ge 1$, which implies that

$$E\left[\left|\int_s^t g(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}B_\tau\right|^4\right] \le 6m^4\left(\int_s^t K^2(t)\mathrm{d}t\right)^2.$$

Hence it follows that

$$\begin{aligned} P\left(\{|x^k_t - x^k_s| > \lambda\}\right) &\le \frac{T^2}{\lambda^4}\left(\int_s^t K^2(t)\mathrm{d}t\right)^2 + \frac{6m^4}{\lambda^4}\left(\int_s^t K^2(t)\mathrm{d}t\right)^2 \\ &\le \frac{1}{\lambda^4}\,|\Gamma(t) - \Gamma(s)|^2 \end{aligned}$$

for $s, t \in [0, T]$, where

$$\Gamma(t) = \sqrt{T^2 + 6m^4}\int_0^t K^2(\tau)\mathrm{d}\tau \quad \text{for} \quad 0 \le t \le T.$$

This, by virtue of Theorem 3.6 of Chap. 1, Theorem 2.2 of Chap. 1, and Theorem 2.3 of Chap. 1, implies that there exist an increasing sequence $(k_r)_{r=1}^\infty$ of positive integers, a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$, and d-dimensional continuous stochastic processes $\tilde{x}$ and $\tilde{x}^{k_r}$ on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ for $r = 1, 2, \ldots$, such that $P(x^{k_r})^{-1} = P(\tilde{x}^{k_r})^{-1}$ for $1, 2, \ldots$ and $\sup_{0\le t\le T}|\tilde{x}^{k_r}_t - \tilde{x}_t| \to 0$ with $(\tilde{P}.1)$ as $r \to \infty$. By Corollary 3.3 of Chap. 1, it follows that $P\tilde{x}_0^{-1} = \mu$, because $P(x_0^{k_r})^{-1} = \mu$ for $r = 1, 2, \ldots$ and $P(x_0^{k_r})^{-1} \Rightarrow P\tilde{x}_0^{-1}$ as $r \to \infty$. Let $\tilde{\mathbb{F}}$ be a filtration defined by a process $\tilde{x}$. Similarly as in the proof of Theorem 1.3, immediately from (1.2), it follows that $\mathbb{L}^{x^{k_r}}_{fg}$ generates on $C^2_b(\mathbb{R}^d)$ a family of continuous local $\mathbb{F}$-martingales for every $r = 1, 2, \ldots$, i.e., that $\mathcal{M}^{x^{k_r}}_{FG} \ne \emptyset$ for every $r = 1, 2, \ldots$, which by Lemma 1.3, implies that $\mathcal{M}^{\tilde{x}}_{FG} \ne \emptyset$. Thus there exist a filtered probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{F}}, \tilde{P})$ and a continuous $\tilde{\mathbb{F}}$-adapted process $\tilde{x}$ such that $P\tilde{x}_0^{-1} = \mu$ and $\mathcal{M}^{\tilde{x}}_{FG} \ne \emptyset$. Therefore, by virtue of Theorem 1.3, for every probability measure μ on $\beta(\mathbb{R}^d)$, one has $\mathcal{X}_\mu(F, G) \ne \emptyset$. □

Remark 1.6. If the assumptions of Theorem 1.5 are satisfied without the convexity of values of F, then $\mathcal{X}^0_\mu(\overline{F}, G) \neq \emptyset$.

Proof. By Lemma 1.7 of Chap. 3, one has $\mathcal{X}^0_\mu(\overline{F}, G) = \mathcal{X}^0_\mu(\mathrm{co}\, F, G)$. Similarly as in the proof of Theorem 1.5, by virtue of Theorem 1.4, one gets $\mathcal{X}^0_\mu(\mathrm{co}\, F, G) \neq \emptyset$. Then $\mathcal{X}^0_\mu(\overline{F}, G) \neq \emptyset$. □

2 Stochastic Differential Inclusions

Assume that $F : [0,T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : [0,T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ satisfy conditions $(\mathcal{H})$. By stochastic differential inclusions $SDI(F,G)$ and $\overline{SDI}(F,G)$, we mean relations of the form

$$x_t - x_s \in \int_s^t F(\tau, x_\tau)\mathrm{d}\tau + \int_s^t G(\tau, x_\tau)\mathrm{d}B_\tau\,, \quad \text{a.s.} \tag{2.1}$$

and

$$x_t - x_s \in \mathrm{cl}\left(\int_s^t F(\tau, x_\tau)\mathrm{d}\tau + \int_s^t G(\tau, x_\tau)\mathrm{d}B_\tau\right)\,, \quad \text{a.s.,} \tag{2.2}$$

which have to be satisfied for every $0 \le s \le t \le T$ by a system $(\mathcal{P}_\mathbb{F}, x, B)$ consisting of a complete filtered probability space $\mathcal{P}_\mathbb{F}$ with a filtration $\mathbb{F} = (\mathcal{F}_t)_{0\le t\le T}$ satisfying the usual conditions, a d-dimensional $\mathbb{F}$-adapted continuous stochastic process $x = (x_t)_{0\le t\le T}$, and an m-dimensional $\mathbb{F}$-Brownian motion $B = (B_t)_{0\le t\le T}$ on $\mathcal{P}_\mathbb{F}$, where $\int_s^t F(\tau, x_\tau)\mathrm{d}\tau$ and $\int_s^t G(\tau, x_\tau)\mathrm{d}B_\tau$ denote Aumann and Itô set-valued integrals of set-valued processes $F \circ x = (F(t, x_t))_{0\le t\le T}$ and $G \circ x = (G(t, x_t))_{0\le t\le T}$, respectively. Similarly as above, systems $(\mathcal{P}_\mathbb{F}, x, P)$ are said to be weak solutions of $SDI(F,G)$ and $\overline{SDI}(F,G)$, respectively. If μ is a given probability measure on $\beta(\mathbb{R}^d)$, then a system $(\mathcal{P}_\mathbb{F}, x, B)$ is said to be a weak solution of the initial value problems $SDI(F,G,\mu)$ or $\overline{SDI}(F,G,\mu)$, if it satisfies conditions (2.1) or (2.2) and $Px_0^{-1} = \mu$. If apart from the set-valued mappings F and G, we are also given a filtered probability space $\mathcal{P}_\mathbb{F}$ and an m-dimensional $\mathbb{F}$-Brownian motion B on $\mathcal{P}_\mathbb{F}$, then a continuous $\mathbb{F}$-adapted process X such that the system $(\mathcal{P}_\mathbb{F}, X, B)$ satisfies (2.1) or (2.2) is said to be a strong solution of $SDI(F,G)$ or $\overline{SDI}(F,G)$, respectively.

Corollary 2.1. *For every measurable set-valued mappings $F : [0,T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : [0,T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ every weak (strong) solution of $\overline{SFI}(F,G)$ is a weak (strong) solution of $\overline{SDI}(F,G)$.*

Proof. If $(\mathcal{P}_\mathbb{F}, x, B)$ is a weak solution of $\overline{SFI}(F,G)$, then $S_\mathbb{F}(F \circ x) \neq \emptyset$ and $S_\mathbb{F}(G \circ x) \neq \emptyset$. A set $\mathrm{cl}_\mathbb{L}\{J_{st}[S_\mathbb{F}(F \circ x)] + \mathcal{J}_{st}[S_\mathbb{F}(G \circ x)]\}$ is a subset of $\mathrm{cl}_\mathbb{L}\{\overline{\mathrm{dec}}\{J_{st}[S_\mathbb{F}(F \circ x)]\} + \overline{\mathrm{dec}}\{\mathcal{J}_{st}[S_\mathbb{F}(G \circ x)]\}\}$ for every $0 \le s \le t \le T$ and every continuous $\mathbb{F}$-adapted d-dimensional stochastic process $x = (x_t)_{0\le t\le T}$. From this and Theorem 2.1 of Chap. 3, it follows that every weak solution of $\overline{SFI}(F,G)$ is a weak solution of $\overline{SDI}(F,G)$. In a similar way, the above result for strong solutions can be obtained. □

Corollary 2.2. *For set-valued mappings* $F : [0,T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ *and* $G : [0,T]\times\mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ *satisfying conditions* ($\mathcal{H}$)*, every weak (strong) solution of* $SFI(F,G)$ *is a weak (strong) solution of* $SDI(F,G)$*.*

Proof. By (iv) of Theorem 2.1 of Chap. 3, a system $(\mathcal{P}_{\mathbb{F}}, x, B)$ is a weak solution of $\overline{SDI}(F,G)$ if and only if $x_t - x_s \in \overline{\mathrm{dec}}\{J_{st}[S_{\mathbb{F}}(F\circ x)]\} + \overline{\mathrm{dec}}\{\mathcal{J}_{st}[S_{\mathbb{F}}(G\circ x)]\}$ for every $0 \le s < t \le T$. But $J_{st}[S_{\mathbb{F}}(F\circ x)] + J_{st}[S_{\mathbb{F}}(G\circ x)] \subset \overline{\mathrm{dec}}\{J_{st}[S_{\mathbb{F}}(F\circ x)]\} + \overline{\mathrm{dec}}\{\mathcal{J}_{st}[S_{\mathbb{F}}(G\circ x)]\}$ for every $0 \le s < t \le T$. Then every weak solution of $SFI(F,G)$ is a weak solution of $\overline{SDI}(F,G)$. But for every F and G satisfying conditions ($\mathcal{H}$), a stochastic differential inclusion $\overline{SDI}(F,G)$ is reduced to the form $SDI(F,G)$, because in this case, $\int_s^t F(\tau,x_\tau)\mathrm{d}\tau + \int_s^t G(\tau,x_\tau)\mathrm{d}B_\tau$ is a closed subset of $\mathbb{R}^d$. Therefore, every weak solution of $SFI(F,G)$ is a weak solution of $SDI(F,G)$. In a similar way, the above result for strong solutions of the above inclusions can be obtained. □

It is natural to expect that for every strong solution $(\mathcal{P}_{\mathbb{F}}, x, B)$ of $SDI(F,G)$ and every $\varepsilon > 0$, there exist a partition $(A_k)_{k=1}^N \in \Pi(\Omega,\mathcal{F}_T)$ and a family $(\mathcal{P}_{\mathbb{F}}, x^k, B)_{k=1}^N$ of strong solutions of $SFI(F,G)$ such that $\|(x_t - x_s) - \sum_{k=1}^N \mathbb{1}_{A_k}(x_t^k - x_s^k)\| \le \varepsilon$ for every $0 \le s < t \le T$, where $\|\cdot\|$ is the norm of $\mathbb{L}^2(\Omega,\mathcal{F}_T,\mathbb{R}^d)$. It seems that the proof of such a result depends in an essential way on the $\mathbb{L}^2$-continuity of the mapping $[0,T] \ni t \to x_t \in \mathbb{L}^2(\Omega,\mathcal{F}_t,\mathbb{R}^d)$. By the definition of solutions of $SDI(F,G)$, it follows that the mapping $[0,T] \ni t \to x_t(\omega) \in \mathbb{R}^d$ is continuous for a.e. $\omega \in \Omega$. Therefore, a family $(x_t)_{0\le t\le T}$ of random variables $x_t : \Omega \to \mathbb{R}^d$ has to be uniformly square integrably bounded. But this depends, among other things, on the uniform square integrable boundedness of $(\int_0^t G(\tau,x_\tau)\mathrm{d}B_\tau)_{0\le t\le T}$. From the properties of set-valued integrals $\int_0^t G(\tau,x_\tau)\mathrm{d}B_\tau$, it follows that such a property of the family $(\int_0^t G(\tau,x_\tau)\mathrm{d}B_\tau)_{0\le t\le T}$ is difficult to obtain. Therefore, the desired above property is difficult to obtain. We can prove the following theorem.

Theorem 2.1. *Let* $B = (B_t)_{t\ge 0}$ *be an* m*-dimensional* $\mathbb{F}$*-Brownian motion on a filtered probability space* $\mathcal{P}_{\mathbb{F}} = (\Omega,\mathcal{F},\mathbb{F},P)$ *with a filtration* $\mathbb{F}$ *satisfying the usual conditions and Hölder continuous with exponential* $\alpha = 3$*. Assume that* $F : [0,T]\times\mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ *and* $G : [0,T]\times\mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ *are measurable, uniformly square integrably bounded, and Lipschitz continuous with respect to the second variable for every fixed* $t \in [0,T]$ *with a Lipschitz function* $k \in \mathbb{L}^2([0,T],\mathbb{R})$*. Then for every* $\varepsilon > 0$ *and every strong solution* x *of* $\overline{SDI}(F,G)$*, there exist a number* $\lambda_\varepsilon > 0$ *and a strong* $\varepsilon\lambda_\varepsilon$*-approximating solution* x^ε *of* $SFI(F,G)$ *such that* $\sup_{0\le t\le T}\|x_t - x_t^\varepsilon\| \le \varepsilon\lambda_\varepsilon$*, i.e., there exists a continuous* $\mathbb{F}$*-adapted stochastic process* x^ε *such that* $x_t^\varepsilon - x_s^\varepsilon \in \{J_{st}[S_{\mathbb{F}}(F\circ x^\varepsilon)] + \mathcal{J}_{st}[S_{\mathbb{F}}(G\circ x^\varepsilon)]\} + \varepsilon\lambda_\varepsilon\mathcal{B}$ *for every* $0 \le s < t \le T$*, where* $\mathcal{B}$ *denotes the closed unit ball of* $\mathbb{L}^2(\Omega,\mathcal{F},\mathbb{R}^d)$*.*

Proof. Let $x = (x_t)_{0\le t\le T}$ be a strong solution of $\overline{SDI}(F, G)$ and $\varepsilon > 0$. By virtue of Remark 2.3 of Chap. 3, for $\bar{\varepsilon} = \varepsilon/L(1 + \sqrt{T})$ there exist a number $\lambda_\varepsilon = 1 + m_\varepsilon\beta\left[3\sqrt{6}d(T + 2\delta_\varepsilon) + T + \delta_\varepsilon^3\sqrt{\delta_\varepsilon}\right]$ and processes $f^\varepsilon \in S_{\mathbb{F}}(F \circ x)$ and $g^\varepsilon \in S_{\mathbb{F}}(G \circ x)$ such that $\sup_{0\le t\le T} \|x_t - x_t^\varepsilon\| \le \lambda_\varepsilon\varepsilon/L(1 + \sqrt{T})$, where $L^2 = \int_0^T k_t^2 \mathrm{d}t$ and $x_t^\varepsilon = x_0 + \int_0^t f_\tau^\varepsilon \mathrm{d}\tau + \int_0^t g_\tau^\varepsilon \mathrm{d}B_\tau$ a.s. for $0 \le t \le T$. Hence in particular, it follows that $x_t^\varepsilon - x_s^\varepsilon \in J_{st}[S_{\mathbb{F}}(F \circ x)] + \mathcal{J}_{st}[S_{\mathbb{F}}(G \circ x)]$ for every $0 \le s < t \le T$. Similarly as in the proof of Remark 4.1 of Chap. 2, we obtain

$$
\begin{aligned}
&H\left(\mathrm{cl}_{\mathbb{L}}\{J_{st}[S_{\mathbb{F}}(F\circ x)]+\mathcal{J}_{st}[S_{\mathbb{F}}(G\circ x)]\}, \mathrm{cl}_{\mathbb{L}}\{J_{st}[S_{\mathbb{F}}(F\circ x^\varepsilon)]+\mathcal{J}_{st}[S_{\mathbb{F}}(G\circ x^\varepsilon)]\}\right)\\
&\quad = H\left(J_{st}[S_{\mathbb{F}}(F\circ x)] + \mathcal{J}_{st}[S_{\mathbb{F}}(G\circ x)], J_{st}[S_{\mathbb{F}}(F\circ x^\varepsilon)] + \mathcal{J}_{st}[S_{\mathbb{F}}(G\circ x^\varepsilon)]\right)\\
&\quad \le L(1+\sqrt{T}) \sup_{0\le t\le T} \|x_t - x_t^\varepsilon\|
\end{aligned}
$$

for every $0 \le s \le t \le T$. Therefore, for every $0 \le s \le t \le T$, we get

$$
\begin{aligned}
&\mathrm{dist}\left(x_t^\varepsilon - x_s^\varepsilon, J_{st}[S_{\mathbb{F}}(F\circ x^\varepsilon)] + \mathcal{J}_{st}[S_{\mathbb{F}}(G\circ x^\varepsilon)]\right)\\
&\quad \le H\left(J_{st}[S_{\mathbb{F}}(F\circ x] + \mathcal{J}_{st}[S_{\mathbb{F}}(G\circ x)], J_{st}[S_{\mathbb{F}}(F\circ x^\varepsilon)] + \mathcal{J}_{st}[S_{\mathbb{F}}(G\circ x^\varepsilon)]\right)\\
&\quad \le L(1+\sqrt{T}) \sup_{0\le t\le T} \|x_t - x_t^\varepsilon\|.
\end{aligned}
$$

Then $x_t^\varepsilon - x_s^\varepsilon \in \{J_{st}[S_{\mathbb{F}}(F\circ x^\varepsilon)] + \mathcal{J}_{st}[S_{\mathbb{F}}(G\circ x^\varepsilon)]\} + \varepsilon\lambda_\varepsilon\mathcal{B}$ for every $0 \le s < t \le T$, where $\mathcal{B}$ denotes the closed unit ball of $\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{R}^d)$. □

Remark 2.1. It is difficult to obtain better properties of $SDI(F, G)$, because up to now, we have not been able to prove that the uniform integrable boundedness of G and continuity of $G(t, \cdot)$ imply the integrable boundedness and continuity of the Itô integral $\int_0^T G(t, \cdot)\mathrm{d}B_t$. □

3 Backward Stochastic Differential Inclusions

We shall consider now a special case of stochastic differential inclusions. They are written as relations of the form $x_s \in E[x_t + \int_s^t F(\tau, x_\tau)\mathrm{d}\tau|\mathcal{F}_s]$ a.s., where $F : [0, T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ is a given measurable set-valued mapping and $E[x_t + \int_s^t F(\tau, x_\tau)\mathrm{d}\tau|\mathcal{F}_s]$ denotes the set-valued conditional expectation of $x_t + \int_s^t F(\tau, x_\tau)\mathrm{d}\tau$. Such relations are considered together with a terminal condition $x_T \in H(x_T)$ a.s. for a given set-valued mapping $H : \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$. In what follows, the terminal problem presented above will be denoted by $BSDI(F, H)$ and called a backward stochastic differential inclusion. By a weak solution of $BSDI(F, H)$, we mean a system $(\mathcal{P}_{\mathbb{F}}, x)$ consisting of a complete filtered probability space $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ with a filtration $\mathbb{F} = (\mathcal{F}_t)_{0\le t\le T}$ satisfying the usual conditions and a càdlàg d-dimensional stochastic process $x = (x_t)_{0\le t\le T}$ such that the following conditions are satisfied:

$$\begin{cases} x_s \in E[x_t + \int_s^t F(\tau, x_\tau)\mathrm{d}\tau|\mathcal{F}_s] \text{ a.s. for } 0 \le s < t \le T, \\ x_T \in H(x_T) \text{ a.s.} \end{cases} \tag{3.1}$$

Similarly as in the theory of stochastic differential inclusions, we can consider the terminal problem $BSDI(F, H)$ if apart from F and H, a filtered probability space $\mathcal{P}_{\mathbb{F}}$ is also given. In such a case, a d-dimensional càdlég process x on $\mathcal{P}_{\mathbb{F}}$ such that a system $(\mathcal{P}_{\mathbb{F}}, x)$ satisfies (3.1) is said of be a strong solution of $BSDI(F, H)$ on $\mathcal{P}_{\mathbb{F}}$. It is clear that if x is a strong solution of $BSDI(F, H)$ on $\mathcal{P}_{\mathbb{F}}$, then the pair $(\mathcal{P}_{\mathbb{F}}, x)$ is a weak solution. The set of all weak solutions of $BSDI(F, H)$ is denoted by $\mathcal{B}(F, H)$, and a subset containing all $(\mathcal{P}_{\mathbb{F}}, x) \in \mathcal{B}(F, H)$ with a continuous process x is denoted by $\mathcal{CB}(F, H)$. We obtain the following result immediately from Theorem 3.1 of Chap. 3.

Corollary 3.1. *If $F : [0, T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $H : \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ are measurable and uniformly integrably bounded, then $(\mathcal{P}_{\mathbb{F}}, x) \in \mathcal{B}(F, H)$ if and only if $x_T \in H(x_T)$ a.s. and there exists $f \in S(\mathrm{co}\, F \circ x)$, a measurable selector of $\mathrm{co}\, F \circ x$, such that $x_t = E[x_T + \int_t^T f_\tau \mathrm{d}\tau|\mathcal{F}_t]$ a.s. for every $0 \le t \le T$.* □

Backward stochastic differential inclusions can be regarded as generalizations of backward stochastic differential equations:

$$x_t = E\left[h(x) + \int_t^T f(\tau, x_\tau, z_\tau)\mathrm{d}\tau|\mathcal{F}_t\right] \text{ a.s.,} \tag{3.2}$$

where the triplet (h, f, z) is called the data set of such an equation. Usually, if we consider strong solutions of (3.2) apart from (h, f, z), a probability space $\mathcal{P} = (\Omega, \mathcal{F}, P)$ is also given, and the filtration $\mathbb{F}^z$ is defined to be the smallest filtration satisfying the usual conditions and such that the process z is $\mathbb{F}^z$-adapted. The process z is called the driving process. In practical applications, the driving process z is taken as a d-dimensional Brownian motion or a strong solution of a forward stochastic differential equation. In the case of weak solutions of (3.2) apart from h and f, a probability measure μ on the space $\mathcal{D}_T(\mathbb{R}^d)$ of d-dimensional càdlàg functions on $[0, T]$ is also given, a weak solution of which with an initial distribution μ is defined as a system $(\mathcal{P}_{\mathbb{F}}, x, z)$ satisfying (3.2) and $Pz^{-1} = \mu$, and such that every $\mathbb{F}^z$-martingale is also an $\mathbb{F}$-martingale. Let us observe that in a particular case, for a given weak solution $(\mathcal{P}_{\mathbb{F}}, x)$ of $BSDI(F, H)$ with $H(x) = \{h(x)\}$ and $F(t, x) = \{f(t, x, z) : z \in \mathcal{Z}\}$ for $(t, x) \in [0, T] \times \mathbb{R}^m$, where f and h are given measurable functions and $\mathcal{Z}$ is a nonempty compact subset of the space $\mathcal{D}_T(\mathbb{R}^d)$, there exists a measurable $\mathbb{F}$-adapted stochastic process $(z_t)_{0\le t\le T}$ with values in $\mathcal{Z}$ such that

$$x_t = E\left[h(x) + \int_t^T f(\tau, x_\tau, z_\tau)\mathrm{d}\tau|\mathcal{F}_t\right] \text{ a.s.} \tag{3.3}$$

For given probability measures μ_0 and μ_T on $\mathbb{R}^d$, we can look for a weak solution $(\mathcal{P}_{\mathbb{F}}, x)$ for $BSDI(F, H)$ such that $Px_0^{-1} = \mu_0$ and $Px_T^{-1} = \mu_T$. If F and H are as above, then there exists a measurable and $\mathbb{F}$-adapted stochastic process $(z_t)_{0\le t\le T}$ such that (3.3) is satisfied and such that $E[h(x) + \int_0^T f(\tau, x_\tau, z_\tau)\mathrm{d}\tau] = \int_{\mathbb{R}^d} u\mathrm{d}\mu_0$. If $f(t, x, z) = f(t, x) + g(z)$ with $g \in C(\mathcal{D}_T(\mathbb{R}^d), \mathbb{R}^d)$, then

$$\int_0^T \int_{\mathcal{D}_T(\mathbb{R}^d)} g(v)\mathrm{d}\lambda_\tau \mathrm{d}\tau = \int_{\mathbb{R}^d} u\mathrm{d}\mu_0 - \int_{\mathbb{R}^d} h(u)\mathrm{d}\mu_T - E\int_0^T f(\tau, x_\tau)\mathrm{d}\tau,$$

where $\lambda_t = Pz_t^{-1}$ for $t \in [0, T]$.

In some special cases, weak solutions of $BSDI(F, H)$ describe a class of recursive utilities under uncertainty. To verify this, suppose $(\mathcal{P}_{\mathbb{F}}, x)$ is a weak solution of $BSDI(F, H)$ with $H(x) = \{h(x)\}$ and $F(t, x) = \{f(t, x, c, z) : (c, z) \in \mathcal{C} \times \mathcal{Z}\}$, where h and f are measurable functions and $\mathcal{C}$, $\mathcal{Z}$ are nonempty compact subsets of $C([0, T], \mathbb{R}^+)$ and $\mathcal{D}_T(\mathbb{R}^d)$, respectively. Similarly as above, we can find a pair of measurable $\mathbb{F}$-adapted stochastic processes $(c_t)_{0\le t\le T}$ and $(z_t)_{0\le t\le T}$ with values in $\mathcal{C}$ and $\mathcal{Z}$, respectively, such that

$$x_t = E\left[h(x) + \int_t^T f(\tau, x_\tau, c_\tau, z_\tau)\mathrm{d}\tau | \mathcal{F}_t\right] \quad \text{a.s.} \tag{3.4}$$

for $0 \le t \le T$. In such a case, (3.4) describes a certain class of recursive utilities under uncertainty, where $(c_t(s, \cdot))_{0\le s\le T}$ denotes for fixed $t \in [0, T]$ the future consumption. Let us observe that in some special cases, a strong solution x of $BSDI(F, H)$ on a filtered probability space $\mathcal{P}_{\mathbb{F}}$ with the "constant" filtration $\mathbb{F} = (\mathcal{F}_t)_{0\le t\le T}$, i.e., such that $\mathcal{F}_t = \mathcal{F}$ for $0 \le t \le T$, is a solution of a backward random differential inclusion $-x_t' \in \overline{\mathrm{co}}\, F(t, x_t)$ with a terminal condition $x_T \in H(x_T)$ that has to be satisfied a.s. for a.e. $t \in [0, T]$.

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 $\mathbb{F} = (\mathcal{F}_t)_{0\le t\le T}$ satisfying the usual hypotheses, and by $\mathbb{D}(\mathbb{F}, \mathbb{R}^d)$ and $\mathbf{C}(\mathbb{F}, \mathbb{R}^d)$, we denote the spaces of all d-dimensional $\mathbb{F}$-adapted càdlàg and continuous, respectively, processes X on $\mathcal{P}_{\mathbb{F}}$ such that $\|X\|^2 = E[\sup_{s\in[0,T]}|X_s|^2] < \infty$. Similarly as above, we denote by $\mathcal{S}(\mathbb{F}, \mathbb{R}^d)$ the set of all d-dimensional $\mathbb{F}$-semimartingales X on $\mathcal{P}_{\mathbb{F}}$ such that $\|X\|^2 = E[\sup_{s\in[0,T]}|X_s|^2] < \infty$. We have $\mathbf{C}(\mathbb{F}, \mathbb{R}^d) \subset \mathbb{D}(\mathbb{F}, \mathbb{R}^d)$ and $\mathcal{S}(\mathbb{F}, \mathbb{R}^d) \subset \mathbb{D}(\mathbb{F}, \mathbb{R}^d)$. It can be proved that $(\mathcal{S}(\mathbb{F}, \mathbb{R}^d), \|\cdot\|)$ is a Banach space. In what follows, we shall assume that $F : [0, T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $H : \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ satisfy the following conditions $(\mathcal{A})$:

(i) F is measurable and uniformly square integrably bounded;
(ii) H is measurable and bounded;
(iii) $F(t, \cdot)$ is Lipschitz continuous for a.e. fixed $t \in [0, T]$;
(iv) there is a random variable $\xi \in \mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$ such that $\xi \in H(\xi)$ a.s.

We shall prove that conditions $(\mathcal{A})$ are sufficient for the existence of strong solutions for $BSDI(F, H)$, which implies that $\mathcal{B}(F, H)$ is nonempty. It is natural to look for weaker conditions implying the nonemptiness of $\mathcal{B}(F, H)$. The problem is quite complicated. It needs new sufficient conditions for tightness of sets of probability measures. We do not consider it in this book.

Lemma 3.1. *Let $F : [0, T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $H : \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ satisfy conditions $(\mathcal{A})$. For every filtered probability space $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ and a random variable $\xi : \Omega \to \mathbb{R}^d$, there exists a sequence $(x^n)_{n=0}^{\infty}$ of $\mathcal{S}(\mathbb{F}, \mathbb{R}^d)$ defined by $x_t^n = E[\xi + \int_t^T f_\tau^{n-1} \mathrm{d}\tau | \mathcal{F}_t]$ a.s. and $0 \le t \le T$ with $x^0 \in \mathcal{S}(\mathbb{F}, \mathbb{R}^d)$ satisfying $x_T^0 = \xi$ a.s. and $f^{n-1} \in S_{\mathbb{F}}(\mathrm{co} F \circ x^{n-1})$ for $n = 1, 2, \ldots$ such that*

$$E[\sup_{t \le u \le T} |x_u^{n+1} - x_u^n|^2] \le 4E\left[\int_t^T K(\tau) \sup_{\tau \le s \le T} |x_s^n - x_s^{n-1}| \mathrm{d}\tau\right]^2$$

for $n = 1, 2 \ldots$ and $0 \le t \le T$, with $K(t) = K_d \cdot k(t)$ for $0 \le t \le T$, where $k \in \mathbb{L}^2([0, T], \mathbb{R}^+)$ is a Lipschitz function of $F(t, \cdot)$ and K_d is the number defined in Remark 2.6 of Chap. 2.

Proof. Let $\mathcal{P}_{\mathbb{F}}$ be a filtered probability space and let $x^0 = (x_t^0)_{0 \le t \le T} \in \mathcal{S}(\mathbb{F}, \mathbb{R}^d)$ be such that $x_T^0 = \xi$ a.s. Put $f_t^0 = s(\overline{\mathrm{co}}\, F(t, x_t^0))$ a.s. for $0 \le t \le T$, where s is the Steiner point mapping defined by formula (2.1) of Chap. 2. It is clear that $f^0 \in S_{\mathbb{F}}(\overline{\mathrm{co}}\, F \circ x^0)$, because by virtue of Corollary 2.2 of Chap. 2, the function $s(\overline{\mathrm{co}}\, F(t, \cdot))$ is Lipschitz continuous for a.e. fixed $0 \le t \le T$, and x^0 is $\mathbb{F}$-adapted. We now define a sequence $(x^n)_{n=1}^{\infty}$ by the successive approximation procedure, i.e., by taking $x_t^n = E[\xi + \int_t^T f_\tau^{n-1} \mathrm{d}\tau | \mathcal{F}_t]$ a.s. for $n = 1, 2, \ldots$ and $0 \le t \le T$, where $f_t^{n-1} = s(\overline{\mathrm{co}}\, F(t, x_t^{n-1}))$ a.s. for $0 \le t \le T$. Similarly as above, we have $f^{n-1} \in S_{\mathbb{F}}(\overline{\mathrm{co}}\, F \circ x^{n-1})$. By Corollary 3.2 of Chap. 3, we have $x^n \in \mathcal{S}(\mathbb{F}, \mathbb{R}^d)$. Immediately from the above definitions and Corollary 2.2 of Chap. 2, it follows that $|f_t^n - f_t^{n-1}| \le K(t) \sup_{t \le s \le T} |x_s^n - x_s^{n-1}|$ a.s. for a.e. $0 \le t \le T$ and $n = 1, 2, \ldots$. Hence it follows that

$$|x_t^{n+1} - x_t^n| \le E\left[\int_t^T |f_\tau^n - f_\tau^{n-1}| \mathrm{d}\tau | \mathcal{F}_t\right] \le E\left[\int_t^T K(\tau) \sup_{\tau \le s \le T} |x_s^n - x_s^{n-1}| \mathrm{d}\tau | \mathcal{F}_t\right]$$

a.s. for $0 \le t \le T$. Therefore,

$$\begin{aligned}\sup_{t \le u \le T} |x_u^{n+1} - x_u^n| &\le \sup_{t \le u \le T} E\left[\int_u^T K(\tau) \sup_{\tau \le s \le T} |x_s^n - x_s^{n-1}| \mathrm{d}\tau | \mathcal{F}_u\right] \\ &\le \sup_{t \le u \le T} E\left[\int_t^T K(\tau) \sup_{\tau \le s \le T} |x_s^n - x_s^{n-1}| \mathrm{d}\tau | \mathcal{F}_u\right]\end{aligned}$$

a.s. for $0 \le t \le T$ and $n = 1, 2, \ldots$. By Doob's inequality, we obtain

$$E\left(\sup_{t \le u \le T} E\left[\int_t^T K(\tau) \sup_{\tau \le s \le T} |x_s^n - x_s^{n-1}| \mathrm{d}\tau | \mathcal{F}_u\right]\right)^2 \le 4E\left(\int_t^T K(\tau) \sup_{\tau \le s \le T} |x_s^n - x_s^{n-1}| \mathrm{d}\tau\right)^2$$

for $0 \le t \le T$. Therefore, for every $n = 1, 2, \ldots$ and $0 \le t \le T$, we have

$$E\left[\sup_{t\le u\le T} |x_u^{n+1} - x_u^n|^2\right] \le 4E\left(\int_t^T K(\tau) \sup_{\tau\le t\le T} |x_s^n - x_s^{n-1}|\mathrm{d}\tau\right)^2. \qquad \square$$

We obtain the following result immediately from the properties of multivalued conditional expectations.

Lemma 3.2. *If F satisfies conditions $(\mathcal{A})$, then for every $x, y \in \mathcal{S}(\mathbb{F}, \mathbb{R}^d)$, one has*

$$E\left[h\left(E\left[\int_s^t F(\tau, x_\tau)\,\mathrm{d}\tau|\mathcal{F}_s\right], E\left[\int_s^t F(\tau, y_\tau)\mathrm{d}\tau|\mathcal{F}_s\right]\right)\right] \le \int_s^t k(\tau)E|x_\tau - y_\tau|\mathrm{d}\tau$$

for every $0 \le s \le t \le T$, where h is the Hausdorff metric on $\mathrm{Cl}(\mathbb{R}^d)$.

We can now prove the following existence theorem.

Theorem 3.1. *If $F : [0, T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $H : \mathbb{R}^m \to \mathrm{Cl}(\mathbb{R}^m)$ satisfy conditions $(\mathcal{A})$, then for every complete filtered probability space $\mathcal{P}_{\mathbb{F}}$ and fixed point ξ of H, there exists a strong solution of (3.1).*

Proof. Let $\mathcal{P}_{\mathbb{F}}$ be given and assume that $\xi \in \mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$ is such that $\xi \in H(\xi)$. By virtue of Lemma 3.1, there exists a sequence $(x^n)_{n=1}^{\infty}$ of $\mathcal{S}(\mathbb{F}, \mathbb{R}^d)$ such that $x_T^n = \xi$, $x_s^n \in E[x_t^n + \int_s^t F(\tau, x_\tau^{n-1}\mathrm{d}\tau|\mathcal{F}_t]$ a.s. for $0 \le s \le t \le T$ and

$$E\left[\sup_{t\le u\le T} |x_u^{n+1} - x_u^n|^2\right] \le 4E\left(\int_t^T K(\tau) \sup_{\tau\le s\le T} |x_s^n - x_s^{n-1}|\mathrm{d}\tau\right)^2$$

for $n = 1, 2, \ldots$ and $0 \le t \le T$. By properties of F and H, one has $E[\sup_{t\le u\le T} |x_u^1 - x_u^0|^2] \le L$, where $L = 4[E|\xi|^2 + \int_0^T m^2(\tau)\mathrm{d}\tau] + 2E[\sup_{0\le t\le T} |x_t^0|^2]$ with $m \in \mathbb{L}^2([0, T], \mathbb{R}^+)$ such that $\|F(t, x)\| \le m(t)$ for every $x \in \mathbb{R}^d$ and a.e. $0 \le t \le T$. Therefore,

$$E\left[\sup_{t\le u\le T} |x_u^2 - x_u^1|^2\right] \le 4TL\int_t^T K^2(\tau)\mathrm{d}\tau.$$

Hence it follows that

$$E\left[\sup_{t\le u\le T} |x_u^3 - x_u^2|^2\right] \le (4T)^2L\int_t^T \left(K^2(\tau)\int_\tau^T K^2(s)\mathrm{d}s\right)\mathrm{d}\tau$$

$$= \frac{(4T)^2L}{2}\left(\int_t^T K^2(\tau)\mathrm{d}\tau\right)^2.$$

By the inductive procedure, for every $n = 1, 2, \ldots$ and $0 \leq t \leq T$, we get

$$E\left[\sup_{t\leq u\leq T}|x_u^{n+1}-x_u^n|^2\right]\leq\frac{(4T)^nL^{n-1}}{n!}\left(\int_t^T K^2(\tau)\mathrm{d}\tau\right)^n.$$

Then $(x^n)_{n=1}^{\infty}$ is a Cauchy sequence of $\mathcal{S}(\mathbb{F}, \mathbb{R}^d)$. Therefore, there exists a process $(x_t)_{0\leq t\leq T} \in \mathcal{S}(\mathbb{F}, \mathbb{R}^d)$ such that $E[\sup_{0\leq t\leq T}|x_t^n - x_t|^2] \to 0$ as $n \to \infty$. By Lemma 3.2, it follows that

$$\begin{aligned}
&E\,\mathrm{dist}\left(x_s, E\left[x_t+\int_s^t F(\tau,x_\tau)\mathrm{d}\tau|\mathcal{F}_s\right]\right)\\
&\leq E|x_s-x_s^n|]+E\left[\mathrm{dist}\left(x_s^n, E\left[x_t^n+\int_s^t F(\tau,x_\tau^{n-1})\mathrm{d}\tau|\mathcal{F}_s\right]\right)\right]\\
&\quad+E\left[h\left(E\left[x_t^n+\int_s^t F(\tau,x_\tau^{n-1})\mathrm{d}\tau|\mathcal{F}_s\right], E\left[x_t+\int_s^t F(\tau,x_\tau)\mathrm{d}\tau|\mathcal{F}_s\right]\right)\right]\\
&\leq E|x_s^n-x_s|+E|x_t^n-x_t|+\int_s^t K(\tau)E|x_\tau^{n-1}-x_\tau|\mathrm{d}\tau\\
&\leq 2\|x^n-x\|+\left(\int_0^T K^2(\tau)\mathrm{d}\tau\right)^{\frac{1}{2}}\|x^{n-1}-x\|
\end{aligned}$$

for every $0 \leq s \leq t \leq T$ and $n = 1, 2, \ldots$. Therefore, $\mathrm{dist}(x_s, E[x_t + \int_s^t F(\tau, x_\tau)\mathrm{d}\tau|\mathcal{F}_s]) = 0$ a.s. for every $0 \leq s \leq t \leq T$, which implies that $x_s \in E\left[x_t + \int_s^t F(\tau, x_\tau)\mathrm{d}\tau|\mathcal{F}_s\right]$ a.s for every $0 \leq s \leq t \leq T$. By the definition of $(x_t^n)_{0\leq t\leq T}$, we have $x_T^n = \xi \in H(\xi)$ a.s. for every $n = 1, 2, \ldots$. Therefore, we also have $x_T = \xi$ a.s. Thus $x_T \in H(x_T)$ a.s. Then x satisfies (3.1). □

4 Weak Compactness of Solution Sets

For given measurable multifunctions $F : [0, T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$, $G : [0, T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ and a probability measure μ on $\beta(\mathbb{R}^d)$, by $\mathcal{X}_\mu(F, G)$ we denote, similarly as above, the set of all weak solutions (equivalence classes defined in Sect. 1) of $SFI(F, G, \mu)$. Elements $[(\mathcal{P}_{\mathbb{F}}, X, B)]$ of $\mathcal{X}_\mu(F, G)$ will be identified with equivalence classes $[X]$ of all d-dimensional continuous processes Z such that $PX^{-1} = PZ^{-1}$. In what follows, $[X]$ will be denoted simply by X. It is clear that we can associate with every $[(\mathcal{P}_{\mathbb{F}}, X, B)] \in \mathcal{X}_\mu(F, G)$ a probability measure PX^{-1}, a distribution of X, defined on a Borel σ-algebra $\beta(C_T)$ of the space $C_T =: C(([0, T], \mathbb{R}^d)$. The family of all such probability measures, corresponding to all classes belonging to $\mathcal{X}_\mu(F, G)$, is denoted by $\mathcal{X}_\mu^P(F, G)$. It is a subset of the space $\mathcal{M}(C_T)$ of probability measures on C_T. The set $\mathcal{X}_\mu(F, G)$ is said to be weakly compact, or weakly compact in distribution, if $\mathcal{X}_\mu^P(F,G)$ is a weakly compact subset of $\mathcal{M}(C_T)$. We now present sufficient conditions for the weak compactness of $\mathcal{X}_\mu(F,G)$.

Theorem 4.1. *Let $F : [0,T]\times\mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : [0,T]\times\mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ be measurable, uniformly square integrably bounded, and convex-valued such that $F(t,\cdot)$ and $G(t,\cdot)$ are continuous for fixed $t \in [0,T]$. If G is diagonally convex, then for every probability measure μ on $\beta(\mathbb{R}^d)$, the set $\mathcal{X}_\mu(F,G)$ is nonempty and weakly compact.*

Proof. The nonemptiness of $\mathcal{X}_\mu(F,G)$ follows from Theorem 1.5. To show that $\mathcal{X}_\mu(F,G)$ is relatively weakly compact in the sense of distributions, let us note that by virtue of Theorem 1.5 of Chap. 3, for every $(\mathcal{P}_{\mathbb{F}},x,B) \in \mathcal{X}_\mu(F,G)$ there are $f \in S_{\mathbb{F}}(F\circ x)$ and $g \in S_{\mathbb{F}}(G\circ x)$ such that $Px_0^{-1} = \mu$ and $x_t = x_0 + \int_0^t f_\tau \mathrm{d}\tau + \int_0^t g_\tau \mathrm{d}B_\tau$ for every $t \in [0,T]$. Similarly as in the proof of Theorem 1.5, we can verify that every sequence $(\mathcal{P}_{\mathbb{F}^n}^n, x^n, B^n)_{n=1}^\infty$ of $\mathcal{X}_\mu(F,G)$ satisfies the conditions of Theorem 3.6 of Chap. 1. Therefore, for every sequence $(\mathcal{P}_{\mathbb{F}^n}^n, x^n, B^n)$ of $\mathcal{X}_\mu(F,G)$, there exists an increasing subsequence $(n_k)_{k=1}^\infty$ of $(n)_{n=1}^\infty$ such that the sequence $\{P(x^{n_k})^{-1}\}_{n=1}^\infty$ is weakly convergent in distribution. Then the sequence $(x^n)_{n=1}^\infty$ possesses a subsequence converging in distribution.

Let $(x^r)_{r=1}^\infty$ be a sequence of $\mathcal{X}_\mu(F,G)$ convergent in distribution. Then there exists a probability measure $\mathcal{P}$ on $\beta(C_T)$ such that $P(x^r)^{-1} \Rightarrow \mathcal{P}$ as $r\to\infty$. By virtue of Theorem 2.3 of Chap. 1, there exist a probability space $(\tilde{\Omega},\tilde{\mathcal{F}},\tilde{P})$ and random variables $\tilde{x}^r : \tilde{\Omega} \to C_T$ and $\tilde{x} : \tilde{\Omega} \to C_T$ for $r = 1,2,\ldots$ such that $P(x^r)^{-1} = P(\tilde{x}^r)^{-1}$ for $r = 1,2,\ldots$, $\tilde{P}(\tilde{x})^{-1} = \mathcal{P}$ and $\lim_{r\to\infty}\sup_{0\le t\le T}|\tilde{x}_t^r - \tilde{x}_t| = 0$ with $(\tilde{P}.1)$. Immediately from Corollary 3.3 of Chap. 1, it follows that $x_0^r \Rightarrow \tilde{x}_0$ as $r\to\infty$, because $P(x^r)^{-1} \Rightarrow P(\tilde{x})^{-1}$ as $r\to\infty$. But $P(x_0^r)^{-1} = \mu$ for every $r \ge 1$. Then $P\tilde{x}_0^{-1} = \mu$. By Theorem 1.3, we have $\mathcal{M}_{FG}^{x_r} \neq \emptyset$ for every $r \ge 1$, which by Lemma 1.3, implies that $\mathcal{M}_{FG}^{\tilde{x}} \neq \emptyset$. Therefore, by virtue of Theorem 1.3, there exist a standard extension $\hat{\mathcal{P}}_{\hat{\mathbb{F}}} = (\hat{\Omega},\hat{\mathcal{F}},\hat{\mathbb{F}},\hat{P})$ of $(\tilde{\Omega},\tilde{\mathcal{F}},\tilde{\mathbb{F}},\tilde{P})$ and an m-dimensional Brownian motion $\hat{B}$ such that $(\hat{\mathcal{P}}_{\hat{\mathbb{F}}},\hat{x},\hat{B})$, with $\hat{x}(\hat{\omega}) = \tilde{x}(\pi(\hat{\omega}))$ for every $\hat{\omega} \in \hat{\Omega}$, is a weak solution of $SFI(F,G,\mu)$, where $\pi : \hat{\Omega} \to \tilde{\Omega}$ is an $(\hat{\mathcal{F}},\tilde{\mathcal{F}})$-measurable mapping as described in the definition of the extension of $(\tilde{\Omega},\tilde{\mathcal{F}},\tilde{\mathbb{F}},\tilde{P})$, because its standard extension $\hat{\mathcal{P}}_{\hat{\mathbb{F}}}$ is also its extension. Let $\hat{x}^r(\hat{\omega}) = \tilde{x}^r(\pi(\hat{\omega}))$ for $\hat{\omega} \in \hat{\Omega}$. For every $A \in \beta(C)$, one has $P(\hat{x}^r)^{-1}(A) = \hat{P}[(\hat{x}^r)^{-1}(A)] = \hat{P}[(\tilde{x}^r\circ\pi)^{-1}(A)] = (\hat{P}\circ\pi^{-1})[(\tilde{x}^r)^{-1}(A)] = \tilde{P}[(\tilde{x}^r)^{-1}(A)] = P(\tilde{x}^r)^{-1}(A)$. Therefore, $P(\hat{x}^r)^{-1} = P(\tilde{x}^r)^{-1} = P(x^r)^{-1}$ for every $r \ge 1$. By the properties of the sequence $(\tilde{x}^r)_{r=1}^\infty$, it follows that $\tilde{x}_t^r(\tilde{\omega}) \to \tilde{x}_t(\tilde{\omega})$ with $(\tilde{P}.1)$ as $r\to\infty$ uniformly with respect to $0 \le t \le T$. Hence in particular, it follows that $\tilde{x}_t^r(\pi(\hat{\omega})) \to \tilde{x}_t(\pi(\hat{\omega}))$ with $(\hat{P}.1)$ as $r\to\infty$ uniformly with respect to $0 \le t \le T$. Therefore, for every $f \in C_b(C)$, one has $f(\hat{x}^r(\hat{\omega})) \to f(\hat{x}(\hat{\omega}))$ with $(\hat{P}.1)$ as $r\to\infty$. By the boundedness of $f \in C_b(C)$, this implies that $\hat{E}\{f(\hat{x}^r)\} \to \hat{E}\{f(\hat{x})\}$ as $r\to\infty$, which by Corollary 2.1 of Chap. 1, is equivalent to $P(\hat{x}^r)^{-1} \Rightarrow P\hat{x}^{-1}$. But $P(\hat{x}^r)^{-1} = P(x^r)^{-1}$ for every $r \geq 1$. Then $x^r \Rightarrow \hat{x}$, which implies that $\mathcal{X}_\mu(F, G)$ is weakly closed. □

In a similar way, we can prove the following theorem.

Theorem 4.2. *Let $F : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times d})$ be measurable and uniformly square integrably bounded such that $F(t,\cdot)$ and $G(t,\cdot)$ are continuous for fixed $t \in [0,T]$. If G is convex-valued and diagonally convex, then for every probability measure μ on $\beta(\mathbb{R}^d)$, the set $\mathcal{X}^0_\mu(\overline{F}, G)$ is nonempty and weakly compact.*

Proof. The nonemptiness of $\mathcal{X}^0_\mu(\overline{F}, G)$ follows from Remark 1.6. In a similar way as above, we can verify that the set $\mathcal{X}^0_\mu(\mathrm{co}\, F, G)$ of all weak solutions $(\mathcal{P}_\mathbb{F}, x, B)$ of $SFI(\mathrm{co}\, F, G)$ with a separable filtered probability space $\mathcal{P}_\mathbb{F}$ is weakly compact in distribution. By virtue of Lemma 1.7 of Chap. 3, one has $\mathcal{X}^0_\mu(\overline{F}, G) = \mathcal{X}^0_\mu(\mathrm{co}\, F, G)$. Then $\mathcal{X}^0_\mu(\overline{F}, G)$ is nonempty and weakly compact. □

5 Some Properties of Exit Times of Continuous Processes

Let D be a domain in $\mathbb{R}^d$ and $(s,x) \in \mathbb{R}^+ \times D$. Assume that $X = (X(\cdot,t))_{t\geq 0}$ and $X^n = (X^n(\cdot,t))_{t\geq 0}$ are continuous stochastic processes on a stochastic base $\mathcal{P}_\mathbb{F} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ such that $X(\cdot,s) = X^n(\cdot,s) = x$ a.s. for $n = 1,2,\ldots$ and $\sup_{t\geq 0} |X^n(\cdot,t) - X(\cdot,t)| \to 0$ a.s. as $n \to \infty$. Let $\tau = \inf\{r > s : X(\cdot,r) \notin D\}$ and $\tau_n = \inf\{r > s : X^n(\cdot,r) \notin D\}$ for $n = 1,2,\ldots$. We shall show that if $\tau_n < \infty$ a.s. for every $n \geq 1$, then $\tau_n \to \tau$ a.s. as $n \to \infty$. We begin with the following lemmas.

Lemma 5.1. *Let D be a domain in $\mathbb{R}^d$, $(s,x) \in \mathbb{R}^+ \times D$, and $X = (X(\cdot,t))_{t\geq 0}$ a continuous d-dimensional stochastic process on $\mathcal{P}_\mathbb{F} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ such that $X(\cdot,s) = x$ a.s. and $\tau = \inf\{r > s : X(\cdot,r) \notin D\} < \infty$ a.s. If $T : \Omega \to \mathbb{R}$ is such that $T > \tau$ a.s., then $\tau = \inf\{r \in (s,T) : X(\cdot,r) \notin D\}$ a.s.*

Proof. For simplicity, assume that the above relations are satisfied for every $\omega \in \Omega$ and let us observe that $\tau(\omega) = \inf X^{-1}(\omega,\cdot)(D^\sim)$, where $D^\sim = \mathbb{R}^d \setminus D$. We have $X^{-1}(\omega,\cdot)(D^\sim) = X^{-1}(\omega,\cdot)(D^\sim) \cap (s,T(\omega)) \cup X^{-1}(\omega,\cdot)(D^\sim) \cap [T(\omega),\infty)$. Therefore, $\inf X^{-1}(\omega,\cdot)(D^\sim) \leq \inf(X^{-1}(\omega,\cdot)(D^\sim) \cap (s,T(\omega)))$. For every $\omega \in \Omega$, there exists $t(\omega) \in X^{-1}(\omega,\cdot)(D^\sim)$ such that $s < t(\omega) < T(\omega)$, because $\tau(\omega) < T(\omega)$ for $\omega \in \Omega$. Therefore, $X^{-1}(\omega,\cdot)(D^\sim) \cap (s,T(\omega)) \neq \emptyset$ and $\inf(X^{-1}(\omega,\cdot)(D^\sim) \cap (s,T(\omega))) \leq T(\omega)$ for a.e. $\omega \in \Omega$. Suppose $\tau = \inf X^{-1}(\omega,\cdot)(D^\sim) < \tau_T(\omega) =: \inf(X^{-1}(\omega,\cdot)(D^\sim) \cap (s,T(\omega)))$ on a set $\Omega_0 \in \mathcal{F}$ such that $P(\Omega_0) > 0$. Then for every $\omega \in \Omega_0$, there exists $\bar{t}(\omega) \in X^{-1}(\omega,\cdot)(D^\sim)$ such that $s < \bar{t}(\omega) < \tau_T(\omega) < T(\omega)$, which is a contradiction, because for every $\omega \in \Omega$ and $t \in X^{-1}(\omega,\cdot)(D^\sim) \cap (s,T(\omega))$, we have $\tau_T(\omega) \leq t$. Then $\tau(\omega) = \inf\{X^{-1}(\omega,\cdot)(D^\sim) \cap (s,T(\omega))\}$ for a.e. $\omega \in \Omega$. □

Lemma 5.2. *Let D be a domain in $\mathbb{R}^d$ and $(s,x) \in \mathbb{R}^+ \times D$. Assume that $X = (X(\cdot,t))_{t\geq 0}$ and $X^n = (X^n(\cdot,t))_{t\geq 0}$ are continuous d-dimensional stochastic processes on a stochastic base $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ such that $X(\cdot,s) = X^n(\cdot,s) = x$ for $n = 1,2,\ldots$ and $\sup_{t\geq 0} |X^n(\cdot,t) - X(\cdot,t)| \to 0$ a.s. as $n \to \infty$. Then $\mathrm{Li}\, X_n^{-1}(\omega,\cdot)(D^{\sim}) = X^{-1}(\omega,\cdot)(D^{\sim}) = \mathrm{Ls}\, X_n^{-1}(\omega,\cdot)(D^{\sim})$ for a.e. $\omega \in \Omega$.*

Proof. For simplicity, assume that $X(\omega,\cdot)$ and $X_n(\omega,\cdot)$ for $n = 1,2,\ldots$ are continuous and $\lim_{n\to\infty} \sup_{t\geq 0} |X^n(\omega,t) - X(\omega,t)| = 0$ for every $\omega \in \Omega$. For every $\omega \in \Omega$ and $\varepsilon > 0$, there exists $N_\varepsilon(\omega) \geq 1$ such that $X_n(\omega,t) \in X(\omega,t) + \varepsilon B$ and $X(\omega,t) \in X_n(\omega,t) + \varepsilon B$ for $t \geq s$ and $n \geq N_\varepsilon(\omega)$, where B is a closed unit ball of $\mathbb{R}^d$. Then $X_n^{-1}(\omega,\cdot)(\{X_n(\omega,t)\}) \subset X_n^{-1}(\omega,\cdot)(\{X(\omega,t) + \varepsilon B\})$ and $X^{-1}(\omega,\cdot)(\{X(\omega,t)\}) \subset X^{-1}(\omega,\cdot)(\{X_n(\omega,t) + \varepsilon B\})$ a.s. for $n \geq N_\varepsilon(\omega)$. Let us observe that for every $A \subset \mathbb{R}^+$ and $C \subset \mathbb{R}^d$, one has $A \subset X_n^{-1}(\omega,\cdot)(X_n(\omega,A))$, $A \subset X^{-1}(\omega,\cdot)(X(\omega,A))$, $X_n(\omega, X^{-1}(\omega,\cdot)(C)) \subset C + \varepsilon B$, and $X(\omega, X^{-1}(\omega,\cdot)(C)) \subset C$ for $n = 1,2,\ldots$. Taking in particular $A = X^{-1}(\omega,\cdot)(D^{\sim})$ and $C = D^{\sim}$ in the above inclusions, we obtain $X^{-1}(\omega,\cdot)(D^{\sim}) \subset X_n^{-1}(\omega,\cdot)(X_n(\omega, X^{-1}(\omega,\cdot)(D^{\sim})) \subset X_n^{-1}(\omega,\cdot)(D^{\sim} + \varepsilon B)$ a.s. for $n \geq N_\varepsilon(\omega)$. Similarly, taking $A = X_n^{-1}(\omega,\cdot)(D^{\sim})$ and $C = D^{\sim}$, we obtain $X_n^{-1}(\omega,\cdot)(D^{\sim}) \subset X^{-1}(\omega,\cdot)(X(\omega, X_n^{-1}(\omega,\cdot)(D^{\sim}))) \subset X^{-1}(\omega,\cdot)(X_n(\omega, X_n^{-1}(\omega,\cdot)(D^{\sim})) + \varepsilon B) \subset X^{-1}(\omega,\cdot)(D^{\sim} + \varepsilon B)$ a.s. for $n \geq N_\varepsilon(\omega)$. Hence it follows that

$$\begin{aligned}
X^{-1}(\omega,\cdot)(D^{\sim}) &\subset \bigcap_{k=0}^{\infty} X_{k+N_\varepsilon}^{-1}(\omega,\cdot)(D^{\sim} + \varepsilon B) \\
&\subset \bigcup_{n=1}^{N_\varepsilon - 1} \bigcap_{k=0}^{\infty} X_{k+n}^{-1}(\omega,\cdot)(D^{\sim} + \varepsilon B) \\
&\cup \bigcap_{k=0}^{\infty} X_{k+N_\varepsilon}^{-1}(\omega,\cdot)(D^{\sim} + \varepsilon B) \\
&\cup \bigcup_{n=N_\varepsilon+1}^{\infty} \bigcap_{k=0}^{\infty} X_{k+n}^{-1}(\omega,\cdot)(D^{\sim} + \varepsilon B) \\
&= \bigcup_{n=1}^{\infty} \bigcap_{k=0}^{\infty} X_{k+n}^{-1}(\omega,\cdot)(D^{\sim} + \varepsilon B) \\
&= \mathrm{Lim\,inf}\, X_n^{-1}(\omega,\cdot)(D^{\sim} + \varepsilon B)
\end{aligned}$$

a.s. for every $\varepsilon > 0$, which by virtue of Corollary 1.1 of Chap. 2, implies $X^{-1}(\omega,\cdot)(D^{\sim}) \subset \bigcap_{\varepsilon>0} \mathrm{Lim\,inf}\, X^{-1}(\omega,\cdot)(D^{\sim} + \varepsilon B) = \mathrm{Lim\,inf}\, X_n^{-1}(\omega,\cdot)(D^{\sim})$ a.s. Hence, by virtue of (ii) of Lemma 1.2 of Chap. 2, we obtain $X^{-1}(\omega,\cdot)(D^{\sim}) \subset \mathrm{Li}\, X_n^{-1}(\omega,\cdot)(D^{\sim})$. In a similar way, we get $\overline{\bigcup_{k=0}^{\infty} X_{k+N_\varepsilon}^{-1}(\omega,\cdot)(D^{\sim} + \varepsilon B)} \subset X^{-1}(\omega,\cdot)(D^{\sim} + \varepsilon B)$. Then

$$\overline{\bigcap_{n=1}^{N_\varepsilon-1}\bigcup_{k=0}^{\infty}X_{k+n}^{-1}(\omega,\cdot)(D^{\sim})}\cup\overline{\bigcup_{k=0}^{\infty}X_{k+N_\varepsilon}^{-1}(\omega,\cdot)(D^{\sim})}\cup\bigcap_{n=N_\varepsilon+1}^{\infty}\overline{\bigcup_{k=0}^{\infty}X_{k+n}^{-1}(\omega,\cdot)(D^{\sim})}$$
$$\subset X^{-1}(\omega,\cdot)(D^{\sim}+\varepsilon B)$$

for every $\varepsilon > 0$. Hence, by virtue of (v) of Lemma 1.2 of Chap. 2, it follows that

$$\text{Ls } X_n^{-1}(\omega,\cdot)(D^{\sim}) = \bigcap_{n=1}^{\infty}\overline{\bigcup_{k=0}^{\infty}X_{k+n}^{-1}(\omega,\cdot)(D^{\sim})}\subset X^{-1}(\omega,\cdot)(D^{\sim}+\varepsilon B)$$

for every $\varepsilon > 0$. Thus Ls $X_n^{-1}(\omega,\cdot)(D^{\sim}) \subset \bigcap_{\varepsilon>0} X^{-1}(\omega,\cdot)(D^{\sim}+\varepsilon B) = X^{-1}(\omega,\cdot)(D^{\sim})$ a.s. From the above inclusions, we obtain $X^{-1}(\omega,\cdot)(D^{\sim}) \subset$ Li $X_n^{-1}(\omega,\cdot)(D^{\sim}) \subset$ Ls $X_n^{-1}(\omega,\cdot)(D^{\sim}) \subset X^{-1}(\omega,\cdot)(D^{\sim})$ a.s. Then

$$\text{Ls } X_n^{-1}(\omega,\cdot)(D^{\sim}) \subset X^{-1}(\omega,\cdot)(D^{\sim}) \subset \text{Li } X_n^{-1}(\omega,\cdot)(D^{\sim}),$$

which by (i) of Corollary 1.2 of Chap. 2, implies that Li $X_n^{-1}(\omega,\cdot)\,(D^{\sim}) =$ Ls $X_n^{-1}(\omega,\cdot)(D^{\sim}) = X^{-1}(\omega,\cdot)(D^{\sim})$. □

Lemma 5.3. *Let D be a domain in $\mathbb{R}^d$ and $(s,x) \in \mathbb{R}^+ \times D$. Assume that $X = (X(\cdot,t))_{t\geq 0}$ and $X^n = (X^n(\cdot,t))_{t\geq 0}$ are continuous d-dimensional stochastic processes on a stochastic base $\mathcal{P}_{\mathbb{F}} = (\Omega,\mathcal{F},\mathbb{F},P)$ such that $X(\cdot,s) = X^n(\cdot,s) = x$ for $n = 1,2,\ldots$ and $\sup_{t\geq 0}|X^n(\cdot,t) - X(\cdot,t)| \to 0$ a.s. as $n \to \infty$. If there exists a mapping $T : \Omega \to \mathbb{R}^+$ such that $\max(\tau,\tau_n) < T$ a.s. for $n = 1,2,\ldots$, where $\tau = \inf\{r > s : X(\cdot,r) \notin D\}$ and $\tau_n = \inf\{r > s : X_n(\cdot,r) \notin D\}$, then $(X^{-1}(\omega,\cdot)(D^{\sim})) \cap [s,T(\omega)) =$ Li $(X_n^{-1}(\omega,\cdot)(D^{\sim}) \cap (s,T(\omega))) =$ Ls $(X_n^{-1}(\omega,\cdot)(D^{\sim}) \cap (s,T(\omega)))$ for a.e. $\omega \in \Omega$.*

Proof. Assume that $X(\omega,\cdot)$ and $X_n(\omega,\cdot)$ for $n = 1,2,\ldots$ are continuous, $\max(\tau(\omega),\tau_n(\omega)) < T(\omega)$ for $n = 1,2,\ldots$, and $\lim_{n\to\infty}\sup_{t\geq 0}|X^n(\omega,t) - X(\omega,t)| = 0$ for every $\omega \in \Omega$. By virtue of (iv) and (vi) of Lemma 1.2 of Chap. 2 and Lemma 5.2, we get

$$\begin{aligned}\text{Ls } (X_n^{-1}(\omega,\cdot)(D^{\sim}) \cap (s,T(\omega))) &\subset \text{Ls } X_n^{-1}(\omega,\cdot)(D^{\sim}) \cap [s,T(\omega))\\ &= X^{-1}(\omega,\cdot)(D^{\sim}) \cap [s,T(\omega)).\end{aligned}$$

Similarly, by virtue of (iii) and (vi) of Lemma 1.2 of Chap. 2, we also have

$$\text{Li } (X_n^{-1}(\omega,\cdot)(D^{\sim}) \cap (s,T(\omega))) \subset (\text{Li } (X_n^{-1}(\omega,\cdot)(D^{\sim})) \cap [s,T(\omega)).$$

By virtue of (ii) of Corollary 1.2 of Chap. 2, for every $t \in (\text{Li } (X_n^{-1}(\omega,\cdot)(D^{\sim})) \cap [s,T(\omega))$, there exists $\bar{n} \geq 1$ such that for every $n > \bar{n}$, there is $t_n \in X_n^{-1}(\omega,\cdot)(D^{\sim}) \cap [s,T(\omega))$ such that $t_n \to t$ as $n \to \infty$. Then $\text{dist}(t, X_n^{-1}(\omega,\cdot)(D^{\sim})) \to 0$ as $n \to \infty$. Therefore, for every $\varepsilon > 0$, there exists $N_\varepsilon > \bar{n}$ such that $t \in X_n^{-1}(\omega,\cdot)(D^{\sim}) + \varepsilon B$ for $n \geq N_\varepsilon$. Hence, similarly as in the proof of

Lemma 5.2, it follows that for every $t \in \mathrm{Li}\,(X_n^{-1}(\omega,\cdot)(D^{\sim})) \cap [s, T(\omega))$ and $\varepsilon > 0$, one has

$$
\begin{aligned}
t &\in \bigcap_{k=0}^{\infty}\{(X_{k+N_\varepsilon}^{-1}(\omega,\cdot)(D^{\sim}) + \varepsilon B) \cap [s, T(\omega))\} \\
&\subset \bigcup_{n=1}^{\infty}\bigcap_{k=0}^{\infty}\{(X_{k+n}^{-1}(\omega,\cdot)(D^{\sim}) + \varepsilon B) \cap [s, T(\omega))\} \\
&= \mathrm{Lim\,inf}\{(X_n^{-1}(\omega,\cdot)(D^{\sim}) + \varepsilon B) \cap [s, T(\omega))\}.
\end{aligned}
$$

Then

$$
\mathrm{Li}(X_n^{-1}(\omega,\cdot)(D^{\sim})) \cap [s, T(\omega)) \subset \mathrm{Lim\,inf}\{X_n^{-1}(\omega,\cdot)(D^{\sim}) \cap [s, T(\omega))\}
$$

$$
\subset \mathrm{Li}\,(X_n^{-1}(\omega,\cdot)(D^{\sim}) \cap [s, T(\omega))) \subset \mathrm{Li}(X_n^{-1}(\omega,\cdot)(D^{\sim})) \cap [s, T(\omega)).
$$

Thus

$$
\begin{aligned}
X^{-1}(\omega,\cdot)(D^{\sim}) \cap [s, T(\omega)) &= (\mathrm{Li}\,X_n^{-1}(\omega,\cdot)(D^{\sim})) \cap [s, T(\omega)) \\
&= \mathrm{Li}\,(X_n^{-1}(\omega,\cdot)(D^{\sim}) \cap [s, T(\omega))).
\end{aligned}
$$

Therefore, by (iv) and (vi) of Lemma 1.2 of Chap. 2 and Lemma 5.2, one has

$$
\begin{aligned}
&\mathrm{Ls}\,(X_n^{-1}(\omega,\cdot)(D^{\sim}) \cap [s, T(\omega))) \\
&\quad \subset X^{-1}(\omega,\cdot)(D^{\sim}) \cap [s, T(\omega)) \\
&\quad = \mathrm{Li}\,(X_n^{-1}(\omega,\cdot)(D^{\sim}) \cap [s, T(\omega))).
\end{aligned}
$$

□

Lemma 5.4. *Let D be a domain in $\mathbb{R}^d$ and $(s,x) \in \mathbb{R}^+ \times D$. Assume that $X = (X(\cdot,t))_{t\geq 0}$ and $X^n = (X^n(\cdot,t))_{t\geq 0}$ are continuous d-dimensional stochastic processes on a stochastic base $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ such that $X(\cdot,s) = X^n(\cdot,s) = x$ for $n = 1,2,\ldots$ and $\sup_{t\geq 0}|X^n(\cdot,t) - X(\cdot,t)| \to 0$ a.s. as $n \to \infty$. If $\inf X_n^{-1}(\omega,\cdot)(D^{\sim}) < \infty$ for a.e. $\omega \in \Omega$ for $n = 1,2,\ldots$, then $\inf X^{-1}(\omega,\cdot)(D^{\sim}) < \infty$ for a.e. $\omega \in \Omega$.*

Proof. Let $\tau_n(\omega) = \inf X_n^{-1}(\omega,\cdot)(D^{\sim}) < \infty$ and $\tau(\omega) = \inf X^{-1}(\omega,\cdot)(D^{\sim})$ for $\omega \in \Omega$. Put $\Lambda = \{\omega \in \Omega : \tau(\omega) = \infty\}$ and $\Lambda_n = \{\omega \in \Omega : \tau_n(\omega) = \infty\}$ for $n = 1,2,\ldots$. For every $\omega \in \Lambda$, one has $X(\omega,t) \in D$ for $t \geq s$. By the properties of the sequence $(X_n)_{n=1}^{\infty}$ for a.e. fixed $\omega \in \Lambda$, there exists a positive integer $N(\omega) \geq 1$ such that $X_n(\omega,t) \in D$ for $t \geq s$ and every $n \geq N(\omega)$. Then for a.e. $\omega \in \Lambda$ and every $n \geq N(\omega)$, we have $\tau_n(\omega) = \infty$. For simplicity, assume that $\tau_n(\omega) = \infty$ for every $n \geq N(\omega)$ and $\omega \in \Lambda$. By the assumption that $\tau_n < \infty$ a.s. and the definition of Λ_n, we have $P(\Lambda_n) = 0$ for every $n \geq 1$. Then $P(\bigcup_{n=1}^{\infty}\Lambda_n) = 0$. But for every $\omega \in \Lambda$ and $n \geq N(\omega)$, we have $\tau_n(\omega) = \infty$. Therefore, $\Lambda \subset \bigcup_{n=1}^{\infty}\Lambda_n$. Then $P(\Lambda) = 0$. □

Lemma 5.5. *Let D be a domain in $\mathbb{R}^d$ and $(s,x) \in \mathbb{R}^+ \times D$. Assume that $X = (X(\cdot,t))_{t\geq 0}$ and $X^n = (X^n(\cdot,t))_{t\geq 0}$ are continuous d-dimensional stochastic processes on a stochastic base $\mathcal{P}_{\mathbb{F}} = (\Omega,\mathcal{F},\mathbb{F},P)$ such that $X(\cdot,s) = X^n(\cdot,s) = x$ for $n = 1,2,\ldots$ and $\sup_{t\geq 0}|X^n(\cdot,t) - X(\cdot,t)| \to 0$ a.s. as $n \to \infty$ and let $\tau_n(\omega) = \inf X_n^{-1}(\omega,\cdot)(\bar{D}^{\sim})$ and $\tau(\omega) = \inf X^{-1}(\omega,\cdot)(D^{\sim})$ for $\omega \in \Omega$. If $\max(\tau_n,\tau) < \infty$ a.s. for $n = 1,2,\ldots$, then there is a mapping $T : \Omega \to \mathbb{R}^+$ such that $\max(\tau_n,\tau) < T$ a.s. for $n \geq 1$.*

Proof. By virtue of Lemma 5.2, we have $\tau(\omega) = \inf(\mathrm{Li}\, X_n^{-1}(\omega,\cdot)(D^{\sim}))$ for a.e. $\omega \in \Omega$. By virtue of (ii) of Corollary 1.2 of Chap. 2, for a.e. $\omega \in \Omega$ there is $\bar{n} \geq 1$ such that for every $n > \bar{n}$, there exists $t_n \in X_n^{-1}(\omega,\cdot)(D^{\sim})$ such that $t_n \to \tau$ a.s. as $n \to \infty$. For every $n > \bar{n}$, we have $\tau_n \leq \tau$, because $X^{-1}(\omega,\cdot)(D^{\sim}) \subset X_n^{-1}(\omega,\cdot)(D^{\sim} + \varepsilon B)$, $\tau_n^{\varepsilon} \leq \tau$ and $\tau_n^{\varepsilon} \to \tau_n$ a.s. as $\varepsilon \to 0$, where $\tau_n^{\varepsilon}(\omega) = \inf X_n^{-1}(\omega,\cdot)(D^{\sim} + \varepsilon B)$ for $n \geq \bar{n}$. Then $\limsup \tau_n \leq \tau$ a.s., which implies that for a.e. $\omega \in \Omega$, there exists a positive integer $N(\omega) \geq 1$ such that $\tau_n(\omega) < \tau(\omega)$ for $n \geq N(\omega)$. Taking $T(\omega) = \max\{\tau_1(\omega)+1, \tau_2(\omega)+1, \ldots, \tau_{N(\omega)}(\omega)+1, \tau(\omega)+1\}$ for a.e. $\omega \in \Omega$, we have defined a mapping $T : \Omega \to \mathbb{R}^+$ such that $\max(\tau_n,\tau) < T$ a.s. for $n \geq 1$. □

Now we can prove the following convergence theorem.

Theorem 5.1. *Let D be a domain in $\mathbb{R}^d$ and $(s,x) \in \mathbb{R}^+ \times D$. Assume that $X = (X(\cdot,t))_{t\geq 0}$ and $X^n = (X^n(\cdot,t))_{t\geq 0}$ are continuous d-dimensional stochastic processes on a stochastic base $\mathcal{P}_{\mathbb{F}} = (\Omega,\mathcal{F},\mathbb{F},P)$ such that $X(\cdot,s) = X^n(\cdot,s) = x$ for $n = 1,2,\ldots$ and $\sup_{t\geq 0}|X^n(\cdot,t) - X(\cdot,t)| \to 0$ a.s. as $n \to \infty$. If $\tau_n = \inf\{r > s : X_n(\cdot,r) \notin D\} < \infty$ a.s. for $n = 1,2,\ldots$, then $\lim_{n\to\infty}\tau_n = \tau$ a.s., where $\tau = \inf\{r > s : X(\cdot,r) \notin D\}$.*

Proof. By virtue of Lemma 5.4, we have $\max(\tau_n,\tau) < \infty$ a.s. for $n = 1,2,\ldots$. Therefore, by virtue of Lemma 5.5, there is a mapping $T : \Omega \to \mathbb{R}^+$ such that $\max(\tau_n,\tau) < T$ a.s. for $n = 1,2,\ldots$. Then by virtue of Lemma 5.1, we have $\tau_n(\omega) = \inf(X_n^{-1}(\omega,\cdot)(D^{\sim}) \cap (s,T(\omega)))$ and $\tau(\omega) = \inf(X^{-1}(\omega,\cdot)(D^{\sim}) \cap (s,T(\omega)))$ for $\omega \in \Omega$ and $n = 1,2,\ldots$. By virtue of Lemma 5.3, Remark 1.2 of Chap. 2, and Theorem 1.1 of Chap. 2, we get

$$
\begin{aligned}
&\lim_{n\to\infty} h((X_n^{-1}(\omega,\cdot)(D^{\sim}) \cap (s,T(\omega))), X^{-1}(\omega,\cdot)(D^{\sim}) \cap (s,T(\omega)))) \\
&\quad = \lim_{n\to\infty} h(\overline{(X_n^{-1}(\omega,\cdot)(D^{\sim}) \cap (s,T(\omega)))}, \overline{X^{-1}(\omega,\cdot)(D^{\sim}) \cap (s,T(\omega))})) \\
&\quad = 0
\end{aligned}
$$

for a.e. $\omega \in \Omega$, where h is the Hausdorff metric on $\mathrm{Cl}([s,T(\omega)])$ for every fixed $\omega \in \Omega$. Let $\varepsilon > 0$ and $t_\varepsilon(\omega) \in X^{-1}(\omega,\cdot)(D^{\sim}) \cap (s,T(\omega)))$ be such that $t_\varepsilon(\omega) < \tau(\omega) + \varepsilon$ for fixed $\omega \in \Omega$. By the above property of the sequence $(X_n^{-1}(\omega,\cdot)(D^{\sim}) \cap (s,T(\omega)))_{n=1}^{\infty}$ and the definition of the Hausdorff metric h, we have $\mathrm{dist}(t_\varepsilon(\omega), X_n^{-1}(\omega,\cdot)(D^{\sim}) \cap (s,T(\omega))) \to 0$ for fixed $\omega \in \Omega$ and every $\varepsilon > 0$ as $n \to \infty$. Therefore, for every fixed $\omega \in \Omega$, there exists a sequence $(t_\varepsilon^n(\omega))_{n=1}^{\infty}$

such that $t_\varepsilon^n(\omega) \in X_n^{-1}(\omega,\cdot)(D^\sim) \cap (s, T(\omega))$ for $n \geq 1$ and $|t_\varepsilon^n(\omega) - t_\varepsilon(\omega)| \to 0$ as $n \to \infty$. Hence it follows that

$$\tau_n(\omega) \leq t_\varepsilon^n(\omega) \leq |t_\varepsilon^n(\omega) - t_\varepsilon(\omega)| + t_\varepsilon(\omega) < |t_\varepsilon^n(\omega) - t_\varepsilon(\omega)| + \tau(\omega) + \varepsilon$$

for $\varepsilon > 0$ and $n \geq 1$. Then $\limsup_{n\to\infty} \tau_n(\omega) \leq \tau(\omega)$.

Similarly, for fixed $\omega \in \Omega$ and every $\varepsilon > 0$ and $n \geq 1$, we can select $t_\varepsilon^n(\omega) \in X_n^{-1}(\omega,\cdot)(D^\sim) \cap [s, T(\omega))$ and $\bar{t}_\varepsilon^n \in X^{-1}(\omega,\cdot)(D^\sim) \cap [s, T(\omega))$ such that $t_\varepsilon^n(\omega) \leq \tau_n(\omega) + \varepsilon$ and $|\bar{t}_\varepsilon^n(\omega) - t_\varepsilon^n(\omega)| \to 0$ as $n \to \infty$. Hence it follows that

$$\tau(\omega) \leq \bar{t}_\varepsilon^n(\omega) \leq |\bar{t}_\varepsilon^n(\omega) - t_\varepsilon^n(\omega)| + t_\varepsilon^n(\omega) \leq |\bar{t}_\varepsilon^n(\omega) - t_\varepsilon^n(\omega)| + \tau_n(\omega) + \varepsilon$$

for every $\varepsilon > 0$ and $n \geq 1$. Therefore, $\tau(\omega) \leq \liminf_{n\to\infty} \tau_n(\omega)$. Then $\limsup_{n\to\infty} \tau_n(\omega) \leq \tau(\omega) \leq \liminf_{n\to\infty} \tau_n(\omega)$ for a.e. $\omega \in \Omega$, which implies that $\lim_{n\to\infty} \tau_n = \tau$ a.s. □

Let D be a domain in $\mathbb{R}^d$ and $(s,x) \in \mathbb{R}^+ \times D$. Assume that $X = (X(t))_{t\geq 0}$ and $\tilde{X} = (\tilde{X}(t))_{t\geq 0}$ are continuous d-dimensional stochastic processes on $(\Omega, \mathcal{F}, P)$ and $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$, respectively, such that $X(s) = x$ a.s. and $PX^{-1} = P\tilde{X}^{-1}$. We shall show that $P(\tau_D)^{-1} = P(\tilde{\tau}_D)^{-1}$, $P(X \circ \tau_D)^{-1} = P(\tilde{X} \circ \tilde{\tau}_D)^{-1}$, and $P(\tau_D, X \circ \tau_D)^{-1} = P(\tilde{\tau}_D, \tilde{X} \circ \tilde{\tau}_D)^{-1}$, where $\tau_D = \inf\{t > s : X_t \notin D\}$ and $\tilde{\tau}_D = \inf\{t > s : \tilde{X}_t \notin D\}$.

The next results will follow from the following fundamental lemma, similar to Lemma 2.1 of Chap. 1.

Lemma 5.6. *Let X and $\tilde{X}$ be as above, $(Y, \mathcal{G})$ a measurable space, and $C =: C(\mathbb{R}^+, \mathbb{R}^d)$. If $\Phi : C \to Y$ is $(\beta, \mathcal{G})$-measurable, where β is a Borel σ-algebra on C, 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 $P(\Phi \circ X)^{-1} = P(\Phi \circ \tilde{X})^{-1}$. □

The following theorem can be derived immediately from the above result.

Theorem 5.2. *Let D be a domain in $\mathbb{R}^d$ and $(s,x) \in \mathbb{R}^+ \times D$. Assume that $X = (X(t))_{t\geq 0}$ and $\tilde{X} = (\tilde{X}(t))_{t\geq 0}$ are continuous d-dimensional stochastic processes on $(\Omega, \mathcal{F}, P)$ and $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$, respectively, such that $X(s) = x$ a.s. and $PX^{-1} = P\tilde{X}^{-1}$. Then $P(\tau_D)^{-1} = P(\tilde{\tau}_D)^{-1}$, $P(X \circ \tau_D)^{-1} = P(\tilde{X} \circ \tilde{\tau}_D)^{-1}$, and $P(\tau_D, X \circ \tau_D)^{-1} = P(\tilde{\tau}_D, \tilde{X} \circ \tilde{\tau}_D)^{-1}$, where $\tau_D = \inf\{t > s : X_t \notin D\}$ and $\tilde{\tau}_D = \inf\{t > s : \tilde{X}_t \notin D\}$.*

Proof. Let $\eta : C \to \mathbb{R}^+$ be defined by $\eta(x) = \inf\{t > s : x(t) \notin D\}$ for $x \in C$. It is clear that η is (β, β_+)-measurable, where β_+ denotes the Borel σ-algebra on $\mathbb{R}^+$. Taking $Y = \mathbb{R}^+$, $\mathcal{G} = \beta_+$, and $\Phi = \eta$, we get $\tau_D = \Phi \circ X$ and $\tilde{\tau}_D = \Phi \circ \tilde{X}$. Therefore, by virtue of Lemma 5.6, we obtain $P(\tau_D)^{-1} = P(\tilde{\tau}_D)^{-1}$. Let $\psi(t,x) = x(t)$ for $x \in C$ and $t \in \mathbb{R}^+$ and put $\Phi(x) = \psi(\eta(x), x))$ for $x \in C$. It is clear that the mapping Φ satisfies the conditions of Lemma 5.6 with $Y = \mathbb{R}^d$

and $\mathcal{G} = \beta$, where β denotes the Borel σ-algebra on $\mathbb{R}^d$. Furthermore, we have $\Phi \circ X = X \circ \tau_D$ and $\Phi \circ \tilde{X} = \tilde{X} \circ \tilde{\tau}_D$. Therefore, by virtue of Lemma 5.6, we obtain $P(X \circ \tau_D)^{-1} = P(\tilde{X} \circ \tilde{\tau}_D)^{-1}$. Finally, let $\Phi(x) = (\eta(x), \psi(\eta(x), x))$ for $x \in C$. Immediately from the properties of the mappings ψ and η, it follows that Φ satisfies the conditions of Lemma 5.6 with $Y = \mathbb{R}^+ \times \mathbb{R}^d$ and $\mathcal{G} = \beta_+ \times \beta$, where β_+ denotes the Borel σ-algebra of $\mathbb{R}^+$. Furthermore, $\Phi \circ X = (\tau_D, X \circ \tau_D)$ and $\Phi \circ \tilde{X} = (\tilde{\tau}_D, \tilde{X} \circ \tilde{\tau}_D)$, which by virtue of Lemma 5.6, implies $P(\tau_D, X \circ \tau_D)^{-1} = P(\tilde{\tau}_D, \tilde{X} \circ \tilde{\tau}_D)^{-1}$. □

Corollary 5.1. *If the assumptions of Theorem 5.2 are satisfied, then for every continuous bounded function* $f : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}$, *one has* $E[f(\tau_D, X \circ \tau_D)] = \tilde{E}[f(\tilde{\tau}_D, \tilde{X} \circ \tilde{\tau}_D)]$, *where* E *and* $\tilde{E}$ *denote the mean value operators with respect to probability measures* P *and* $\tilde{P}$, *respectively.* □

6 Notes and Remarks

The first papers concerning stochastic functional inclusions written in the set-valued integral form are due to Hiai [38] and Kisielewicz [51, 55], where stochastic functional inclusions containing set-valued stochastic integrals were independently investigated. In the above papers, only strong solutions were considered. An extension of the Fillipov theorem for stochastic differential inclusions was given by Da Prato and Frankowska [23]. Existence and stability of solutions of stochastic differential inclusions were considered by Motyl in [82] and [83], resp. Weak solutions of stochastic functional inclusions have been considered by Aubin and Da Prato [9], Kisielewicz [53] and Levakov [71]. Weak compactness with respect to convergence in distribution of solution sets of weak solutions of stochastic differential inclusions was considered in Kisielewicz [56, 58, 60]. Also, Levakov in [71] considered weak compactness of all distributions of weak solutions of some special type of stochastic differential inclusions. Compactness of solutions of second order dynamical systems was considered by Michta and Motyl in [78]. The results of the last three sections of this chapter are based on Kisielewicz [56, 58], where stochastic functional inclusions in the finite intervals $[0, T]$ are considered. The results dealing with backward stochastic differential inclusions were first considered in the author's paper [59]. The results contained in Sect. 5 are taken entirely from Kisielewicz [55]. The properties of stochastic differential inclusions presented in Sect. 2 are the first dealing with such inclusions. By Theorem 2.1 of Chap. 3, stochastic differential inclusions $SDI(F, G)$ are equivalent to stochastic functional inclusions of the form $x_t - x_s \in \overline{\text{dec}}\{J(F \circ x)\} + \overline{\text{dec}}\{\mathcal{J}(G \circ x)\}$. Therefore, for multifunctions F and G satisfying the assumptions of Theorem 1.5, the set $S_w(F, G, \mu)$ of all weak solutions of $SDI(F, G)$ with an initial distribution μ contains a set considered in optimal control problems described by $SDI(F, G)$.

For the existence of solutions of such optimal control problems, it is necessary to have some sufficient conditions implying weak compactness of a solution set $S_w(F, G, \mu)$. Such results are difficult to obtain by the methods used in the proof of Theorem 4.1, because boundedness or square integrable boundedness of $\overline{\text{dec}}\{J(F \circ x)\}$ and $\overline{\text{dec}}\{\mathcal{J}(G \circ x)\}$ is necessary in such a proof.

Chapter 5
Viability Theory

The results of this chapter deal with the existence of viable solutions for stochastic functional and backward inclusions. Weak compactness of sets of all viable weak solutions of stochastic functional inclusions is also considered.

1 Some Properties of Set-Valued Stochastic Functional Integrals Depending on Parameters

Let $F : [0, T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : [0, T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ be measurable and square integrably bounded set-valued mappings. Given a set-valued stochastic process $(K(t))_{0\le t\le T}$ with values in $\mathrm{Cl}(\mathbb{R}^d)$, we denote by $\overline{SFI}(F, G, K)$ the following viability problem:

$$\begin{cases} x_t - x_s \in \mathrm{cl}_{\mathbb{L}}\{J_{st}[S_{\mathbb{F}}(F \circ x)] + \mathcal{J}_{st}[S_{\mathbb{F}}(G \circ x)]\} & \text{for} \quad 0 \le s \le t \le T, \\ x_t \in K(t) \;\; \text{a.s.} \;\; \text{for} \;\; t \in [0, T], \end{cases} \tag{1.1}$$

associated with $\overline{SFI}(F, G)$. Similarly, we denote by $BSDI(F, K)$ the backward viability problem:

$$\begin{cases} x_s \in E[x_t + \int_s^t F(\tau, x_\tau)\mathrm{d}\tau|\mathcal{F}_s] & \text{a.s.} \;\; \text{for} \quad 0 \le s \le t \le T, \\ x_t \in K(t) \quad \text{a.s.} \;\; \text{for} \quad t \in [0, T], \end{cases} \tag{1.2}$$

associated with $BSDI(F, K(T))$.

We precede the existence theorems for such problems by some properties of set-valued stochastic functional integrals depending on parameters. Given a Banach space $(X, \|\cdot\|)$, by $\mathrm{Cl}(X)$ we denote the space of all nonempty closed subsets

M. Kisielewicz, *Stochastic Differential Inclusions and Applications*, Springer Optimization and Its Applications 80, DOI 10.1007/978-1-4614-6756-4_5,

of X. In particular, we shall consider X to be equal to $\mathbb{R}^d$, $\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{R}^r)$, and $\mathbb{L}^2([0,T] \times \Omega, \Sigma_{\mathbb{F}}, \mathbb{R}^r$ with $r = d$ and $r = d \times m$, respectively. The Hausdorff metrics on these spaces will be denoted by h, D, and H, respectively.

Let $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a filtered probability space with a filtration $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$ satisfying the usual conditions. Similarly as above, for set-valued mappings F and G as given above and an $\mathbb{F}$-nonanticipative d-dimensional stochastic process $x = (x_t)_{0 \leq t \leq T}$, we shall denote by $S_{\mathbb{F}}(F \circ x)$ and $S_{\mathbb{F}}(G \circ x)$ the sets of all $\mathbb{F}$-nonanticipative stochastic processes $f = (f_t)_{0 \leq t \leq T}$ and $g = (g_t)_{0 \leq t \leq T}$, respectively, such that $f_t \in F(t, x_t)$ and $g_t \in G(t, x_t)$ a.s. for a.e. $t \in [0,T]$. It is clear that $S_{\mathbb{F}}(F \circ x)$ and $S_{\mathbb{F}}(G \circ x)$ are decomposable closed subsets of $\mathbb{L}^2([0,T] \times \Omega, \Sigma_{\mathbb{F}}, \mathbb{R}^d)$ and $\mathbb{L}^2([0,T] \times \Omega, \Sigma_{\mathbb{F}}, \mathbb{R}^{d \times m})$, respectively, where $\Sigma_{\mathbb{F}}$ denotes the σ-algebra of all $\mathbb{F}$-nonanticipative subsets of $[0,T] \times \Omega$. Therefore, by virtue of Theorem 3.2 of Chap. 2, there exist $\Sigma_{\mathbb{F}}$-measurable mappings Φ and Ψ such that $S_{\mathbb{F}}(F \circ x) = S_{\mathbb{F}}(\Phi)$ and $S_{\mathbb{F}}(G \circ x) = S_{\mathbb{F}}(\Psi)$, which by virtue of Corollary 3.1 of Chap. 2, implies that $\Phi = F \circ x$ and $\Psi = G \circ x$.

In what follows, we shall denote by $|\cdot|$ the norm of the Banach space $\mathcal{X}^r = \mathbb{L}^2([0,T] \times \Omega, \Sigma_{\mathbb{F}}, \mathbb{R}^r)$ with $r = d$ or $r = d \times m$. Similarly as above, $\mathbf{C}(\mathbb{F}, \mathbb{R}^d)$ denotes the space of all d-dimensional continuous $\mathbb{F}$-adapted stochastic processes $x = (x_t)_{0 \leq t \leq T}$ with norm $\|x\| = (E[\sup_{0 \leq t \leq T} |x_t|^2])^{1/2}$. Given a measurable and uniformly square integrably bounded set-valued mapping $K : [0,T] \times \Omega \to \mathrm{Cl}(\mathbb{R}^d)$, we shall assume that the set $\mathcal{K}(t) = \{u \in \mathbb{L}^2(\Omega, \mathcal{F}_t, \mathbb{R}^d) : u \in K(t, \cdot) \ a.s.\}$ is nonempty for every $0 \leq t \leq T$. It is clear that this requirement is satisfied for a square integrably bounded multifunction $K : [0,T] \to \mathrm{Cl}(\mathbb{R}^d)$. Recall that $K : [0,T] \times \Omega \to \mathrm{Cl}(\mathbb{R}^d)$ is said to be uniformly square integrably bounded if there exists $\lambda \in \mathbb{L}^2([0,T], \mathbb{R}^+)$ such that $\|K(t,\omega)\| \leq \lambda(t)$ for a.e. $(t,\omega) \in [0,T] \times \Omega$, where $\|K(t,\omega)\| = h(K(t,\omega), \{0\})$. Let us observe that for the above multifunctions F and G and a d-dimensional $\mathcal{F}_t$-measurable random variable X, the set-valued processes $F \circ X$ and $G \circ X$ are $\beta_T \otimes \mathcal{F}_t$-measurable.

Assume that the above set-valued mappings F and G satisfy the following conditions $(\mathcal{H}_1)$:

(i) $F : [0,T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : [0,T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d \times m})$ are measurable and uniformly square integrably bounded, i.e., there exists $m \in \mathbb{L}^2([0,T], \mathbb{R}^+)$ such that $\max(\|F(t,x)\|, \|G(t,x)\|) \leq m(t)$ for a.e. $t \in [0,T]$ and $x \in \mathbb{R}^d$, where $\|F(t,x)\| = \sup\{|z| : z \in F(t,x)\}$ and $\|G(t,x)\| = \sup\{|z| : z \in G(t,x)\}$;
(ii) $F(t,\cdot)$ and $G(t,\cdot)$ are Lipschitz continuous for a.e. fixed $t \in [0,T]$, i.e., there exists $k \in \mathbb{L}^2([0,T], \mathbb{R}^+)$ such that $H(F(t,x), F(t,z)) \leq k(t)|x - z|$ and $H(G(t,x), G(t,z)) \leq k(t)|x - z|$ for a.e. $t \in [0,T]$ and $x, z \in \mathbb{R}^d$.

Lemma 1.1. *If F and G satisfy conditions $(\mathcal{H}_1)$, then the set-valued mappings $\mathbf{C}(\mathbb{F}, \mathbb{R}^d) \ni x \to S_{\mathbb{F}}(F \circ x) \in \mathrm{Cl}(\mathcal{X}^d)$ and $\mathbf{C}(\mathbb{F}, \mathbb{R}^d) \ni x \to S_{\mathbb{F}}(G \circ x) \in \mathrm{Cl}(\mathcal{X}^{d \times m})$ are Lipschitz continuous with Lipschitz constant $L = [\int_0^T k^2(t) \mathrm{d}t]^{1/2}$.*

Proof. The proof is quite similar to the proof of Lemma 3.7 of Chap. 2. Let $x, z \in \mathbf{C}(\mathbb{F}, \mathbb{R}^d)$ and $f^x \in S_{\mathbb{F}}(F \circ x)$. By virtue of Theorem 3.1 of Chap. 2 applied to the $\Sigma_{\mathbb{F}}$-measurable set-valued mapping $F \circ z$, we get

$$\begin{aligned}\operatorname{dist}^2(f^x, S_{\mathbb{F}}(F \circ z)) &= \inf\left\{E\int_0^T |f_\tau^x - f_\tau|^2 d\tau : f \in S_{\mathbb{F}}(F \circ z)\right\}\\ &= E\int_0^T \operatorname{dist}^2(f_\tau^x, F(\tau, z_\tau))d\tau\\ &\le E\int_0^T k^2(t)|x_t - z_t|^2 dt \le L^2\|x - z\|^2.\end{aligned}$$

Then $\overline{H}(S_{\mathbb{F}}(F \circ x), S_{\mathbb{F}}(F \circ z)) \le L\|x - z\|$. In a similar way, we also get $\overline{H}(S_{\mathbb{F}}(F \circ z), S_{\mathbb{F}}(F \circ x)) \le L||x - z||$. Therefore, $H(S_{\mathbb{F}}(F \circ x), S_{\mathbb{F}}(F \circ z)) \le L\|x - z\|$. In a similar way, we obtain $H(S_{\mathbb{F}}(G \circ x), S_{\mathbb{F}}(G \circ z)) \le L\|x - z\|$. □

Lemma 1.2. *Let $K : [0, T] \times \Omega \to \mathrm{Cl}(\mathbb{R}^d)$ be $\mathbb{F}$-adapted and square integrably bounded uniformly with respect to $t \in [0, T]$. If $K(\cdot, \omega)$ is continuous for a.e. $\omega \in \Omega$, then the set-valued mapping $\mathcal{K} : [0, T] \to \mathrm{Cl}(\mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d))$ is continuous.*

Proof. Let $t_0 \in [0, T]$ be fixed and let $(t_k)_{k=1}^{\infty}$ be a sequence of $[0, T]$ converging to t_0. By virtue of Theorem 3.1 of Chap. 2, for every $u \in \mathcal{K}(t_0)$ and $k \ge 1$, one has

$$\begin{aligned}\operatorname{dist}^2(u, \mathcal{K}(t_k)) &= \inf\left\{E|u - v|^2 : v \in \mathcal{K}(t_k)\right\}\\ &\le E\left[\operatorname{dist}^2(u, K(t_k, \cdot)\right]\\ &\le E\left[h^2(K(t_k, \cdot), K(t_0, \cdot))\right].\end{aligned}$$

Then $\overline{D}^2(\mathcal{K}(t_0), \mathcal{K}(t_k)) \le E\left[h^2(K(t_k, \cdot), K(t_0, \cdot))\right]$. In a similar way, we also get $\overline{D}^2(\mathcal{K}(t_k), \mathcal{K}(t_0)) \le E\left[h^2(K(t_k, \cdot), K(t_0, \cdot))\right]$. Therefore, for every $k \ge 1$, one has $D^2(\mathcal{K}(t_k), \mathcal{K}(t_0)) \le E\left[h^2(K(t_k, \cdot), K(t_0, \cdot))\right]$. Hence, by the continuity of $K(\cdot, \omega)$ and its uniformly square integrable boundedness, it follows that $\lim_{k\to\infty} D(\mathcal{K}(t_k), \mathcal{K}(t_0)) = 0$. □

Lemma 1.3. *If F and G satisfy conditions $(\mathcal{H}_1)$, then the set-valued mappings $\mathbf{C}(\mathbb{F}, \mathbb{R}^d) \ni x \to \mathrm{cl}_{\mathbb{L}}\{J_{st}[S_{\mathbb{F}}(F \circ x)]\} \subset \mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$ and $\mathbf{C}(\mathbb{F}, \mathbb{R}^d) \ni x \to \mathrm{cl}_{\mathbb{L}}\{\mathcal{J}_{st}[S_{\mathbb{F}}(G \circ x)]\} \subset \mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$ are Lipschitz continuous uniformly with respect to $0 \le s < t \le T$ with Lipschitz constants equal to $\sqrt{T}L$ and L, respectively, where L is as in Lemma 1.1.*

Proof. Let $x, z \in \mathcal{C}(\mathbb{F}, \mathbb{R}^d)$ and $f^x \in S_{\mathbb{F}}(F \circ x)$. For fixed $0 \le s < t \le T$, we have $\operatorname{dist}^2(J_{st}(f^x), J_{st}[S_{\mathbb{F}}(F \circ z]) = \inf\left\{E|J_{st}(f^x - f^z)|^2 : f^z \in S_{\mathbb{F}}(F \circ z)\right\}$. But for every $0 \le s < t \le T$, one has

$$E\,|J_{st}(f^x - f^z)|^2 \le TE\left[\int_0^T |f^x - f^z|^2 dt\right].$$

Therefore, by Lemma 3.6 of Chap. 2, it follows that

$$\begin{aligned}\operatorname{dist}^2\left(J_{st}(f^x), J_{st}[S_{\mathbb{F}}(F\circ z)]\right) &\leq T\inf\left\{E\int_0^T |f^x - f^z|^2 \mathrm{d}t : f^z \in S_{\mathbb{F}}(F\circ z)\right\}\\ &= T\operatorname{dist}^2\left(f^x, S_{\mathbb{F}}(F\circ z)\right)\\ &\leq TH(S_{\mathbb{F}}(F\circ x), S_{\mathbb{F}}(F\circ z)) \leq TL^2\|x-z\|^2.\end{aligned}$$

Then for every $0 \leq s < t \leq T$, one obtains

$$\overline{D}^2\left(J_{st}[S_{\mathbb{F}}(F\circ x)], J_{st}[S_{\mathbb{F}}(F\circ z)]\right) \leq TL^2\|x-z\|^2.$$

Similarly, for every fixed $0 \leq s < t \leq T$, we also get

$$\overline{D}^2\left(J_{st}[S_{\mathbb{F}}(F\circ z)], J_{st}[S_{\mathbb{F}}(F\circ x)]\right) \leq TL^2\|x-z\|^2.$$

Therefore, for every $0 \leq s < t \leq T$, one has

$$D\left(J_{st}[S_{\mathbb{F}}(F\circ x)], J_{st}[S_{\mathbb{F}}(F\circ z)]\right) \leq \sqrt{T}L\|x-z\|.$$

In a similar way, for fixed $0 \leq s < t \leq T$, we obtain

$$D\left(\mathcal{J}_{st}[S_{\mathbb{F}}(G\circ x)], \mathcal{J}_{st}[S_{\mathbb{F}}(G\circ z)]\right) \leq L\|x-z\|.$$

Hence it follows that

$$\sup_{0\leq s<t\leq T} D\left(\mathrm{cl}_{\mathbb{L}}\{J_{st}[S_{\mathbb{F}}(F\circ x)]\}, \mathrm{cl}_{\mathbb{L}}\{J_{st}[S_{\mathbb{F}}(F\circ z)]\}\right) \leq \sqrt{T}L\|x-z\|$$

and

$$\sup_{0\leq s<t\leq T} D\left(\mathrm{cl}_{\mathbb{L}}\{\mathcal{J}_{st}[S_{\mathbb{F}}(G\circ x)]\}, \mathrm{cl}_{\mathbb{L}}\{\mathcal{J}_{st}[S_{\mathbb{F}}(G\circ z)]\}\right) \leq L\|x-z\|.$$

□

Lemma 1.4. *Assume that F and G satisfy (i) of $(\mathcal{H}_1)$ and let $x_n, x \in \mathbf{C}(\mathbb{F}, \mathbb{R}^d)$ for $n = 1, 2, \ldots$ be such that $\sup_{0\leq t\leq T}|x_n(t) - x(t)| \to 0$ a.s. as $n \to \infty$. If $F(t,\cdot)$ and $G(t,\cdot)$ are continuous for a.e. fixed $0 \leq t \leq T$, and $(\theta_n)_{n=1}^{\infty}$ is a sequence of functions $\theta_n : [0,T] \to [0,T]$ such that $\theta_n(t) \to t$ as $n \to \infty$ for every $t \in [0,T]$, then $\mathrm{cl}_{\mathbb{L}}\{J_{st}[S_{\mathbb{F}}(F\circ(x_n\circ\theta_n))] + \mathcal{J}_{st}[S_{\mathbb{F}}(G\circ(x_n\circ\theta_n))]\} \to \mathrm{cl}_{\mathbb{L}}\{J_{st}[S_{\mathbb{F}}(F\circ x)] + \mathcal{J}_{st}[S_{\mathbb{F}}(G\circ x)]\}$ in the D-metric topology of $\mathrm{Cl}(\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{R}^d)$ as $n \to \infty$ for every $0 \leq s \leq t \leq T$.*

Proof. Let $0 \leq s \leq t \leq T$ be fixed and set $y^n = x_n\circ\theta_n$ for every $n = 1, 2, \ldots$. One has

$$\begin{aligned}|y_t^n(\tau)-x(\tau)| &= |x_n(\theta_n(\tau))-x(\tau)|\\ &\le |x_n(\theta_n(\tau))-x(\theta_n(\tau))|+|x(\theta_n(\tau))-x(\tau)|\\ &\le \sup_{0\le u\le T}|x_n(u)-x(u)|+|x(\theta_n(t))-x(t)|\end{aligned}$$

for $n=1,2,\ldots$ and $0\le\tau\le T$. Then $y_t^n(\tau)\to x(\tau)$ a.s. for every $0\le\tau\le T$ as $n\to\infty$. Similarly as in the proof of Lemma 1.3, we can verify that the set-valued mappings $\mathbf{C}(\mathbb{F},\mathbb{R}^d)\ni x\to \mathrm{cl}_{\mathbb{L}}\{J_{st}[S_{\mathbb{F}}(F\circ x)]\}\in \mathrm{Cl}(\mathbb{L}^2(\Omega,\mathcal{F},\mathbb{R}^d))$ and $\mathbf{C}(\mathbb{F},\mathbb{R}^d)\ni x\to \mathrm{cl}_{\mathbb{L}}\{\mathcal{J}_{st}[S_{\mathbb{F}}(G\circ x)]\}\in \mathrm{Cl}(\mathbb{L}^2(\Omega,\mathcal{F},\mathbb{R}^d))$ are continuous. Therefore, $\mathrm{cl}_{\mathbb{L}}\{J_{st}[S_{\mathbb{F}}(F\circ(x_n\circ\theta_n))]+\mathcal{J}_{st}[S_{\mathbb{F}}(G\circ(x_n\circ\theta_n))]\}\to \mathrm{cl}_{\mathbb{L}}\{J_{st}[S_{\mathbb{F}}(F\circ x)]+\mathcal{J}_{st}[S_{\mathbb{F}}(G\circ x)]\}$ in the D-metric topology as $n\to\infty$. □

2 Viable Approximation Theorems

The existence of solutions of viability problems (1.1) and (1.2) will follow from some viable approximation theorems by applying the standard methods presented in the proofs of the existence of strong and weak solutions for stochastic functional inclusions. We shall now present such approximation theorems. In what follows, it will be convenient to denote by $d(x,A)$ the distance $\mathrm{dist}(x,A)$ of $x\in X$ to a nonempty set $A\subset X$. We shall also denote the set-valued functional integrals $J_{st}[S_{\mathbb{F}}\Phi)]$ and $\mathcal{J}_{st}[S_{\mathbb{F}}\Psi)]$ of $\mathbb{F}$-nonanticipative set-valued processes $\Phi\in\mathcal{L}^2_{\mathbb{F}}(T,\Omega,\mathbb{R}^d)$ and $\Psi\in\mathcal{L}^2_{\mathbb{F}}(T,\Omega,\mathbb{R}^{d\times m})$ by $\int_s^t\Phi_\tau \mathrm{d}\tau$ and $\int_s^t\Psi_\tau \mathrm{d}B_\tau$, respectively.

We shall prove the following approximation theorems.

Theorem 2.1. *Assume that F and G satisfy condition (i) of $(\mathcal{H}_1)$ and let $\mathcal{P}_{\mathbb{F}}=(\Omega,\mathcal{F},\mathbb{F},P)$ be a complete filtered probability space with a filtration $\mathbb{F}=(\mathcal{F}_t)_{0\le t\le T}$ such that there exists an m-dimensional $\mathbb{F}$-Brownian motion $B=(B_t)_{0\le t\le T}$ defined on $\mathcal{P}_{\mathbb{F}}$. Let $K:[0,T]\to \mathrm{Cl}(\mathbb{R}^d)$ be such that a set-valued process $(\mathcal{K}(t))_{0\le t\le T}$ is continuous. If*

$$\liminf_{h\to 0+}\frac{1}{h}\overline{D}\left[x+\mathrm{cl}_{\mathbb{L}}\left(\int_t^{t+h}F(\tau,x)\mathrm{d}\tau+\int_t^{t+h}G(\tau,x)\mathrm{d}B_\tau\right),\mathcal{K}(t+h)\right]=0 \tag{2.1}$$

for every $(t,x)\in Graph(\mathcal{K}^\varepsilon)$ and every $\varepsilon\in(0,1)$, where $\mathcal{K}^\varepsilon(t)=\{u\in\mathbb{L}^2(\Omega,\mathcal{F}_t,\mathbb{R}^d):d(u,\mathcal{K}(t))\le\varepsilon\}$ for every $0\le t\le T$, then for every $\varepsilon\in(0,1)$ and $x_0\in\mathcal{K}(0)$, there exist a step function $\theta_\varepsilon:[0,T]\to[0,T]$ and $\mathbb{F}$-nonanticipative stochastic processes $f^\varepsilon=(f_t^\varepsilon)_{0\le t\le T}$ and $g^\varepsilon=(g_t^\varepsilon)_{0\le t\le T}$ such that

(i) $f^\varepsilon\in S_{\mathbb{F}}(F\circ(x^\varepsilon\circ\theta_\varepsilon))$ and $g^\varepsilon\in S_{\mathbb{F}}(G\circ(x^\varepsilon\circ\theta_\varepsilon))$, where $x^\varepsilon(t)=x_0+\int_0^t f_\tau^\varepsilon\mathrm{d}\tau+\int_0^t g_\tau^\varepsilon\mathrm{d}B_\tau$ for $0\le t\le T$;
(ii) $E[\mathrm{dist}(x^\varepsilon(\theta_\varepsilon(t)),K(\theta_\varepsilon(t))]\le\varepsilon$ for $0\le t\le T$;

(iii) $E\left[l(x^{\varepsilon}(s))\left(h(x^{\varepsilon}(t))-h(x^{\varepsilon}(s))-\int_s^t(\mathbb{L}^{x^{\varepsilon}}_{f^{\varepsilon}g^{\varepsilon}}h)_{\tau}\mathrm{d}\tau\right)\right]=0$ *for every* $0\leq s\leq t\leq T$, $l\in C_b(\mathbb{R}^d,\mathbb{R})$ *and* $h\in C_b^2(\mathbb{R}^d,\mathbb{R})$.

Proof. Let $\varepsilon\in(0,1)$ and $x_0\in\mathcal{K}(0)$ be fixed. Select $\delta\in(0,\varepsilon)$ such that $\int_t^{t+\delta}m^2(\tau)\mathrm{d}\tau\leq\varepsilon^2/2^4$ and $D(\mathcal{K}(t+\delta),\mathcal{K}(t))\leq\varepsilon/2^2$ for $t\in[0,T]$. By virtue of (2.1), there exists $h_0\in(0,\delta)$ such that

$$\overline{D}\left[x_0+\mathrm{cl}_{\mathbb{L}}\left(\int_0^{h_0}F(\tau,x_0)\mathrm{d}\tau+\int_0^{h_0}G(\tau,x_0)\mathrm{d}B_{\tau}\right),\mathcal{K}(h_0)\right]\leq\frac{\varepsilon h_0}{2^2}.$$

Then for every $u_0\in x_0+cl_{\mathbb{L}}\left(\int_0^{h_0}F(\tau,x_0)\mathrm{d}\tau+\int_0^{h_0}G(\tau,x_0)\mathrm{d}B_{\tau}\right)$, one has $d(u_0,\mathcal{K}(h_0))\leq\varepsilon h_0/2^2$. Let $t_0=0$ and $t_1=h_0$. Select arbitrarily $\beta_T\otimes\mathcal{F}_0$-measurable selectors f^0 and g^0 of $F\circ x_0$ and $G\circ x_0$, respectively. It is clear that $f^0\in S_{\mathbb{F}}(F\circ x_0))$ and $g^0\in S_{\mathbb{F}}(G\circ x_0))$. Let $x^{\varepsilon}(t)=x_0+\int_0^t f_{\tau}^0\mathrm{d}\tau+\int_0^t g_{\tau}^0\mathrm{d}B_{\tau}$ for $0\leq t\leq t_1$. Put $\theta_{\varepsilon}(t)=0$ for $0\leq t<t_1$ and $\theta_{\varepsilon}(t_1)=t_1$. We have

$$x^{\varepsilon}(h_0)\in x_0+\mathrm{cl}_{\mathbb{L}^2}\left(\int_0^{h_0}F(\tau,x_0)\mathrm{d}\tau+\int_0^{h_0}G(\tau,x_0)\mathrm{d}B_{\tau}\right).$$

Therefore, $d(x^{\varepsilon}(h_0),\mathcal{K}(h_0))\leq\varepsilon h_0/2^2\leq\varepsilon/2^2$. Together with the properties of the number $\delta>0$, it follows that

$$\begin{aligned}d(x^{\varepsilon}(t),\mathcal{K}(t))&\leq\|x^{\varepsilon}(t)-x^{\varepsilon}(h_0)\|+d(x^{\varepsilon}(h_0),\mathcal{K}(h_0))\\&\leq\varepsilon/2+\varepsilon h_0/2^2+D(\mathcal{K}(h_0),\mathcal{K}(t))\leq\varepsilon\end{aligned}$$

for $0\leq t\leq t_1$, because

$$\begin{aligned}&\|x^{\varepsilon}(t)-x^{\varepsilon}(h_0)\|\\&\quad\leq\sqrt{h_0}\left[E\int_0^{h_0}|f_{\tau}^0|^2\mathrm{d}\tau\right]^{1/2}+\left[E\int_0^{h_0}|g_{\tau}^0|^2\mathrm{d}\tau\right]^{1/2}\leq 2\varepsilon/2^2=\varepsilon/2\end{aligned}$$

for $0\leq t\leq t_1$. Let $x_1\in\mathcal{K}(t_1)$ be such that $\|x^{\varepsilon}(h_0)-x_1\|\leq d(x^{\varepsilon}(h_0),\mathcal{K}(h_0))+\varepsilon/2^2$. Hence, by Theorem 3.1 of Chap. 2, it follows that

$$E[\mathrm{dist}(x^{\varepsilon}(h_0),K(h_0))]=\inf\{E|x^{\varepsilon}(h_0)-u|:u\in\mathcal{K}(h_0)\}\leq$$

$$E[|x^{\varepsilon}(h_0)-x_1|]\leq(E[|x^{\varepsilon}(h_0)-x_1|^2])^{1/2}=||x^{\varepsilon}(h_0)-x_1||\leq\varepsilon/2^2+\varepsilon/2^2\leq\varepsilon.$$

By Itô's formula, for every $h\in C_b^2(\mathbb{R}^d,\mathbb{R})$ and $0\leq s\leq t\leq T$, we have

$$h(x^{\varepsilon}(t))-h(x^{\varepsilon}(s))-\int_s^t(\mathbb{L}^{x^{\varepsilon}}_{f^0g^0}h)_{\tau}\mathrm{d}\tau=\sum_{i=1}^n\sum_{j=1}^m\int_s^t h''_{x_ix_j}(x^{\varepsilon}(\tau))g_{ij}^0(\tau)\mathrm{d}B_{\tau}^j,$$

a.s., where $B = (B^1, \ldots, B^m)$ and $g_\tau^0 = (g_{ij}^0(\tau))_{d\times m}$. But $x^\varepsilon(s)$ is $\mathcal{F}_s$-measurable. Then for $0 \le s \le t \le t_2$, $i = 1, 2, \ldots, n$, and $j = 1, 2, \ldots, m$, we have

$$E\left[l(x^\varepsilon(s))\int_s^t h''_{x_i x_j}(x^\varepsilon(\tau))g_{ij}^0(\tau)\mathrm{d}B_\tau^j\right] = E\left[\int_s^t l(x^\varepsilon(s))h''_{x_i x_j}(x^\varepsilon(\tau))g_{ij}^0(\tau)\mathrm{d}B_\tau^j\right] = 0.$$

Therefore, for every $l \in C_b(\mathbb{R}^d, \mathbb{R})$, $h \in C_b^2(\mathbb{R}^d, \mathbb{R})$, and $0 \le s \le t \le t_1$, we get

$$E\left[l(x^\varepsilon(s))\left(h(x^\varepsilon(t)) - h(x^\varepsilon(s)) - \int_s^t (\mathbb{L}_{f^0 g^0}^{x^\varepsilon} h)_\tau \mathrm{d}\tau\right)\right] = 0.$$

Suppose $h_0 < T$. We have $(h_0, x^\varepsilon(h_0)) \in Graph(\mathcal{K}^\varepsilon)$ because $d(x^\varepsilon(h_0), \mathcal{K}(h_0)) \le \varepsilon$. Therefore, we can repeat the above procedure and select $h_1 \in (0, \delta)$ such that

$$\overline{D}\left[x^\varepsilon(h_0) + \mathrm{cl}_{\mathbb{L}}\left(\int_{t_1}^{t_1+h_1} F(\tau, x^\varepsilon(h_0))\mathrm{d}\tau + \int_{t_1}^{t_1+h_1} G(\tau, x^\varepsilon(h_0))\mathrm{d}B_\tau\right), \mathcal{K}(t_1+h_1)\right] \le \frac{\varepsilon h_1}{2^2}.$$

Similarly as above, we can select $f^1 \in S_{\mathbb{F}}(F \circ x^\varepsilon(h_0))$ and $g^1 \in S_{\mathbb{F}}(G \circ x^\varepsilon(h_0))$, and define $x^\varepsilon(t) = x^\varepsilon(t_1) + \int_{t_1}^t f_\tau^1 \mathrm{d}\tau + \int_{t_1}^t g_\tau^1 \mathrm{d}B_\tau$ for $t_1 \le t \le t_2$, where $t_2 = t_1 + h_1$. We can also extend the function θ on $[0, t_2]$ by taking $\theta(t) = t_1$ for $t_1 \le t < t_2$ and $\theta_\varepsilon(t_2) = t_2$. We have

$$x^\varepsilon(t_2) \in x^\varepsilon(t_1) + \mathrm{cl}_{\mathbb{L}}\left(\int_{t_1}^{t_2} F(\tau, x^\varepsilon(t_1))\mathrm{d}\tau + \int_{t_1}^{t_2} G(\tau, x^\varepsilon(t_1))\mathrm{d}B_\tau\right).$$

Therefore, for every $t_1 \le t \le t_2$, one has

$$d(x^\varepsilon(t), \mathcal{K}(t)) \le \|x^\varepsilon(t) - x^\varepsilon(t_2)\| + d(x^\varepsilon(t_2), \mathcal{K}(t_2)) + H(\mathcal{K}(t_2), \mathcal{K}(t)) \le \varepsilon,$$

because similarly as above, we get $\|x^\varepsilon(t) - x^\varepsilon(t_2)\| \le \varepsilon/2$ for every $t_1 \le t \le t_2$. Similarly as above, for every $l \in C_b(\mathbb{R}^d, \mathbb{R})$, $h \in C_b^2(\mathbb{R}^d, \mathbb{R})$, and $t_1 \le s \le t \le t_2$, we also get

$$E\left[l(x^\varepsilon(s))\left(h(x^\varepsilon(t)) - h(x^\varepsilon(s)) - \int_s^t (\mathbb{L}_{f^\varepsilon g^\varepsilon}^{x^\varepsilon} h)_\tau \mathrm{d}\tau\right)\right] = 0.$$

Let $x_2 \in \mathcal{K}(t_2)$ be such that $\|x^\varepsilon(t_2) - x_2\| \le d(x^\varepsilon(t_2), \mathcal{K}(t_2)) + \varepsilon/2^2$. Hence it follows that $E[\mathrm{dist}(x^\varepsilon(t_2), K(t_2))] \le \varepsilon$. Let us observe that the above relations can be written in the form presented in (i)–(iii) above with $T = t_2$, where $f^\varepsilon = \mathbb{1}_{[0,t_1)} f^0 + \mathbb{1}_{(t_1,t_2]} f^1$, $g^\varepsilon = \mathbb{1}_{[0,t_1)} g^0 + \mathbb{1}_{(t_1,t_2]} g^1$ and $x^\varepsilon(t) = x_0 + \int_0^t f_\tau^\varepsilon \mathrm{d}\tau + \int_0^t g_\tau^\varepsilon \mathrm{d}B_\tau$ for $0 \le t \le t_2$.

Continuing the above procedure, we can extend the function θ_ε and processes f^ε, g^ε, and x^ε on the whole interval $[0, T]$ such that the above conditions (i)–(iii) are satisfied. To see this, let us denote by Λ_ε the set of all extensions of the vector function $\Phi_\varepsilon = (\theta_\varepsilon, f^\varepsilon, g^\varepsilon, x^\varepsilon)$ on $[0, \alpha] \times \Omega$ with $\alpha \in (0, T]$ and $\theta_\varepsilon|_{[0,\alpha]}$ not depending on $\omega \in \Omega$. We have $\Lambda_\varepsilon \ne \emptyset$. Let us introduce in Λ_ε the partial order relation $\preceq$ by setting $\Phi_\varepsilon^\alpha \preceq \Phi_\varepsilon^\beta$ if and only if $\alpha \le \beta$ and $\Phi_\varepsilon^\alpha = \Phi_\varepsilon^\beta|_{[0,\alpha]}$, where Φ_ε^α

and Φ_ε^β denote extensions of Φ_ε to $[0,\alpha]$ and $[0,\beta]$, respectively. Let P_ε^α be a set containing an extension Φ_ε^α and all its restrictions $\Phi_\varepsilon^\alpha|_{[0,a]}$ for every $a \in (0,\alpha]$. It is clear that each completely ordered subset of Λ_ε is of the form P_ε^α determined by some extension Φ_ε^α. It is also clear that every set P_ε^α has Φ_ε^α as its upper bound. Then by the Kuratowski–Zorn lemma, there exists a maximal element Ψ_ε of Λ_ε defined on $[0,b]\times\Omega$ with $b \le T$. It has to be $b = T$. Indeed, if it were $b < T$, then we could repeat the above procedure and extend Ψ_ε to the vector function Γ_ε defined on $[0,\gamma]\times\Omega$ with $b \le \gamma$. It would be $\Psi_\varepsilon \preceq \Gamma_\varepsilon$, in contradiction to the assumption that Ψ_ε is a maximal element of Λ_ε. Then Φ_ε can be extended on $[0,T]\times\Omega$ in such a way that conditions (i)–(iii) are satisfied. □

Remark 2.1. Theorem 2.1 is also true if instead of (2.1), we assume that

$$\liminf_{h\to 0+}\frac{1}{h}\overline{D}\left[x + \mathrm{cl}_{\mathbb{L}}\left(\int_t^{t+h} F(\tau,x)\mathrm{d}\tau\right) + \int_t^{t+h} G(\tau,x)\mathrm{d}B_\tau, \mathcal{K}(t+h)\right] = 0 \tag{2.2}$$

for every $(t,x) \in$ Graph $(\mathcal{K}^\varepsilon)$. □

Theorem 2.2. *Assume that F and G satisfy conditions $(\mathcal{H}_1)$. Suppose $\mathcal{P}_{\mathbb{F}} = (\Omega,\mathcal{F},\mathbb{F},P)$ is a complete filtered probability space with a filtration $\mathbb{F} = (\mathcal{F}_t)_{0\le t\le T}$ such that there exists an m-dimensional $\mathbb{F}$-Brownian motion $B = (B_t)_{0\le t\le T}$ defined on $\mathcal{P}_{\mathbb{F}}$. Let $K : [0,T]\times\Omega \to \mathrm{Cl}(\mathbb{R}^d)$ be $\mathbb{F}$-nonanticipative such that $\mathcal{K}(t) \ne \emptyset$ for every $0 \le t \le T$ and $(\mathcal{K}(t))_{0\le t\le T}$ is continuous. If (2.1) is satisfied for every $(t,x) \in Graph(\mathcal{K})$, then for every $\varepsilon \in (0,1)$, $a \in (0,T)$, $x_0 \in \mathcal{K}(0)$, and $\mathbb{F}$-nonanticipative processes $\phi = (\phi_t)_{0\le t\le T}$ and $\psi = (\psi_t)_{0\le t\le T}$ with $\phi_t \in \mathbb{L}^2(\Omega,\mathcal{F}_T,\mathbb{R}^d)$, $\psi_t \in \mathbb{L}^2(\Omega,\mathcal{F}_T,\mathbb{R}^{d\times m})$ for $0 \le t \le T$ and $(\phi_0,\psi_0) \in F(0,x_0)\times G(0,x_0)$ a.s., there exist a partition $0 = t_0 < t_1 < \cdots < t_p = a$ of the interval $[0,a]$, a step function $\theta_\varepsilon : [0,a] \to [0,a]$, $\mathbb{F}$-nonanticipative stochastic processes $f^\varepsilon = (f_t^\varepsilon)_{0\le t\le a}$ and $g^\varepsilon = (g_t^\varepsilon)_{0\le t\le a}$, and a step stochastic process $z^\varepsilon = (z^\varepsilon(t))_{0\le t\le a}$ such that*

(i) $t_{j+1}-t_j \le \delta$, *where* $\delta \in (0,\varepsilon)$ *is such that* $\max\left(\int_t^{t+\delta} k^2(\tau)\mathrm{d}\tau, \int_t^{t+\delta} m^2(\tau)\mathrm{d}\tau\right) \le \varepsilon^2/2^4$ *and* $D(\mathcal{K}(t+\delta),\mathcal{K}(t)) \le \varepsilon/2^2$ *for* $t \in [0,T]$;
(ii) $\theta_\varepsilon(t) = t_j$ *for* $t_j \le t < t_{j+1}$ *for* $j = 0,1,\dots p-2$ *and* $\theta_\varepsilon(t) = t_{p-1}$ *for* $t_{p-1} \le t \le a$;
(iii) $f^\varepsilon \in S_{\mathbb{F}}(F\circ(x^\varepsilon\circ\theta_\varepsilon))$, $g^\varepsilon \in S_{\mathbb{F}}(G\circ(x^\varepsilon\circ\theta_\varepsilon))$, $|\phi_t(\omega) - f_t^\varepsilon(\omega)| \le \mathrm{dist}(\phi_t, F(t,(x^\varepsilon\circ\theta_\varepsilon)(t)))$ *and* $|\psi_t(\omega) - g_t^\varepsilon(\omega)| \le \mathrm{dist}(\psi_t, G(t,(x^\varepsilon\circ\theta_\varepsilon)(t)))$ *for* $(t,\omega) \in [0,a]\times\Omega$, *where* $x^\varepsilon(t) = x_0 + \int_0^t (f_\tau^\varepsilon + z^\varepsilon(\tau))\mathrm{d}\tau + \int_0^t g_\tau^\varepsilon \mathrm{d}B_\tau$ *a.s. for* $0 \le t \le a$;
(iv) $\|z^\varepsilon(t)\| \le \varepsilon/2^2$ *for* $0 \le t \le a$, *where* $\|z^\varepsilon(t)\|^2 = E|z(t)|^2$;
(v) $d(x^\varepsilon(\theta_\varepsilon(t)),\mathcal{K}(\theta_\varepsilon(t))) = 0$ *for* $0 \le t \le a$;
(vi) $d\left(x^\varepsilon(t) - x^\varepsilon(s), \mathrm{cl}_{\mathbb{L}}\left(\int_s^t F(\tau,(x^\varepsilon\circ\theta_\varepsilon)(\tau))\mathrm{d}\tau + \int_s^t G(\tau,(x^\varepsilon\circ\theta_\varepsilon)(\tau))\mathrm{d}B_\tau\right)\right) \le \varepsilon$ *for every* $0 \le s < t \le a$.

Proof. Let $x_0 \in \mathcal{K}(0)$, $\varepsilon \in (0,1)$ and $a \in (0,T)$ be fixed. Without loss of generality, we can assume that $T = 1$. By virtue of (2.1), there exists $h_0 \in (0,\delta)$ such that

$$\overline{D}\left[x_0 + \mathrm{cl}_{\mathbb{L}}\left(\int_0^{h_0} F(\tau,x_0)\mathrm{d}\tau + \int_0^{h_0} G(\tau,x_0)\mathrm{d}B_\tau\right), \mathcal{K}(h_0)\right] \le \frac{\varepsilon h_0}{2^2},$$

where $\delta > 0$ is such that condition (i) is satisfied. By virtue of Corollary 2.3 of Chap. 2 applied to $\Sigma_{\mathbb{F}}$-measurable multifunctions $F \circ x_0$ and $G \circ x_0$, and given the above processes ϕ and ψ, there exist $f^0 \in S_{\mathbb{F}}(F \circ x_0)$ and $g^0 \in S_{\mathbb{F}}(G \circ x_0)$ such that $|\phi_t(\omega) - f_t^0(\omega)| = \mathrm{dist}(\phi_t, F(t,x_0))$ and $|\psi_t(\omega) - g_t^0(\omega)| = \mathrm{dist}(\psi_t, G(t,x_0))$ for $(t,\omega) \in [0,a] \times \Omega$. Similarly as in the proof of Theorem 2.1, we define now the function θ_ε by taking $\theta_\varepsilon(t) = 0$ for $0 \le t < t_1$ and $\theta_\varepsilon(t_1) = t_1$, where $t_1 = h_0$. Hence it follows that $f_t^0 \in F(t,\theta_\varepsilon(t))$ and $g_t^0 \in G(t,\theta_\varepsilon(t))$ a.s. for $0 \le t < t_1$. Let $y_0 = x_0 + \int_0^{t_1} f_\tau^0 \mathrm{d}\tau + \int_0^{t_1} g_\tau^0 \mathrm{d}B_\tau$ a.s. We have

$$y_0 \in x_0 + \mathrm{cl}_{\mathbb{L}}\left(\int_0^{h_0} F(\tau,x_0)\mathrm{d}\tau + \int_0^{h_0} G(\tau,x_0)\mathrm{d}B_\tau\right).$$

Then $d(y_0,\mathcal{K}(h_0)) \le \varepsilon h_0/2^2$, which by Theorem 3.1 of Chap. 2, implies that $d^2(y_0,\mathcal{K}(h_0)) = E[\mathrm{dist}(y_0, K(h_0,\cdot)]^2$. Therefore, by Corollary 2.3 of Chap. 2, there exists an $\mathcal{F}_{t_1}$-measurable random variable x_1 such that $x_1 \in K(h_0,\cdot)$ for $\omega \in \Omega$ and

$$\|y_0 - x_1\| = \left(E\left[dist^2(y_0, K(h_0,\cdot)\right]\right)^{1/2} = d(y_0,\mathcal{K}(h_0)) \le \varepsilon h_0/2^2.$$

Define $z_t^\varepsilon = (1/h_0)(x_1 - y_0)$ a.s. for $0 \le t \le t_1$. We get $||z^\varepsilon(t)|| \le (1/h_0)||x_1 - y_0|| \le (1/h_0)(\varepsilon h_0/2^2) = \varepsilon/4$ for $0 \le t \le t_1$. We define now a process x^ε on $[0,t_1)$ by setting

$$x^\varepsilon(t) = x_0 + \int_0^t (f_\tau^0 + z^\varepsilon(\tau))\mathrm{d}\tau + \int_0^t g_\tau^0 \mathrm{d}B_\tau \quad \text{a.s.} \quad \text{for} \quad 0 \le t \le t_1.$$

We have $x^\varepsilon(0) = x_0 \in \mathcal{K}(0)$ and $x^\varepsilon(t_1) = y_0 + h_0(1/h_0)(x_1 - y_0) = x_1 \in \mathcal{K}(h_0)$, which is equivalent to $d(x^\varepsilon(\theta_\varepsilon(t), \mathcal{K}(\theta_\varepsilon(t))) = 0$ for $t \in [0,t_1]$. Similarly, for $0 \le s \le t < t_1$, one obtains

$$\begin{aligned}
&d\left[x^\varepsilon(t) - x^\varepsilon(s), \mathrm{cl}_{\mathbb{L}}\left(\int_s^t F(\tau,(x^\varepsilon \circ \theta_\varepsilon)(\tau))\mathrm{d}\tau + \int_s^t G(\tau,(x^\varepsilon \circ \theta_\varepsilon)(\tau))\mathrm{d}B_\tau\right)\right] \\
&\quad \le d\left[\int_s^t f_\tau^0 \mathrm{d}\tau + \int_s^t g_\tau^0 \mathrm{d}B_\tau, \mathrm{cl}_{\mathbb{L}}\left(\int_s^t F(\tau,x^0)\mathrm{d}\tau + \int_s^t G(\tau,x^0)\mathrm{d}B_\tau\right)\right] \\
&\qquad + (t-s)\sup_{0 \le \tau \le t_1} ||z^\varepsilon(\tau)|| \le \frac{\varepsilon}{4} < \varepsilon.
\end{aligned}$$

If $h_0 < a$, we can repeat the above procedure. Applying (2.1) to $(t_1, x_1) \in Graph(\mathcal{K})$, we can select $h_1 \in (0, \delta)$ such that

$$\overline{D}\left[x_1 + \mathrm{cl}_{\mathbb{L}}\left(\int_{t_1}^{t_1+h_1} F(\tau, x_1)\mathrm{d}\tau + \int_{t_1}^{t_1+h_1} G(\tau, x_1)\mathrm{d}B_\tau\right), \mathcal{K}(t_1 + h_1)\right] \leq \frac{\varepsilon h_1}{2^2}.$$

Similarly as above, we can select $x_2 \in \mathcal{K}(t_1 + h_1)$, $f^1 \in S_{\mathbb{F}}(F \circ x_1)$, and $g^1 \in S_{\mathbb{F}}(G \circ x_1)$ such that $|\phi_t(\omega) - f_t^1(\omega)| = \mathrm{dist}(\phi_t, F(t, x^1))$ and $|\psi_t(\omega) - g_t^1(\omega)| = \mathrm{dist}(\psi_t, G(t, x^1))$ for $(t, \omega) \in [0, a] \times \Omega$ and $\|y_1 - x_2\| \leq \varepsilon h_1/2^2$, where $y_1 = x_1 + \int_{t_1}^{t_1+h_1} f_\tau^1 \mathrm{d}\tau + \int_{t_1}^{t_1+h_1} g_\tau^1 \mathrm{d}B_\tau$ a.s. We can extend the function θ_ε and the process z^ε on the interval $[0, t_2]$ by setting $\theta_\varepsilon(t) = t_1$ for $t_1 \leq t < t_2$, $\theta(t_2) = t_2$, and $z^\varepsilon(t) = (1/h_1)(x_2 - y_1)$ for $t_1 < t \leq t_2$, where $t_2 = t_1 + h_2$. Define on the interval $[0, t_2]$ the process x^ε by setting

$$x^\varepsilon(t) = x_0 + \int_0^t (f_\tau^\varepsilon + z^\varepsilon(\tau))\mathrm{d}\tau + \int_0^t g_\tau^\varepsilon \mathrm{d}B_\tau \quad \text{a.s.} \quad \text{for} \quad 0 \leq t \leq t_2,$$

where $f^\varepsilon = \mathbb{1}_{[0,t_1)} f^0 + \mathbb{1}_{[t_1,t_2)} f^1$ and $g^\varepsilon = \mathbb{1}_{[0,t_1)} g^0 + \mathbb{1}_{[t_1,t_2)} g^1$. Similarly as above, we obtain $d(x^\varepsilon(\theta_\varepsilon(t), \mathcal{K}(\theta_\varepsilon(t))) = 0$ for $0 \leq t < t_2$ and $d(x^\varepsilon(\theta_\varepsilon(t_2), \mathcal{K}(\theta_\varepsilon(t_2))) = 0$, because $x^\varepsilon(t_2) = x_2$. Then $d(x^\varepsilon(\theta_\varepsilon(t), \mathcal{K}(\theta_\varepsilon(t))) = 0$ for $0 \leq t \leq t_2$. It is clear that $\|z_t^\varepsilon\| \leq \varepsilon/4 \leq \varepsilon$ for every $0 \leq t \leq t_2$. Then for every $0 \leq s \leq t \leq t_2$, we get

$$\begin{aligned} & d\left[x^\varepsilon(t) - x^\varepsilon(s), \mathrm{cl}_{\mathbb{L}}\left(\int_s^t F(\tau, (x^\varepsilon \circ \theta_\varepsilon)(\tau))\mathrm{d}\tau + \int_s^t G(\tau, (x^\varepsilon \circ \theta_\varepsilon)(\tau))\mathrm{d}B_\tau\right)\right] \\ & \leq d\left[\int_s^t f_\tau^\varepsilon \mathrm{d}\tau + \int_s^t g_\tau^\varepsilon \mathrm{d}B_\tau, \mathrm{cl}_{\mathbb{L}}\left(\int_s^t F(\tau, (x^\varepsilon \circ \theta_\varepsilon)(\tau))\mathrm{d}\tau\right.\right. \\ & \quad \left.\left. + \int_s^t G(\tau, (x^\varepsilon \circ \theta_\varepsilon)(\tau))\mathrm{d}B_\tau\right)\right] + (t - s) \sup_{0 \leq \tau \leq t_2} ||z^\varepsilon(\tau|| \leq \frac{\varepsilon}{4} < \varepsilon. \end{aligned}$$

Suppose that for some $i \geq 1$, the inductive procedure is realized on $[0, t_i) \subset [0, a]$ and the above step function θ_ε, and stochastic processes z^ε f^ε, g^ε, and x^ε are extended to $[0, t_i]$ and $[0, t_i)$, respectively, with the above properties on this interval. Denote by S_i the set of all positive numbers h such that $h \in (0, \min(\delta, a - t_i))$ and

$$\overline{D}\left[x_i + \mathrm{cl}_{\mathbb{L}}\left(\int_{t_i}^{t_i+h} F(\tau, x_i)\mathrm{d}\tau + \int_{t_i}^{t_i+h} G(\tau, x_i)\mathrm{d}B_\tau\right), \mathcal{K}(t_i + h)\right] \leq \frac{\varepsilon h}{2^3},$$

where $x_i = x^\varepsilon(t_i)$. We have $S_i \neq \emptyset$ and $\sup S_i > 0$. Choose $h_i \in S_i$ such that $\sup S_i - (1/2) \sup S_i \leq h_i$. Put $t_{i+1} = t_i + h_i$ and let $f^i \in S_{\mathbb{F}}(F \circ x_i)$ and $g^i \in S_{\mathbb{F}}(G \circ x_i)$ be such that $|\phi_t(\omega) - f_t^i(\omega)| = \mathrm{dist}(\phi_t, F(t, x_i))$ and $|\psi_t(\omega) - g_t^i(\omega)| = \mathrm{dist}(\psi_t, G(t, x_i))$. We can now extend θ_ε, f^ε, and g^ε to the interval $[0, t_{i+1}]$ by taking $\theta_\varepsilon(t) = t_i$ for $t_i \le t < t_{i+1}$ and $\theta_\varepsilon(t_{i+1}) = t_{i+1}$, $f_t^\varepsilon = f_t^i$, $g_t^\varepsilon = g_t^i$ for $t_i \le t < t_{i+1}$. Then $|\phi_t(\omega) - f_t^\varepsilon(\omega)| \le \mathrm{dist}(\phi_t, F(t, (x^\varepsilon \circ \theta_\varepsilon)(t)))$ and $|\psi_t(\omega) - g_t^\varepsilon(\omega)| \le \mathrm{dist}(\psi_t, G(t, (x^\varepsilon \circ \theta_\varepsilon)(t)))$ for $(t, \omega) \in [0, t_{i+1}) \times \Omega$, where

$$x^\varepsilon(t) = x_0 + \int_0^t (f_\tau^\varepsilon + z^\varepsilon(\tau))\mathrm{d}\tau + \int_0^t g_\tau^\varepsilon \mathrm{d}B_\tau$$

a.s. for $0 \le t \le t_{i+1}$ with

$$z^\varepsilon(t) = (1/h_i)\left(x_{i+1} - x_i - \int_{t_i}^{t_{i+1}} f_\tau^\varepsilon \mathrm{d}\tau - \int_{t_i}^{t_{i+1}} g_\tau^\varepsilon \mathrm{d}B_\tau\right)$$

a.s for $t_i < t \le t_{i+1}$, where $x_{i+1} \in \mathcal{K}(t_{i+1})$ is such that

$$\left\| x_i + \int_{t_i}^{t_{i+1}} f_\tau^\varepsilon \mathrm{d}\tau + \int_{t_i}^{t_{i+1}} g_\tau^\varepsilon \mathrm{d}B_\tau - x_{i+1} \right\| \le \varepsilon h_i/4.$$

Similarly as above, we obtain $||z^\varepsilon(t)|| \le \varepsilon/4$ for $t_i < t \le t_{i+1}$. Hence it follows that

$$\begin{aligned} &d\left[x^\varepsilon(t) - x^\varepsilon(s), \mathrm{cl}_{\mathbb{L}}\left(\int_s^t F(\tau, (x^\varepsilon \circ \theta_\varepsilon)(\tau))\mathrm{d}\tau + \int_s^t G(\tau, (x^\varepsilon \circ \theta_\varepsilon)(\tau))\mathrm{d}B_\tau\right)\right] \\ &\le d\left[\int_s^t f_\tau^\varepsilon \mathrm{d}\tau + \int_s^t g_\tau^\varepsilon \mathrm{d}B_\tau, \mathrm{cl}_{\mathbb{L}}\left(\int_s^t F(\tau, (x^\varepsilon \circ \theta_\varepsilon)(\tau))\mathrm{d}\tau + \int_s^t G(\tau, (x^\varepsilon \circ \theta_\varepsilon)(\tau))\mathrm{d}B_\tau\right)\right] \\ &\quad +(t-s) \sup_{0 \le \tau \le t_2} ||z^\varepsilon(\tau|| \le \frac{\varepsilon}{4} < \varepsilon \end{aligned}$$

for $0 \le s < t < t_{i+1}$ and $d(x^\varepsilon(\theta_\varepsilon(t), \mathcal{K}(\theta_\varepsilon(t))) = 0$ for $0 \le t \le t_2$.

We can continue the above procedure up to $n > 1$ such that $t_n \in [a, 1]$. Suppose to the contrary that such $n > 1$ does not exist, i.e., that for every $n > 1$, one has $0 < t_n < a$. Then we obtain a sequence $(t_i)_{i=1}^\infty$ converging to $t^* \le a$ such that for every $0 \le j < k \le i + 1$ and $i \ge 0$, we have

$$\begin{aligned} ||x^\varepsilon(t_k) - x^\varepsilon(t_j)|| &\le \left\| \int_{t_j}^{t_k} f_\tau^\varepsilon \mathrm{d}\tau \right\| + \left\| \int_{t_j}^{t_k} g_\tau^\varepsilon \mathrm{d}B_\tau \right\| + \left\| \int_{t_j}^{t_k} z^\varepsilon(\tau)\mathrm{d}\tau \right\| \\ &\le 2\left(\int_{t_j}^{t_k} m^2(\tau)\mathrm{d}\tau\right)^{1/2} + \varepsilon \cdot (t_k - t_j)/4. \end{aligned}$$

Let $x_j = x^\varepsilon(t_j)$ and $x_k = x^\varepsilon(t_k)$ for $0 \le j < k < \infty$. For every $0 \le j < k < \infty$, one gets

$$\|x_k - x_j\| \le 2\left(\int_{t_j}^{t_k} m^2(\tau)\mathrm{d}\tau\right)^{1/2} + \varepsilon \cdot (t_k - t_j)/4.$$

Then $(x_i)_{i=1}^{\infty}$ is a Cauchy sequence of $\mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$. Therefore, there exists $x^* \in \mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$ such that $\|x_i - x^*\| \to 0$ as $i \to \infty$. By the continuity of the set-valued mapping $\mathcal{K}$, we get $(t^*, x^*) \in Graph(\mathcal{K})$. Then by (2.1), there exists $h^* \in (0, \min(\delta, 1-a))$ such that

$$\overline{D}\left[x^* + \mathrm{cl}_{\mathbb{L}}\left(\int_{t^*}^{t^*+h^*} F(\tau, x^*)\mathrm{d}\tau + \int_{t^*}^{t^*+h^*} G(\tau, x^*)\mathrm{d}B_\tau\right), \mathcal{K}(t^* + h^*)\right] \leq \frac{\varepsilon h^*}{2^3}.$$

Let $N > 1$ be such that for every $i \geq N$, one has $0 < t^* - t_i < \min(h^*, a, \eta_\varepsilon)$, $\|x_i - x^*\| \leq \varepsilon h^*/(2^6 A)$, and $D(\mathcal{K}(t_i), \mathcal{K}(t^*)) \leq \varepsilon h^*/2^6$, where $A = 1 + 2\left(\int_0^1 k^2(t)\mathrm{d}t\right)^{1/2}$ and $\eta_\varepsilon \in (0, 1-a)$ is such that $\left(\int_t^{t+\eta_\varepsilon} m^2(\tau)\mathrm{d}\tau\right)^{1/2} \leq \varepsilon h^*/2^7$ for every $0 \leq t \leq a$. For every $i \geq N$ and arbitrarily taken $\phi^i \in S_{\mathbb{F}}(F \circ x_i)$ and $\psi^i \in S_{\mathbb{F}}(G \circ x_i)$, we can select $f^* \in S_{\mathbb{F}}(F \circ x^*)$ and $g^* \in S_{\mathbb{F}}(G \circ x^*)$ such that $|\phi_t^i(\omega) - f_t^*(\omega)| = \mathrm{dist}(\phi^i, F(t, x^*))$ and $|\psi_t^i(\omega) - g_t^*(\omega)| = \mathrm{dist}(\psi^i, G(t, x^*))$ for $(t, \omega) \in [t_i, t^* + h^*] \times \Omega$. In particular, this implies

$$\|\phi^i - f^*\|_*^2 \leq E\int_{t_i}^{t^*+h^*} [h(F(t, x_i), F(t, x^*))]^2\mathrm{d}t \leq \int_{t_i}^{t^*+h^*} k^2(t)\|x_i - x^*\|^2\mathrm{d}t$$

and

$$\|\psi^i - f^*\|_*^2 \leq E\int_{t_i}^{t^*+h^*} [h(G(t, x_i), G(t, x^*))]^2\mathrm{d}t \leq \int_{t_i}^{t^*+h^*} k^2(t)\|x_i - x^*\|^2\mathrm{d}t$$

for $i \geq 1$. Therefore, for every $i \geq N$, we get

$$\begin{aligned}
& d\left[x_i + \int_{t_i}^{t_i+h^*} \phi_\tau^i\mathrm{d}\tau + \int_{t_i}^{t_i+h^*} \psi_\tau^i\mathrm{d}B_\tau, \mathcal{K}(t_i + h^*)\right] \\
& \leq \left\|\left[x_i + \int_{t_i}^{t_i+h^*} \phi_\tau^i\mathrm{d}\tau + \int_{t_i}^{t_i+h^*} \psi_\tau^i\mathrm{d}B_\tau\right] - \left[x^* + \int_{t^*}^{t^*+h^*} f_\tau^*\mathrm{d}\tau + \int_{t^*}^{t^*+h^*} g_\tau^*\mathrm{d}B_\tau\right]\right\| \\
& \quad + d\left[x^* + \int_{t^*}^{t^*+h^*} f_\tau^*\mathrm{d}\tau + \int_{t^*}^{t^*+h^*} g_\tau^*\mathrm{d}B_\tau, \mathcal{K}(t^* + h^*)\right] + D(\mathcal{K}(t^* + h^*), \mathcal{K}(t_i + h^*)) \\
& \leq ||x_i - x^*|| + \left\|\int_{t_i}^{t^*+h^*} (\phi_\tau^i - f_\tau^*)\mathrm{d}\tau\right\| + \left\|\int_{t_i}^{t^*+h^*} (\psi_\tau^i - g_\tau^*)\mathrm{d}B_\tau\right\| \\
& \quad + \left\|\int_{t_i+h^*}^{t^*+h^*} \phi_\tau^i\mathrm{d}\tau\right\| + \left\|\int_{t_i+h^*}^{t^*+h^*} \psi_\tau^i\mathrm{d}B_\tau\right\| + \left\|\int_{t_i}^{t^*} f_\tau^*\mathrm{d}\tau\right\| + \left\|\int_{t_i}^{t^*} g_\tau^*\mathrm{d}B_\tau\right\| \\
& \quad + d\left[x^* + \int_{t^*}^{t^*+h^*} f_\tau^*\mathrm{d}\tau + \int_{t^*}^{t^*+h^*} g_\tau^*\mathrm{d}B_\tau, \mathcal{K}(t^* + h^*)\right] + D(\mathcal{K}(t^* + h^*), \mathcal{K}(t_i + h^*)) \\
& \leq \|x_i - x^*\| + 2\sqrt{(t^* - t_i) + h^*}\,\|x_i - x^*\|\left(\int_{t_n}^{t^*+h^*} k^2(\tau)\mathrm{d}\tau\right)^{1/2}
\end{aligned}$$

$$+\left(1+\sqrt{(t^*-t_i)}\right)\left(\int_{t_i+h^*}^{t^*+h^*} m^2(\tau)\mathrm{d}\tau\right)^{1/2}+\left(1+\sqrt{(t^*-t_i)}\right)\left(\int_{t_i}^{t^*} m^2(\tau)\mathrm{d}\tau\right)^{1/2}$$

$$+2\frac{\varepsilon h^*}{2^6}+\frac{\varepsilon h^*}{2^6}\le\left[1+2\sqrt{(t^*-t_i)+h^*}\left(\int_{t_i}^{t^*+h^*} k^2(\tau)\mathrm{d}\tau\right)^{1/2}\right]\|x_i-x^*\|$$

$$+\left(1+\sqrt{(t^*-t_i)}\right)\max\left\{\left[\int_{t_i+h^*}^{t^*+h^*} m^2(\tau)\mathrm{d}\tau\right]^{1/2},\left[\int_{t_i}^{t^*} m^2(\tau)\mathrm{d}\tau\right]^{1/2}\right\}+2\frac{\varepsilon h^*}{2^6}$$

$$+\frac{\varepsilon h^*}{2^6}\le\left[1+2\left(\int_0^1 k^2(\tau)\mathrm{d}\tau\right)^{1/2}\right]\|x_i-x^*\|$$

$$+\left(1+\sqrt{(t^*-t_i)}\right)\max\left\{\left[\int_{t_i+h^*}^{t^*+h^*} m^2(\tau)\mathrm{d}\tau\right]^{1/2},\left[\int_{t_i}^{t^*} m^2(\tau)\mathrm{d}\tau\right]^{1/2}\right\}$$

$$+2\frac{\varepsilon h^*}{2^6}+\frac{\varepsilon h^*}{2^6}\le\frac{\varepsilon h^*}{2^6\cdot A}\cdot A+2\cdot\frac{\varepsilon h^*}{2\cdot 2^6}+2\cdot\frac{\varepsilon h^*}{2^6}+\frac{\varepsilon h^*}{2^6}$$

$$=5\cdot\frac{\varepsilon h^*}{2^6}=\frac{5}{8}\cdot\frac{\varepsilon h^*}{2^3}<\frac{\varepsilon h^*}{2^3}.$$

Then for every $i\ge N$, we have

$$D\left[x_i+\mathrm{cl}_{\mathbb{L}}\left(\int_{t_i}^{t_i+h^*} F(\tau,x_i)\mathrm{d}\tau+\int_{t_i}^{t_i+h^*} G(\tau,x_i)\mathrm{d}B_\tau\right),\mathcal{K}(t_i+h^*)\right]\le\frac{\varepsilon h^*}{2^3}$$

and $h^*\in(0,\min(\delta,1-a))$. But $t_i<a$ for every $i\ge 1$. Then $1-a<1-t_i$ for every $i\ge 1$. Therefore, for every $i\ge N$, we have $h^*\in(0,\min(\delta,1-t_i))$. Hence it follows that $h^*\in S_i$ for every $i\ge N$. Then for every $i\ge N$, one has $(1/2)h^*\le(1/2)\sup S_i\le h_i=t_{i+1}-t_i$, which contradicts the convergence of the sequence $(t_i)_{i=1}^{\infty}$. Therefore, there exists $p\ge 1$ such that $0=t_0<t_1<\cdots<t_p=a$. □

Remark 2.2. Theorem 2.2 is also true if instead of (2.1), we assume that (2.2) is satisfied for every $(t,x)\in Graph(\mathcal{K})$. □

Theorem 2.3. *Assume that F satisfies conditions $(\mathcal{H}_1)$, and let $\mathcal{P}_{\mathbb{F}}=(\Omega,\mathcal{F},\mathbb{F},P)$ be a complete filtered probability space with a continuous filtration $\mathbb{F}=(\mathcal{F}_t)_{0\le t\le T}$ such that $\mathcal{F}_T=\mathcal{F}$. Suppose $K:[0,T]\times\Omega\to\mathrm{Cl}(\mathbb{R}^d)$ is an $\mathbb{F}$-adapted set-valued process such that $\mathcal{K}(t)\ne\emptyset$ for every $0\le t\le T$ and such that the set-valued mapping $\mathcal{K}:[0,T]\to\mathrm{Cl}(\mathbb{L}(\Omega,\mathcal{F}_T,\mathbb{R}^d)$ is continuous. If*

$$\liminf_{h\to 0+}\frac{1}{h}\overline{D}\left[S\left(E\left[x+\int_{t-h}^{t} F(\tau,x)\mathrm{d}\tau|\mathcal{F}_{t-h}\right]\right),\mathcal{K}(t-h)\right]=0 \tag{2.3}$$

is satisfied for every $(t,x)\in Graph(\mathcal{K})$, where $S(E[x+\int_{t-h}^t F(\tau,x)\mathrm{d}\tau|\mathcal{F}_{t-h}])=\{E[x+\int_{t-h}^t f_\tau\mathrm{d}\tau|\mathcal{F}_{t-h}]: f\in S(coF\circ x)\}$, then for every $\varepsilon\in(0,1)$, $x_T\in\mathcal{K}(x_T)$, $a\in(0,T)$ and measurable process $\phi=(\phi)_{0\le t\le T}$ such that $\phi_t\in\mathbb{L}(\Omega,\mathcal{F}_T,\mathbb{R}^d)$ for $0\le t\le T$ and $\phi_T\in F(T,x_T)$ a.s., there exist a partition $a=t_p<t_{p-1}<$

$\cdots < t_1 < t_0 = T$ *of the interval* $[a, T]$*, a step function* $\theta_\varepsilon : [a, T] \to [a, T]$*, a step stochastic process* $z^\varepsilon = (z_t^\varepsilon)_{a \le t \le T}$*, and a measurable process* $f^\varepsilon = (f_t^\varepsilon)_{a \le t \le T}$ *on* $\mathcal{P}_{\mathbb{F}}$ *such that*

(i) $t_j - t_{j+1} \le \delta$*, where* $\delta \in (0, \varepsilon)$ *is such that* $\max\{\int_t^{t+\delta} k(\tau)\mathrm{d}\tau, \int_t^{t+\delta} m(\tau)\mathrm{d}\tau\} \le \varepsilon^2/2^4$ *and* $D(\mathcal{K}(t+\delta), \mathcal{K}(t)) \le \varepsilon/2$ *for* $t \in [0, T]$*;*
(ii) $\|z_t^\varepsilon\| \le \varepsilon/2$ *for every* $a \le t \le T$*, where* $\|z_t^\varepsilon\| = E|z_t^\varepsilon|$*;*
(iii) $\theta_\varepsilon(t) = t_{j-1}$ *for* $t_j < t \le t_{j-1}$ *and* $\theta_\varepsilon(t_j) = t_j$ *with* $j = 1, \ldots, p-1$ *and* $\theta_\varepsilon(t) = t_{p-1}$ *for* $a \le t \le t_{p-1}$*;*
(iv) $f^\varepsilon \in S(\mathrm{co}\, F \circ (x^\varepsilon \circ \theta_\varepsilon))$*,* $|\phi_t(\omega) - f_t^\varepsilon(\omega)| = \mathrm{dist}(\phi_t, \mathrm{co}\, F(t, (x^\varepsilon \circ \theta_\varepsilon)(t)))$ *for* $(t, \omega) \in [a, T] \times \Omega$*, where* $x^\varepsilon(t) = E[x_T + \int_t^T f_\tau^\varepsilon \mathrm{d}\tau | \mathcal{F}_t] + \int_t^T z_\tau^\varepsilon \mathrm{d}\tau$ *a.s. for* $a \le t \le T$ *and* $S(\mathrm{co}\, F \circ (x^\varepsilon \circ \theta_\varepsilon)) = \{f \in \mathbb{L}([a, T] \times \Omega, \beta_T \otimes \mathcal{F}_T, \mathbb{R}^d) : f_t \in \mathrm{co}\, F(t, x^\varepsilon(\theta_\varepsilon(t)))$ a.s. for $a.e.\ a \le t \le T\}$*;*
(v) $E[\mathrm{dist}(x^\varepsilon(s), E[x^\varepsilon(t) + \int_s^t F(\tau, (x^\varepsilon \circ \theta_\varepsilon)(\tau)\mathrm{d}\tau | \mathcal{F}_s])] \le \varepsilon$ *for* $a \le s \le t \le T$*,*
(vi) $d(x^\varepsilon(\theta_\varepsilon(t)), \mathcal{K}(\theta_\varepsilon(t))) = 0$ *for* $a \le t \le T$*.*

Proof. Let $\varepsilon \in (0, 1)$, $a \in (0, T)$, $x_T \in \mathcal{K}(T)$, and a measurable process $\phi = (\phi)_{0 \le t \le T}$ be given. By virtue of (2.3), there exists $h_0 \in (0, \min(\delta, T))$ such that

$$\overline{D}\left[S\left(E\left[x_T + \int_{T-h_0}^T F(\tau, x_T)\mathrm{d}\tau | \mathcal{F}_{T-h_0}\right]\right), \mathcal{K}(T-h_0)\right] \le \varepsilon h_0/2.$$

Let $t_1 = T - h_0$. By virtue of Corollary 2.3 of Chap. 2, there exists $f^0 \in S(\mathrm{co}\, F \circ x_T)$ such that $|\phi_t(\omega) - f_t^0(\omega)| = \mathrm{dist}(\phi_t(\omega), \mathrm{co}\, F(t, x_T(\omega))$ for $(t, \omega) \in [t_1, T] \times \Omega$. Let $y_0 = E[x_T + \int_{t_1}^T f_\tau^0 \mathrm{d}\tau | \mathcal{F}_{t_1}]$ a.s. We have $y_0 \in E[x_T + \int_{t_1}^T F(\tau, x_T)\mathrm{d}\tau | \mathcal{F}_{t_1}]$ a.s., i.e., $y_0 \in S(E[x_T + \int_{t_1}^T F(\tau, x_T)\mathrm{d}\tau | \mathcal{F}_{t_1}])$. Therefore, $d(y_0, \mathcal{K}(t_1)) \le \varepsilon h_0/2$. Similarly as above, we can see that there exists $x_1 \in \mathcal{K}(t_1)$ such that $E|y_0 - x_1| = E[\mathrm{dist}(y_0, K(t_1 \cdot))] = d(y_0, \mathcal{K}(t_1)) \le \varepsilon h_0/2$. Then $\|y_0 - x_1\| \le \varepsilon h_0/2$. Let $z_t^\varepsilon = 1/h_0(x_1 - y_0)$ a.s. for $t_1 \le t \le T$. We have $\|z_t^\varepsilon\| \le (1/h_0)\|y_0 - x_1\| \le \varepsilon/2$. Furthermore, by the definition of z_t^ε, it follows that $\int_t^T z_\tau^\varepsilon \mathrm{d}\tau$ is $\mathcal{F}_{t_1}$-measurable. Define $\theta_\varepsilon(t) = T$ for $t_1 < t \le T$ and $\theta(t_1) = t_1$. One has $f_t^0 \in \mathrm{co} F(t, x_T)$ a.s. for $t_1 \le t \le T$. Let

$$x^\varepsilon(t) = E\left[x_T + \int_t^T f_\tau^0 \mathrm{d}\tau | \mathcal{F}_t\right] + \int_t^T z_\tau^\varepsilon \mathrm{d}\tau$$

for $t_1 \le t \le T$. We have $x^\varepsilon(T) = x_T$ and $x^\varepsilon(t_1) = y_0 + h_0(1/h_0)(x_1 - y_0) = x_1$. Therefore, $d(x^\varepsilon(\theta(t)), \mathcal{K}(\theta(t))) = 0$ for $t_1 \le t \le T$ and $|\phi_t(\omega) - f_t^0(\omega)| = \mathrm{dist}(\phi_t(\omega), \mathrm{co}\, F(t, x^\varepsilon(\theta_\varepsilon(t)(\omega)))$ for $(t, \omega) \in [t_1, T] \times \Omega$. By the definition of x^ε, it follows that it is $\mathbb{F}$-adapted. By properties of f^0 and x^ε, it follows that

$$E\left[\mathrm{dist}\left(x^\varepsilon(s), E[x_T + \int_s^t F(\tau, x^\varepsilon(\theta(\tau)))\mathrm{d}\tau | \mathcal{F}_s]\right)\right] \le \varepsilon/2 \quad \text{for} \quad t_1 \le s \le t \le T.$$

If $t_1 > a$, we can repeat the above procedure starting with $(t_1, x_1) \in Graph(\mathcal{K})$. Immediately from (2.3), it follows that there exists an $h_1 \in (0, \delta)$ such that

$$\overline{D}\left[S(E[x_1 + \int_{t_1-h_1}^{t_1} F(\tau, x_1)\mathrm{d}\tau|\mathcal{F}_{t_1-h_1}]), \mathcal{K}(t_1 - h_1)\right] \leq \varepsilon h_1/2.$$

Similarly as above, we can select $f^1 \in S(\mathrm{co}F \circ x_1)$ and $x_2 \in \mathcal{K}(t_1 - h_1)$ such that $|\phi_t(\omega) - f_t^1(\omega))| = \mathrm{dist}(\phi_t(\omega), \mathrm{co}(F \circ x_1)(t, \omega)$ for $(t, \omega) \in [t_1 - h_1, t_1] \times \Omega$ and $\|y_1 - x_2\| \leq \varepsilon h_1/2^2$, where $y_1 = E[x_1 + \int_{t_1-h_1}^{t_1} f_\tau^1 \mathrm{d}\tau|\mathcal{F}_{t_1-h_1}]$ and $t_2 = t_1 - h_1$. We can now extend the step function θ_ε and step process z^ε on the interval $[t_2, T]$ by taking $\theta_\varepsilon(t_2) = t_2$, $\theta_\varepsilon(t) = t_1$ for $t_2 < t \leq t_1$ and $z_t^\varepsilon = (1/h_1)(x_2 - y_1)$ for $t_2 \leq t \leq t_1$. We have $f_t^1 \in \mathrm{co}F(t, \theta_\varepsilon(t))$ a.s. for $t_2 \leq t \leq t_1$. We can also extend the process x^ε to the interval $[t_2, T]$ by taking

$$x^\varepsilon(t) = E\left[x_1 + \int_t^{t_1} f_\tau^1 \mathrm{d}\tau|\mathcal{F}_t\right] + \int_t^{t_1} z_\tau^\varepsilon \mathrm{d}\tau$$

a.s. for $t_2 \leq t \leq t_1$. We have $d(x^\varepsilon(\theta_\varepsilon(t)), \mathcal{K}(\theta(t))) = 0$ for $t_2 \leq t \leq T$, because $x^\varepsilon(t_2) = x_2$. Let $f^\varepsilon = \mathbb{1}_{(t_2,t_1]}f^1 + \mathbb{1}_{(t_1,T]}f^0$. We have $x^\varepsilon(t) = E[x_T + \int_t^T f_\tau^\varepsilon \mathrm{d}\tau|\mathcal{F}_t] + \int_t^T z_\tau^\varepsilon \mathrm{d}\tau$ a.s. for $t_2 < t \leq T$. Similarly as above, we can verify that $f_t^\varepsilon \in \mathrm{co}F(t, x^\varepsilon(\theta_\varepsilon(t)))$ a.s. for $t_2 < t \leq T$ and $|\phi_t - f_t^\varepsilon)| \leq \mathrm{dist}(\phi_t, \mathrm{co}\, F(t, x^\varepsilon(\theta(t)))$ a.s. for $t_2 < t \leq T$. Furthermore, $d(x^\varepsilon(\theta_\varepsilon(t)), \mathcal{K}(\theta_\varepsilon(t))) = 0$ and $E[\mathrm{dist}(x^\varepsilon(s), E[\int_s^t F(\tau, x^\varepsilon(\theta(\tau)))\mathrm{d}\tau|\mathcal{F}_s] \leq \varepsilon/2$ for $t_2 \leq t \leq T$ and $t_2 \leq s \leq t \leq T$, respectively.

Suppose that for some $i \geq 1$, the inductive procedure is realized. Then there exist $t_{i-1} \in [a, T)$ and $x_{i-1} \in \mathcal{K}(t_{i-1})$ such that we can extend the step function θ_ε, step process z^ε, and process f^ε to the whole interval $[t_{i-1}, T]$ such that $f_t^\varepsilon \in \mathrm{co}\, F(t, x^\varepsilon(\theta_\varepsilon(t))$ and $|\phi_t - f_t^\varepsilon| = \mathrm{dist}(\phi_t, \mathrm{co}\, F(t, x^\varepsilon(\theta_\varepsilon(t)))$ for $t_{i-1} \leq t \leq T$. Define

$$x^\varepsilon(t) = E\left[x_T + \int_t^T f_\tau^\varepsilon \mathrm{d}\tau|\mathcal{F}_t\right] + \int_t^T z_\tau^\varepsilon \mathrm{d}\tau$$

a.s. for $t_{i-1} \leq t \leq T$. We have $x^\varepsilon(t_{i-1}) = x_{i-1}$, $d(x^\varepsilon(\theta_\varepsilon(t)), \mathcal{K}(\theta_\varepsilon(t))) = 0$, and

$$E\left[\mathrm{dist}\left(x^\varepsilon(s), E[x^\varepsilon(t) + \int_s^t F(\tau, (x_{i-1}^\varepsilon \circ \theta_\varepsilon)(\tau))\mathrm{d}\tau|\mathcal{F}_s]\right)\right] \leq \varepsilon/2$$

for $t_{i-1} \leq s \leq t \leq T$.

Denote by S_i the set of all positive numbers $h \in (0, \min(\delta, t_{i-1}))$ such that

$$\overline{D}\left[S(E[x^\varepsilon(t_{i-1}) + \int_{t_{i-1}-h}^{t_{i-1}} F(\tau, x_{i-1})\mathrm{d}\tau|\mathcal{F}_{t_{i-1}-h}]), \mathcal{K}(t_{i-1})\right] \leq \varepsilon h/2.$$

By the properties of x^ε, we have $x^\varepsilon(t_{i-1}) = x_{i-1}$ and $(t_{i-1}, x^\varepsilon(t_{i-1})) \in Graph(\mathcal{K})$. Therefore, by virtue of (2.3), we have $S_i \neq \emptyset$ and $\sup S_i > 0$. Choose $h_{i-1} \in S_i$ such that $(1/2)\sup S_i \leq h_{i-1}$. Put $t_i = t_{i-1} - h_{i-1}$. We can extend again the step function θ_ε, step process z^ε, and processes f^ε and x^ε to the interval $[t_i, T]$ such that $d(x^\varepsilon(\theta_\varepsilon(t)), \mathcal{K}(\theta(t)))) = 0$ for $t_i \le t \le T$, and $f_t^\varepsilon \in \mathrm{co}F(t, x^\varepsilon(\theta_\varepsilon(t))$ and $|\phi_t - f_t^\varepsilon| = \mathrm{dist}(\phi_t, \mathrm{co}F(t, x^\varepsilon(\theta_\varepsilon(t))$ a.s. for $t_i \le t \le T$. Furthermore,

$$E[\mathrm{dist}(x^\varepsilon(s), E[x^\varepsilon(t) + \int_s^t F(\tau, (x_{i-1}^\varepsilon \circ \theta_\varepsilon)(\tau))\mathrm{d}\tau|\mathcal{F}_s])] \le \varepsilon/2$$

for $t_i \le s \le t \le T$. We can continue the above procedure up to $n \ge 1$ such that $0 < t_n \le a < t_{n-1}$. Suppose to the contrary that there does not exist such $n \ge 1$, i.e., that for every $n \ge 1$, one has $a < t_n < T$. Then we can extend the step function θ_ε, the step process z^ε, and the stochastic processes f^ε and x^ε to the interval $[t_n, T]$ for every $n \ge 1$ such that $x^\varepsilon(t_n) \in \mathcal{K}(t_n)$ a.s. for every $n \ge 1$ and so that the above properties are satisfied on $[t_n, T]$ for every $n \ge 1$. By the boundedness of the sequence $(t_n)_{n=1}^\infty$, we can select a decreasing subsequence $(t_i)_{i=1}^\infty$ converging to $t^* \in [a, T]$. Let $(x_i)_{i=1}^\infty$ be a sequence defined by $x_i = x^\varepsilon(t_i)$ a.s. for every $i \ge 0$. In particular, we have $x_i \in \mathcal{K}(t_i)$ a.s. for every $i \ge 1$. For every $j > k \ge 0$, we obtain

$$\begin{aligned} E|x_k - x_j| \le{}& E|E[x_T|\mathcal{F}_{t_k}] - E[x_T|\mathcal{F}_{t_j}]| + \int_{t^*}^{t_k} m(t)\mathrm{d}t + \int_{t^*}^{t_j} m(t)\mathrm{d}t \\ &+ (t_k - t_j)E|z_t^\varepsilon| + E\left|E\left[\int_{t^*}^T f_t^\varepsilon \mathrm{d}t|\mathcal{F}_{t_k}\right] - E\left[\int_{t^*}^T f_t^\varepsilon \mathrm{d}t|\mathcal{F}_{t^*}\right]\right| \\ &+ E\left|E\left[\int_{t^*}^T f_t^\varepsilon \mathrm{d}t|\mathcal{F}_{t_j}\right] - E\left[\int_{t^*}^T f_t^\varepsilon \mathrm{d}t|\mathcal{F}_{t^*}\right]\right|. \end{aligned}$$

By the continuity of the filtration $\mathbb{F}$, it follows that $\lim_{j,k\to\infty} E|x_k - x_j| = 0$. Then $(x_i)_{i=1}^\infty$ is a Cauchy sequence of $\mathbb{L}(\Omega, \mathcal{F}_T, \mathbb{R}^d)$. Therefore, there is $x^* \in \mathbb{L}(\Omega, \mathcal{F}_T, \mathbb{R}^d$ such that $\|x_i - x^*\| \to 0$ as $i \to \infty$. But $x_i \in \mathcal{K}(t_i))$ for every $i \ge 1$ and $\mathcal{K}$ is continuous. Then $(t^*, x^*) \in Graph(\mathcal{K})$, which by virtue of (3), implies that we can select $h^* \in (0, \min(\delta, t^*))$ such that

$$\overline{D}\left[S(E[x^* + \int_{t^*-h^*}^{t^*} F(\tau, x^*)\mathrm{d}\tau|\mathcal{F}_{t^*-h^*}]), \mathcal{K}(t^* - h^*)\right] \le \varepsilon h^*/2^5.$$

Similarly as above, for every $i \ge 1$ and $\phi_i \in S(\mathrm{co}F \circ x_i)$, we can select $f^* \in S(\mathrm{co}F \circ x^*)$ such that $|\phi_t^i - f_t^*)| = \mathrm{dist}(\phi_t^i, F(t, x^*))$ a.s. for every $t^* - h^* \le t \le t^*$. By the continuity of the filtration $\mathbb{F}$, we obtain $\|E[x^*|\mathcal{F}_{t_i-h^*}] - E[x^*|\mathcal{F}_{t^*-h^*}]\| \to 0$ and

$$E\left|E\left[\int_{t^*-h^*}^{t^*} f_\tau^* \mathrm{d}\tau|\mathcal{F}_{t_i-h^*}\right] - E\left[\int_{t^*-h^*}^{t^*} f_\tau^* \mathrm{d}\tau|\mathcal{F}_{t^*-h*}\right]\right| \to 0$$

as $i \to \infty$. Let $N \ge 1$ be such that for every $i \ge N$, we have $0 < t_i - t^* < \min(h^*, \delta)$, $\|x_i - x^*\| < \varepsilon h^*/(2^5 \cdot A)$, $D(\mathcal{K}(t_i - h^*), \mathcal{K}(t^* - h^*)) \le$

$\varepsilon h^*/2^5$, $\|E[x^*|\mathcal{F}_{t_i-h^*}] - E[x^*|\mathcal{F}_{t^*-h^*}]\| \le \varepsilon h^*/2^5$, $E\int_{t_i-h^*}^{t^*-h^*} |\phi_\tau^i|\mathrm{d}\tau \le \varepsilon h^*/2^5$, $E\int_{t^*}^{t_i} |\phi_\tau^i|\mathrm{d}t \le \varepsilon h^*/2^5$, and $E|E[\int_{t^*-h^*}^{t^*} f_\tau^*\mathrm{d}\tau|\mathcal{F}_{t_i-h^*}] - E[\int_{t^*-h^*}^{t^*} f_\tau^*\mathrm{d}\tau|\mathcal{F}_{t^*-h*}]| \le \varepsilon h^*/2^5$, where $A = 1 + \int_0^T k(t)\mathrm{d}t$. By the properties of the multifunction $F(t,\cdot)$ and selector f^* of $F \circ x^*$, it follows that

$$\begin{aligned}\|\mathbb{1}_{[t^*-h^*,t^*]}(\phi^i - f^*)\| &= E\int_{t^*-h^*}^{t^*} |\phi_\tau^i - f_\tau^*|\mathrm{d}\tau \\ &\le E\int_{t^*-h^*}^{t^*} h((F(t,x_i), F(t,x^*))]\mathrm{d}t \\ &\le \|x_i - x^*\| \int_{t^*-h^*}^{t^*} k(t)\mathrm{d}t.\end{aligned}$$

For every $i \ge N$, one gets

$$\begin{aligned}&d\left(E[x_i + \int_{t_i-h^*}^{t_i} \phi_\tau^i\mathrm{d}\tau|\mathcal{F}_{t_i-h^*}], \mathcal{K}(t_i - h^*)\right) \\ &\quad\le E\left|E[x_i + \int_{t_i-h^*}^{t_i} \phi_\tau^i\mathrm{d}\tau|\mathcal{F}_{t_i-h^*}] - E[x^* + \int_{t^*-h^*}^{t^*} f_\tau^*\mathrm{d}\tau|\mathcal{F}_{t^*-h*}]\right| \\ &\qquad + d\left(E[x^* + \int_{t^*-h^*}^{t^*} f_\tau^*\mathrm{d}\tau|\mathcal{F}_{t^*-h*}], \mathcal{K}(t^*-h^*)\right) + D(\mathcal{K}(t^* - h^*), \mathcal{K}(t_i - h^*)).\end{aligned}$$

But for every $i \ge N$, we have

$$\begin{aligned}&E\left|E\left[x_i + \int_{t_i-h^*}^{t_i} \phi_\tau^i\mathrm{d}\tau|\mathcal{F}_{t_i-h^*}\right] - E\left[x^* + \int_{t^*-h^*}^{t^*} f_\tau^*\mathrm{d}\tau|\mathcal{F}_{t^*-h*}\right]\right| \\ &\quad\le E|E[(x_i - x^*)|\mathcal{F}_{t_i-h*}]| + E|E[x^*|\mathcal{F}_{t_i-h^*}] - E[x^*|\mathcal{F}_{t^*-h^*}]| \\ &\qquad + E\left|E\left[\int_{t^*}^{t^*-h^*} (\phi_\tau^i - f_\tau^*)\mathrm{d}\tau|\mathcal{F}_{t_i-h^*}\right]\right| + E\int_{t_i-h^*}^{t^*-h^*} |\phi_\tau^i|\mathrm{d}\tau + E\int_{t^*}^{t_i} |\phi_\tau^i|\mathrm{d}t \\ &\qquad + E\left|E\left[\int_{t^*}^{t^*-h^*} f_\tau^*\mathrm{d}\tau|\mathcal{F}_{t_i-h^*}\right] - E\left[\int_{t^*-h^*}^{t^*} f_\tau^*\mathrm{d}\tau|\mathcal{F}_{t^*-h*}\right]\right| \le 6\varepsilon h^*/2^5.\end{aligned}$$

Therefore, for every $i \ge N$, one gets

$$d\left[E\left[x_i + \int_{t_i-h^*}^{t_i} \phi_\tau^i\mathrm{d}\tau|\mathcal{F}_{t_i-h^*}\right], \mathcal{K}(t_i)\right] \le 8\varepsilon h^*/2^5 = \varepsilon h^*/2^2,$$

which implies that

$$\overline{D}(S(E[x_i + \int_{t_i - h^*}^{t_i} F(\tau, x_i)\mathrm{d}\tau|\mathcal{F}_{t_i - h^*}], \mathcal{K}(t_i)) \leq \varepsilon h^*/2^2.$$

But $t^* \leq t_i$ for $i \geq 1$. Therefore, for every $i \geq N$, one has $h^* \in S_{i+1}$ and $(1/2)h^* \leq \sup S_{i+1} \leq h_i = t_i - t_{i+1}$, which contradicts the convergence of the sequence $(t_i)_{i=1}^{\infty}$. Then there is a $p > 1$ such that $a = t_p < t_{p-1}, \ldots, t_1 < t_0 = T$. Taking $f^\varepsilon = \mathbb{1}_{[a,t_{p-1}]} f^p + \sum_{i=p-2}^{0} \mathbb{1}_{(t_{i+1},t_i]} f^i$, we obtain the desired selector of $\mathrm{co}F \circ (x^\varepsilon \circ \theta_\varepsilon)$. □

Remark 2.3. The above results are also true if instead of continuity of the set-valued mapping $\mathcal{K}$, we assume that it is uniformly upper semicontinuous on $[0, T]$, i.e., that $\lim_{\delta \to 0} \sup_{0 \leq t \leq T} \overline{D}(\mathcal{K}(t + \delta), \mathcal{K}(t)) = 0$. □

Conditions (2.1) and (2.3) can be expressed by certain types of stochastic tangent sets. To see this, let $(t, x) \in Graph(\mathcal{K})$ and denote by $\mathcal{T}_K(t, x)$ the set of all pairs $(f, g) \in \mathbb{L}^2([t, T] \times \Omega, \Sigma_{\mathbb{F}}^t, \mathbb{R}^d) \times \mathbb{L}^2([t, T] \times \Omega, \Sigma_{\mathbb{F}}^t, \mathbb{R}^{d \times m})$ such that

$$\liminf_{h \to 0+} (1/h) d\left[x + \int_t^{t+h} f_\tau \mathrm{d}\tau + \int_t^{t+h} g_\tau \mathrm{d}B_\tau, \mathcal{K}(t + h)\right] = 0,$$

where $\Sigma_{\mathbb{F}}^t$ denotes the σ-algebra of all $\mathbb{F}$-nonanticipative subsets of $[t, T] \times \Omega$. In a similar way, for $(t, x) \in Graph(\mathcal{K}^\varepsilon)$ and $\varepsilon \in (0, 1)$, we can define a backward stochastic tangent set $\mathcal{T}_K^b(t, x)$ with respect to a filtration $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$ as the set of all measurable processes $f \in \mathbb{L}([0, T] \times \Omega, \mathcal{F}_T, \mathbb{R}^d)$ such that

$$\liminf_{h \to 0+} (1/h) d\left(E\left[x + \int_{t-h}^{t} f_\tau \mathrm{d}\tau|\mathcal{F}_{t-h}\right], \mathcal{K}(t - h)\right) = 0.$$

Lemma 2.1. *Let $\mathcal{P}_{\mathbb{F}}$ be a complete filtered probability space. Assume that F and G satisfy condition (i) of $(\mathcal{H}_1)$ and let $K : [0, T] \times \Omega \to \mathrm{Cl}(\mathbb{R}^d)$ be $\mathbb{F}$-adapted and such that $\mathcal{K}(t) \neq \emptyset$ for every $0 \leq t \leq T$. The condition (2.1) is satisfied for every $(t, x) \in Graph(\mathcal{K})$ if and only if $S_{\mathbb{F}}^t(F \circ x) \times S_{\mathbb{F}}^t(G \circ x) \subset \mathcal{T}_K(t, x)$ for every $(t, x) \in Graph(\mathcal{K})$, where $S_{\mathbb{F}}^t(F \circ x)$ and $S_{\mathbb{F}}^t(G \circ x)$ denote the sets of all restrictions of all elements of $S_{\mathbb{F}}(F \circ x)$ and $S_{\mathbb{F}}(G \circ x)$, respectively, to the set $[t, T] \times \Omega$.*

Proof. It is clear that if (2.1) is satisfied for every $(t, x) \in Graph(\mathcal{K})$, then $S_{\mathbb{F}}^t(F \circ x) \times S_{\mathbb{F}}^t(G \circ x) \subset \mathcal{T}_K(t, x)$ for every $(t, x) \in Graph(\mathcal{K})$. Let $S_{\mathbb{F}}^t(F \circ x) \times S_{\mathbb{F}}^t(G \circ x) \subset \mathcal{T}_K(t, x)$ for fixed $(t, x) \in Graph(\mathcal{K})$. Then for every $(f, g) \in S_{\mathbb{F}}^t(F \circ x) \times S_{\mathbb{F}}^t(G \circ x)$, one has

$$\liminf_{h \to 0+} (1/h) d\left[x + \int_t^{t+h} f_\tau \mathrm{d}\tau + \int_t^{t+h} g_\tau \mathrm{d}B_\tau, \mathcal{K}(t + h)\right] = 0.$$

Thus for every $(t, x) \in Graph(\mathcal{K})$ and $(f, g) \in S^t_{\mathbb{F}}(F \circ x) \times S^t_{\mathbb{F}}(G \circ x)$ and every $\varepsilon \in (0, 1)$, there exists $h^{f,g}_\varepsilon(t) \in (0, \varepsilon)$ such that

$$d\left[x + \int_t^{t+h} f_\tau \mathrm{d}\tau + \int_t^{t+h} g_\tau \mathrm{d}B_\tau, \mathcal{K}(t+h)\right] \le h^{f,g}_\varepsilon(t) \cdot \varepsilon.$$

Let $h_\varepsilon = \sup\{h^{f,g}_\varepsilon(t) : (f, g) \in S^t_{\mathbb{F}}(F \circ x) \times S^t_{\mathbb{F}}(G \circ x)),\ 0 \le t \le T\}$. We have

$$d\left[x + \int_t^{t+h} f_\tau \mathrm{d}\tau + \int_t^{t+h} g_\tau \mathrm{d}B_\tau, \mathcal{K}(t+h)\right] \le h_\varepsilon \cdot \varepsilon$$

for every $(t, x) \in Graph(\mathcal{K})$ and $(f, g) \in S_{\mathbb{F}}(F \circ x) \times S_{\mathbb{F}}(G \circ x)$. Then

$$\overline{D}\left[x + \int_t^{t+h} F(\tau, x)\mathrm{d}\tau + \int_t^{t+h} G(\tau, x)\mathrm{d}B_\tau, \mathcal{K}(t+h)\right] \le h_\varepsilon \varepsilon,$$

which implies that

$$\liminf_{h\to 0+}(1/h)\overline{D}\left(x + \int_t^{t+h} F(\tau, x)\mathrm{d}\tau + \int_t^{t+h} G(\tau, x)\mathrm{d}B_\tau, \mathcal{K}(t+h)\right) = 0$$

for every $(t, x) \in Graph(\mathcal{K})$. □

Remark 2.4. The results of Theorems 2.1 and 2.2 also hold if instead of condition (2.1), we assume that $[S^t_{\mathbb{F}}(F \circ x) \times S^t_{\mathbb{F}}(G \circ x)] \cap \mathcal{T}_K(t, x) \ne \emptyset$ for every $\varepsilon \in (0, 1)$ and $(t, x) \in Graph(\mathcal{K}^\varepsilon)$. □

There are another types of stochastic tangent sets. For a given $\mathbb{F}$-adapted set-valued stochastic process $K : [0, T] \times \Omega \to \mathrm{Cl}(\mathbb{R}^d)$ and $(t, x) \in Graph(\mathcal{K})$, by $\mathcal{S}_K(t, x)$ we denote the stochastic "tangent set" to K at (t, x) with respect to the filtration $\mathbb{F}$ defined as the set of all pairs $(f, g) \in \mathbb{L}^2([t, T] \times \Omega, \Sigma^t_{\mathbb{F}}, \mathbb{R}^d) \times \mathbb{L}^2([t, T] \times \Omega, \Sigma^t_{\mathbb{F}}, \mathbb{R}^{d\times m})$ such that for every $(f, g) \in \mathcal{S}_K(t, x)$, there exist a sequence $(h_n)_{n=1}^\infty$ of positive numbers converging to 0 and sequences $(a^n)_{n=1}^\infty$ and $(b^n)_{n=1}^\infty$ of d- and $d \times m$-dimensional $\mathbb{F}$-adapted stochastic processes $a^n = (a^n_t)_{0\le t\le T}$ and $b^n = (b^n_t)_{0\le t\le T}$, respectively, such that

$$\sup_{n\ge 1} d\left[x + \int_t^{t+h_n} (f_\tau + a^n_s)\mathrm{d}\tau + \int_t^{t+h_n} (g_\tau + b^n_s)\mathrm{d}B_\tau, \mathcal{K}(t+h_n)\right] = 0$$

and

$$\lim_{n\to\infty} (1/h_n)E\left[\left|\int_t^{t+h_n} a^n_\tau \mathrm{d}\tau + \int_t^{t+h_n} b^n_\tau \mathrm{d}B_\tau\right|^2\right]^{1/2} = 0.$$

We shall show that such stochastic tangent sets are smaller then $\mathcal{T}_K(t,x)$, i.e., that $\mathcal{S}_K(t,x) \subset \mathcal{T}_K(t,x)$ for every $(t,x) \in Graph(\mathcal{K})$.

Lemma 2.2. *Let* $K : [0,T] \times \Omega \to \mathrm{Cl}(\mathbb{R}^d)$ *be an* $\mathbb{F}$*-adapted set-valued process such that* $\mathcal{K}(t) \neq \emptyset$ *for every* $0 \leq t \leq T$*. For every* $(t,x) \in Graph(\mathcal{K})$*, one has* $\mathcal{S}_K(t,x) \subset \mathcal{T}_K(t,x)$.

Proof. Let $(t,x) \in Graph(\mathcal{K})$ be fixed and $(f,g) \in \mathcal{S}_K(t,x)$. There exist a sequence $(h_n)_{n=1}^{\infty}$ of positive numbers converging to 0 and sequences $(a^n)_{n=1}^{\infty}$ and $(b^n)_{n=1}^{\infty}$ of d- and $d \times m$-dimensional $\mathbb{F}$-adapted stochastic processes $a^n = (a_t^n)_{0\leq t\leq T}$ and $b^n = (b_t^n)_{0\leq t\leq T}$, respectively, such that the above conditions are satisfied. For every $n \geq 1$, one has

$$d^2\left[x + \int_t^{t+h_n} f_\tau \mathrm{d}\tau + \int_t^{t+h_n} g_\tau \mathrm{d}B_\tau, \mathcal{K}(t+h)\right] \leq 2E\left[\left|\int_t^{t+h_n} a_\tau^n \mathrm{d}\tau + \int_t^{t+h_n} b_\tau^n \mathrm{d}B_\tau\right|^2\right].$$

Hence, by the properties of sequences $(a^n)_{n=1}^{\infty}$ and $(b^n)_{n=1}^{\infty}$, it follows that

$$\lim_{n\to\infty}(1/h_n)d\left[x + \int_t^{t+h_n} f_\tau \mathrm{d}\tau + \int_t^{t+h_n} g_\tau \mathrm{d}B_\tau, \mathcal{K}(t+h)\right] = 0,$$

which implies

$$\liminf_{h\to 0+}(1/h)d\left(x + \int_t^{t+h} f_\tau \mathrm{d}\tau + \int_t^{t+h} g_\tau \mathrm{d}B_\tau, \mathcal{K}(t-h)\right) = 0.$$

Then $(f,g) \in \mathcal{T}_K(t,x)$ for every $(f,g) \in \mathcal{S}_K(t,x)$. □

Denote by $\tau_K(t,x)$ that stochastic "contingent set" to K at (t,x) with respect to $\mathbb{F}$, defined as the set of all pairs $(f,g) \in \mathbb{L}^2([t,T] \times \Omega, \Sigma_{\mathbb{F}}^t, \mathbb{R}^d) \times \mathbb{L}^2([t,T] \times \Omega, \Sigma_{\mathbb{F}}^t, \mathbb{R}^{d\times m})$ such that for every such pair (f,g), there exist a sequence $(h_n)_{n=1}^{\infty}$ of positive numbers converging to 0 and sequences $(a_n)_{n=1}^{\infty}$ and $(b_n)_{n=1}^{\infty}$ of d- and $d \times m$-dimensional $\mathcal{F}_t$-measurable random variables a_n and b_n, respectively, such that $x + \int_t^{t+h_n} f_s ds + \int_t^{t+h_n} g_s \mathrm{d}B_s + h_n a_n + \sqrt{h_n} b_n \in \mathcal{K}(t+h_n)$ for every $n \geq 1$ and $\max\left\{E|a_n|^2, (1/h_n)E|b_n|^2\right\} \to 0$ as $n \to \infty$. Similarly as above, we obtain the following result.

Lemma 2.3. *Let* $K : [0,T] \times \Omega \to \mathrm{Cl}(\mathbb{R}^d)$ *be an* $\mathbb{F}$*-adapted set-valued process such that* $\mathcal{K}(t) \neq \emptyset$ *for every* $0 \leq t \leq T$*. For every* $(t,x) \in Graph(\mathcal{K})$*, one has* $\tau_K(t,x) \subset \mathcal{S}_K(t,x)$.

Proof. Let $(t,x) \in Graph(\mathcal{K})$ be fixed and $(f,g) \in \tau_K(t,x)$. There are a sequence $(h_n)_{n=1}^{\infty}$ of positive numbers converging to zero and sequences $(a_n)_{n=1}^{\infty}$, $(b_n)_{n=1}^{\infty}$ of $\mathcal{F}_t$-measurable random variables $a_n : \Omega \to \mathbb{R}^d$ and $b_n : \Omega \to \mathbb{R}^{d\times m}$ such that the above conditions are satisfied. For every $n \geq 1$, one gets

$$\sup_{n\geq 1} d\left[x + \int_t^{t+h_n}(f_s + a_n)\mathrm{d}s + \int_t^{t+h_n}(g_s + b_n)\mathrm{d}B_s, \mathcal{K}(t+h)\right] = 0$$

and

$$\left[E\left|\int_t^{t+h_n} a_n \mathrm{d}s + \int_t^{t+h_n} b_n \mathrm{d}B_s\right|^2\right]^{1/2} \leq h_n\left[E|a_n|^2\right]^{1/2} + \sqrt{h_n}\left[E|b_n|^2\right]^{1/2}.$$

Hence, for $n \geq 1$ sufficiently large, it follows that

$$(1/h_n)\left[E\left|\int_t^{t+h_n} a_n \mathrm{d}s + \int_t^{t+h_n} b_n \mathrm{d}B_s\right|^2\right]^{1/2} \leq \left[E|a_n|^2\right]^{1/2} + \left[1/h_n E|b_n|^2\right]^{1/2},$$

which implies that

$$(1/h_n)\left[E\left|\int_t^{t+h_n} a_n \mathrm{d}s + \int_t^{t+h_n} b_n \mathrm{d}B_s\right|^2\right]^{1/2} \to 0 \quad \text{as} \quad n \to \infty.$$

Then $(f, g) \in \mathcal{S}_K(t, x)$. □

Remark 2.5. The results of Theorems 2.1 and 2.2 are also true if instead of condition (2.1), we assume that $[S^t_{\mathbb{F}}(F \circ x) \times S^t_{\mathbb{F}}(G \circ x)] \cap \tau_K(t, x) \neq \emptyset$ for every $\varepsilon \in (0, 1)$ and $(t, x) \in Graph(\mathcal{K}^\varepsilon)$. □

3 Existence of Viable Solutions

We shall prove now that if F and G satisfy conditions $(\mathcal{H}_1)$, then for every continuous set-valued $\mathbb{F}$-adapted process $K : [0, T] \times \Omega \to \mathrm{Cl}(\mathbb{R}^d)$, the viability problems $\overline{SFI}(F, G, K)$ and $BSDI(F, K)$ possess viable strong solutions. Furthermore, the existence of viable weak solutions of $\overline{SFI}(F, G, K)$ is considered. Similarly as above, we define $\mathcal{K}(t)$ and $\mathcal{K}^\varepsilon(t)$ by setting $\mathcal{K}(t) = \{u \in \mathbb{L}^2(\Omega, \mathcal{F}_t, \mathbb{R}^d) : d(u, \mathcal{K}(t)) = 0\}$ and $\mathcal{K}^\varepsilon(t) = \{u \in \mathbb{L}^2(\Omega, \mathcal{F}_t, \mathbb{R}^d) : d(u, \mathcal{K}(t)) \leq \varepsilon\}$.

Theorem 3.1. *Let $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a complete filtered probability space and $B = (B_t)_{0 \leq t \leq T}$ an m-dimensional $\mathbb{F}$-Brownian motion on $\mathcal{P}_{\mathbb{F}}$. Assume that F and G satisfy conditions $(\mathcal{H}_1)$ and let $K : [0, T] \times \Omega \to \mathrm{Cl}(\mathbb{R}^d)$ be an $\mathbb{F}$-adapted set-valued process such that $\mathcal{K}(t) \neq \emptyset$ for every $0 \leq t \leq T$ and such that the mapping $\mathcal{K} : [0, T] \ni t \to \mathcal{K}(t) \in \mathrm{Cl}(\mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d))$ is continuous. If $\mathcal{P}_{\mathbb{F}}$, B, F, G, and K are such that (2.1) is satisfied for every $(t, x) \in Graph(\mathcal{K})$, then the problem $\overline{SFI}(F, G, K)$ possesses on $\mathcal{P}_{\mathbb{F}}$ a strong viable solution.*

Proof. Let $a \in (0, T)$ and select arbitrarily $x_0 \in \mathcal{K}(0)$. Let $u_0 \in \mathbb{L}^2(\Omega, \mathcal{F}_0, \mathbb{R}^d)$ and $v_0 \in \mathbb{L}^2(\Omega, \mathcal{F}_0, \mathbb{R}^{d \times m})$ be such that $u_0 \in F(0, x_0)$ and $v_0 \in G(0, x_0)$ a.s. By virtue of Theorem 2.2, for $\varepsilon_1 = 1/2^{3/2}$ and stochastic processes $\phi^1 = (\phi^1_t)_{a \leq t \leq T}$

and $\psi^1 = (\psi_t^1)_{a\le t\le T}$ defined by $\phi_t^1 = u_0$ and $\psi_t^1 = v_0$ a.s. for every $a \le t \le T$, there exist a partition $0 = t_0^1 < t_1^1 < \cdots < t_{p_1-1}^1 < t_{p_1}^1 = a$, a step function θ_1, and stochastic processes f^1, g^1, and z^1 such that conditions (i)–(v) of Theorem 2.2 are satisfied with

$$x^1(t) = x_0 + \int_0^t (f_\tau^1 + z_\tau^1)\mathrm{d}\tau + \int_0^t g_\tau^1 \mathrm{d}B_\tau$$

a.s. for $a \le t \le T$. Similarly, for $\varepsilon_2 = 1/2$ and $\phi^2 = f^1$ and $\psi^2 = g^1$, we can select a partition $0 = t_0^2 < t_1^2 < \cdots < t_{p_2-1}^2 < t_{p_1}^2 = a$, a step function θ_2, and stochastic processes f^2, g^2, and z^2 such that conditions (i)–(v) of Theorem 2.2 are satisfied with

$$x^2(t) = x_0 + \int_0^t (f_\tau^2 + z_\tau^2)\mathrm{d}\tau + \int_0^t g_\tau^2 \mathrm{d}B_\tau$$

a.s. for $a \le t \le T$. Continuing this procedure for $\varepsilon_k = 1/2^{3k/2}$ and $\phi^k = f^{k-1}$ and $\psi^k = g^{k-1}$, we obtain a partition $0 = t_0^k < t_1^k < \cdots < t_{p_k-1}^k < t_{p_k}^k = a$, a step function θ_k, and stochastic processes f^k, g^k, and z^k such that conditions (i)–(v) of Theorem 2.2 are satisfied for every $k \ge 1$ with

$$x^k(t) = x_0 + \int_0^t (f_\tau^k + z_\tau^k)\mathrm{d}\tau + \int_0^t g_\tau^k \mathrm{d}B_\tau$$

a.s. for $a \le t \le T$ such that $d(x^k(\theta_k(t)), \mathcal{K}(\theta_k(t)) = 0$. For every $k \ge 1$, one has $f^k \in S_{\mathbb{F}}(F \circ (x^{k-1} \circ \theta_{k-1}))$, $g^k \in S_{\mathbb{F}}(G \circ (x^{k-1} \circ \theta_{k-1}))$, $|f^k - f^{k-1}| \le \mathrm{dist}(f_t^{k-1}, F(t, (x^{k-1}(\theta_{k-1}(t)))$, $|g^k - g^{k-1}| \le \mathrm{dist}(g_t^{k-1}, G(t, (x^{k-1}(\theta_{k-1}(t)))$, $\|z^k(t)\| \le \varepsilon_k$, and

$$d\left(x^k(t) - x^k(s), \mathrm{cl}_{\mathbb{L}}\left(\int_s^t F(\tau, (x^k \circ \theta_k)(\tau))\mathrm{d}\tau + \int_s^t G(\tau, (x^k \circ \theta_k)(\tau))\mathrm{d}B_\tau\right)\right) \le \varepsilon_k$$

for $0 \le s \le t \le a$. Furthermore, one has $|\theta_k(t) - \theta_{k-1}(t)| \le \delta_{k-1}$,

$$\int_{\theta_{k-1}(t)}^{\theta_k(t)} |f_\tau^{k-1}|^2 \mathrm{d}\tau \le \varepsilon_{k-1}^2/2^4 \quad \text{and} \quad \int_{\theta_{k-1}(t)}^{\theta_k(t)} |g_\tau^{k-1}|^2 \mathrm{d}\tau \le \varepsilon_{k-1}^2/2^4$$

for $0 \le t \le a$, because by (i) of Theorem 2.2, $\delta_k \in (0, \varepsilon_k)$ is such that

$$\max\left[\sup_{0\le s<t\le T} \int_t^{t+\delta_k} k^2(\tau)\mathrm{d}\tau, \sup_{0\le s<t\le T} \int_t^{t+\delta_k} m^2(\tau)\mathrm{d}\tau\right] \le \varepsilon_k^2/2^4.$$

We shall now evaluate $E[\sup_{0\le\tau\le t} |x^{k+1}(\tau) - x^k(\tau)|^2]$ for $k = 1, 2, \ldots$ and $0 \le t \le a$. Let us observe first that $E[\sup_{0\le\tau\le t} |x^k(\theta_{k+1}(\tau)) - x^k(\theta_k(\tau))|^2] \to 0$ as $k \to \infty$, because $|\theta_{k+1}(t) - \theta_k(t)| \le \delta_k$ and

$$E[\sup_{0\le\tau\le t} |x^k(\theta_{k+1}(\tau)) - x^k(\theta_k(\tau))|^2] \le 3(\delta_k + 1) \int_{\theta_k(t)}^{\theta_{k+1}(t)} m^2(\tau)\mathrm{d}\tau + \varepsilon_k^2$$

for $k = 2, 3, \ldots$ and $0 \le t \le a$. Hence it follows that

$$E[\sup_{0\le\tau\le t} |x^{k+1}(\tau) - x^k(\tau)|^2] \le \alpha\varepsilon_k^2 + \beta \int_0^t k^2(\tau) E[\sup_{0\le u\le\tau} |x^k(u) - x^{k-1}(u)|^2] \mathrm{d}\tau$$

for every $k = 1, 2, \ldots$ and $0 \le t \le a$, where $x_t^0 = x_0$, $\alpha = (4T)^2$ and $\beta = 2^2(T+1)$.

Now, by the definition of the processes x^1 and x^0, one gets $E[\sup_{0\le\tau\le t} |x^1(\tau) - x^0(\tau)|^2] \le \gamma$ with $\gamma = 2^2[(T+1)\int_0^T m^2(t)\mathrm{d}t + T^2]$. Therefore,

$$E[\sup_{0\le\tau\le t} |x^2(\tau) - x^1(\tau)|^2] \le \alpha\varepsilon_1^2 + \beta\gamma \int_0^t k^2(\tau)\mathrm{d}\tau$$

for $0 \le t \le a$. From this and (3), it follows that

$$E[\sup_{0\le\tau\le t} |x^3(\tau) - x^2(\tau)|^2] \le \alpha\varepsilon_2^2 + \alpha\beta\varepsilon_1^2 \int_0^t k^2(\tau)\mathrm{d}\tau + \frac{\beta^2\gamma}{2!}\left(\int_0^t k^2(\tau)\mathrm{d}\tau\right)^2.$$

Similarly, we get

$$\begin{aligned} &E[\sup_{0\le\tau\le t} |x^4(\tau) - x^3(\tau)|^2] \\ &\le \alpha\varepsilon_3^2 + \alpha\beta\varepsilon_2^2 \int_0^t k^2(\tau)\mathrm{d}\tau + \alpha\frac{\beta^2\varepsilon_1^2}{2!}\left(\int_0^t k^2(\tau)\mathrm{d}\tau\right)^2 + \gamma\frac{\beta^3}{3!}\left(\int_0^t k^2(\tau)\mathrm{d}\tau\right)^3 \end{aligned}$$

for $0 \le t \le a$. By the inductive procedure, we obtain

$$\begin{aligned} &E[\sup_{0\le\tau\le t} |x^{n+1}(\tau) - x^n(\tau)|^2] \\ &\le \alpha\varepsilon_n^2 + \alpha\beta\varepsilon_{n-1}^2 \int_0^t k^2(\tau)\mathrm{d}\tau + \alpha\varepsilon_{n-2}^2\frac{\beta^2}{2!}\left(\int_0^t k^2(\tau)\mathrm{d}\tau\right)^2 + \cdots + \gamma\frac{\beta^n}{n!}\left(\int_0^t k^2(\tau)\mathrm{d}\tau\right)^n \\ &\le M\varepsilon_n^2\left[1 + 8\beta\int_0^t k^2(\tau)\mathrm{d}\tau + \frac{(8\beta)^2}{2!}\left(\int_0^t k^2(\tau)\mathrm{d}\tau\right)^2 + \cdots + \frac{(8\beta)^n}{n!}\left(\int_0^t k^2(\tau)\mathrm{d}\tau\right)^n\right] \\ &\le M\varepsilon_n^2 \exp\left[8\beta\int_0^t k^2(\tau)\mathrm{d}\tau\right] \end{aligned}$$

for $n \ge 1$ with $M = \max(\alpha, \gamma)$. By Chebyshev's inequality, we obtain

$$\begin{aligned} &P\left[\sup_{0\le\tau\le a} |x^{n+1}(\tau) - x^n(\tau)| > 2^{-n}\right] \\ &\le 2^{2n} E[\sup_{0\le\tau\le T} |x^{n+1}(\tau) - x^n(\tau)|^2] \le 2^{2n}\varepsilon_n^2 M \exp\left[8\beta\int_0^t k^2(\tau)\mathrm{d}\tau\right] \\ &= 2^{-n} M \exp\left[8\beta\int_0^t k^2(\tau)\mathrm{d}\tau\right]. \end{aligned}$$

Therefore, by the Borel–Cantelli lemma, one gets

$$P\left[\sup_{0\le\tau\le a}|x^{n+1}(\tau)-x^n(\tau)|>2^{-n}\quad\text{for infinitely many } n\right]=0.$$

Thus for a.e. $\omega\in\Omega$, there exists $n_0=n_0(\omega)$ such that $\sup_{0\le\tau\le a}|x^{n+1}(\tau)-x^n(\tau)|\le 2^{-n}$ for $n\ge n_0(\omega)$. Therefore, the sequence $\{x^n(\cdot)(\omega)\}_{n=1}^{\infty}$ is uniformly convergent on $[0,a]$ for a.a. $\omega\in\Omega$, because $x^n(t)(\omega)=x^1(t)(\omega)+\sum_{k=1}^{n-1}[x^{k+1}(t)(\omega)-x^k(t)(\omega)]$ for every $0\le t\le T$ and a.a. $\omega\in\Omega$. Denote the limit of the above sequence by $x_t(\omega)$ for $0\le t\le a$ and a.a. $\omega\in\Omega$. By virtue of (3), it follows that $E[\sup_{0\le\tau\le t}|x^{n+1}(\tau)-x^n(\tau)|^2]\to 0$ as $n\to\infty$. On the other hand, by the properties of sequences $(f^k)_{k=1}^{\infty}$ and $(f^k)_{k=1}^{\infty}$, we get

$$\begin{aligned}\int_0^a E[|f_\tau^{k+1}-f_\tau^k|^2]\mathrm{d}\tau &\le \int_0^a E[h^2(F(\tau,(x^k\circ\theta_k)(\tau))),F(\tau,(x^{k-1}\circ\theta_{k-1})(\tau))))]\mathrm{d}\tau\\ &\le \int_0^a k^2(\tau)E[\sup_{0\le u\le\tau}|x^k(u)-x^{k-1}(u)|^2]\mathrm{d}\tau\end{aligned}$$

and

$$\begin{aligned}\int_0^a E[|g_\tau^{k+1}-g_\tau^k|^2]\mathrm{d}\tau &\le \int_0^a E[H^2(G(\tau,(x^k\circ\theta_k)(\tau))),G(\tau,(x^{k-1}\circ\theta_{k-1})(\tau))))]\mathrm{d}\tau\\ &\le \int_0^a k^2(\tau)E[\sup_{0\le u\le\tau}|x^k(u)-x^{k-1}(u)|^2]\mathrm{d}\tau\end{aligned}$$

for every $k=0,1,\ldots$. Hence it follows that $(f^k)_{k=1}^{\infty}$ and $(g^k)_{k=1}^{\infty}$ are Cauchy sequences of Banach spaces $(\mathbb{L}^2([0,a]\times\Omega,\Sigma_{\mathbb{F}},\mathbb{R}^d),|\cdot|)$ and $(\mathbb{L}^2([0,a]\times\Omega,\Sigma_{\mathbb{F}},\mathbb{R}^{d\times m}),|\cdot|)$, respectively. Then there exist $f\in\mathbb{L}^2([0,a]\times\Omega,\Sigma_{\mathbb{F}},\mathbb{R}^d)$ and $g\in\mathbb{L}^2([0,a]\times\Omega,\Sigma_{\mathbb{F}},\mathbb{R}^{d\times m})$ such that $|f^n-f|\to 0$ and $|g^n-g|\to 0$ as $n\to\infty$. Let $y_t=x_0+\int_0^t f_\tau\mathrm{d}\tau+\int_0^t g_\tau\mathrm{d}B_\tau$ for $0\le t\le a$. For every $n\ge 1$, one gets

$$\begin{aligned}&E[\sup_{0\le t\le a}|x^n(t)-y_t|^2]\\ &\le E[\sup_{0\le t\le a}\left|\int_0^t(f_\tau^n-f_\tau)\mathrm{d}\tau+\int_0^t(g_\tau^n-g_\tau)\mathrm{d}B_\tau+\int_0^t z^n(\tau)\mathrm{d}\tau\right|^2\\ &\le 3T|f^n-f|^2+3|g^n-g||^2+3T^2\varepsilon_n.\end{aligned}$$

Therefore, we have $E[\sup_{0\le t\le a}|x^n(t)-y_t|^2]\to 0$ and $E[\sup_{0\le t\le a}|x^n(t)-x(t)|^2]\to 0$ as $n\to\infty$, which implies that $x(t)=y_t$ a.s. for every $0\le t\le a$. Then $x(t)=x_0+\int_0^t f_\tau\mathrm{d}\tau+\int_0^t g_\tau\mathrm{d}B_\tau$ a.s. for $0\le t\le a$. Now, by Lemma 1.3 and Theorem 2.2, we obtain

$$
\begin{aligned}
0 &\le d\left(x(t)-x(s), \mathrm{cl}_{\mathbb{L}}\left(\int_s^t F(\tau,x(\tau))\mathrm{d}\tau + \int_s^t G(\tau,x(\tau))\mathrm{d}B_\tau\right)\right)\\
&\le ||(x(t)-x(s))-(x^n(t)-x^n(s))||\\
&\quad + d\left(x^n(t)-x^n(s), \mathrm{cl}_{\mathbb{L}}\left(\int_s^t F(\tau,(x^n\circ\theta_n)(\tau))\mathrm{d}\tau + \int_s^t G(\tau,(x^n\circ\theta_n)(\tau))\mathrm{d}B_\tau\right)\right)\\
&\quad + H\left(\mathrm{cl}_{\mathbb{L}}\left(\int_s^t F(\tau,(x^n\circ\theta_n)(\tau))\mathrm{d}\tau + \int_s^t G(\tau,(x^n\circ\theta_n)(\tau))\mathrm{d}B_\tau\right),\right.\\
&\qquad \left.\mathrm{cl}_{\mathbb{L}}\left(\int_s^t F(\tau,x(\tau))\mathrm{d}\tau + \int_s^t G(\tau,x(\tau))\mathrm{d}B_\tau\right)\right)\\
&\le 2||x^n-x|| + \varepsilon_n + (1+\sqrt{T})\left(\int_0^T k^2(t)\mathrm{d}t\right)^{1/2}\|x^n\circ\theta_n - x\|
\end{aligned}
$$

for every $0 \le s \le t \le a$. But

$$
\begin{aligned}
\|x^n\circ\theta_n - x\|^2 &= E[|x^n(\theta_n(t)) - x(t)|]\\
&\le E[\sup_{0\le u\le a}|x^n(u)-x(u)|^2] + E[\sup_{0\le t\le a}|x(\theta_n(t))-x(t)|].
\end{aligned}
$$

Then $\lim_{n\to\infty}\|x^n\circ\theta_n - x\| = 0$. Therefore, for every $0 \le s \le t \le a$, we get

$$
d\left(x(t)-x(s), \mathrm{cl}_{\mathbb{L}}\left(\int_s^t F(\tau,x(\tau))\mathrm{d}\tau + \int_s^t G(\tau,x(\tau))\mathrm{d}B_\tau\right)\right) = 0.
$$

Thus

$$
x(t)-x(s) \in \mathrm{cl}_{\mathbb{L}}\left(\int_s^t F(\tau,x(\tau))\mathrm{d}\tau + \int_s^t G(\tau,x(\tau))\mathrm{d}B_\tau\right)
$$

for every $0 \le s \le t \le a$. In a similar way, we get $d(x(t),\mathcal{K}(t)) \le \|x^n - x\| + d(x^n(t),\mathcal{K}(t)) \le \|x^n - x\| + \varepsilon_n$ for every $n \ge 1$ and $0 \le s \le t \le a$. Therefore, $d(x(t),\mathcal{K}(t)) = 0$ for every $0 \le t \le a$, which by Theorem 3.1 of Chap. 2, implies that $x(t) \in K(t,\cdot)$ a.s. for $0 \le t \le a$.

We can now extend our solution to the whole interval $[0,T]$. Let us denote by Λ_x the set of all extensions of the viable solution x of $\overline{SFI}(F,G,K)$ obtained above. We have $\Lambda_x \ne \emptyset$, because we can repeat the above procedure for every interval $[a,\alpha]$ with $\alpha \in (a,T)$. Let us introduce in Λ_x the partial order relation $\preceq$ by setting $x \preceq z$ if and only if $a_x \le a_z$ and $x = z|_{[0,a_x]}$, where $a_z \in (0,T)$ is such that z is a strong viable solution for $\overline{SFI}(F,G,K)$ on $[0,a_z]$, and $z|_{[0,a_x]}$ denotes the restriction of the solution z to the interval $[0,a_x]$. Let $\psi : [0,\alpha] \to \mathbb{R}^d$ be an extension of x to $[0,\alpha]$ with $\alpha \in (a,T)$ and denote by $P_x^\psi \subset \Lambda_x$ the set containing ψ and all its restrictions $\psi|_{[0,\beta]}$ for every $\beta \in [a,\alpha)$. It is clear that each completely ordered subset of Λ_x is of the form P_x^ψ, determined by some extension ψ of x. It is also clear that every P_x^ψ has ψ as its upper bound. Then by the Kuratowski–Zorn lemma, there exists a maximal element γ of Λ_x. It has to be $a_\gamma = T$. Indeed, if we had $a_\gamma < T$, then we could repeat the above procedure and extend γ as a viable strong solution $\xi \in \Lambda_x$ of $\overline{SFI}(F, G, K)$ to the interval $[0, b]$ with $a_\gamma < b$, which would imply that $\gamma \preceq \xi$, a contradiction to the assumption that γ is a maximal element of Λ_x. Then x can be extended to the whole interval $[0, T]$. □

In a similar way, by virtue of Remark 2.2, we can prove the following existence theorem for $SFI(\overline{F}, G)$.

Theorem 3.2. *Let $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a complete filtered separable probability space and $B = (B_t)_{0 \le t \le T}$ an m-dimensional $\mathbb{F}$-Brownian motion on $\mathcal{P}_{\mathbb{F}}$. Assume that F and G satisfy conditions $(\mathcal{H}_1)$ and that $K : [0, T] \times \Omega \to \mathrm{Cl}(\mathbb{R}^d)$ is $\mathbb{F}$-adapted such that $\mathcal{K}(t) \neq \emptyset$ for every $0 \le t \le T$ and such that the mapping $\mathcal{K} : [0, T] \to \mathrm{Cl}(\mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d))$ is continuous. If $\mathcal{P}_{\mathbb{F}}$, B, F, G, and K are such that (2.2) is satisfied for every $(t, x) \in Graph(\mathcal{K})$, then the problem*

$$\begin{cases} x_t - x_s \in \mathrm{cl}_{\mathbb{L}}\{\int_s^t F(\tau, x_\tau)\mathrm{d}\tau\} + \int_s^t G(\tau, x_\tau)\mathrm{d}B_\tau \quad \text{for } 0 \le s \le t \le T, \\ x_t \in K(t) \ \text{ a.s. for } \ t \in [0, T], \end{cases}$$

possesses on $\mathcal{P}_{\mathbb{F}}$ a strong viable solution. □

We shall now prove the existence of weak viable solutions for stochastic functional inclusions. The proof of such an existence theorem is based on the first viable approximation theorem presented above.

Theorem 3.3. *Assume that $F : [0, T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : [0, T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d \times m})$ are measurable, bounded, convex-valued and are such that $F(t, \cdot)$ and $G(t, \cdot)$ are continuous for a.e. fixed $t \in [0, T]$. Let G be diagonally convex and let $K : [0, T] \to \mathrm{Cl}(\mathbb{R}^d)$ be continuous. If there exist a complete filtered probability space $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ and a d-dimensional $\mathbb{F}$-Brownian motion on $\mathcal{P}_{\mathbb{F}}$ such that (2.1) is satisfied for every $\varepsilon \in (0, 1)$ and $(t, x) \in Graph(\mathcal{K}^\varepsilon)$, then $\overline{SFI}(F, G, K)$ possesses a weak viable solution.*

Proof. Let $x_0 \in \mathcal{K}(0)$ be fixed and let $\varepsilon_n = 1/2^n$. By virtue of Theorem 2.1, we can define on $[0, T]$ a step function $\theta_n = \theta_{\varepsilon_n}$ and $\mathbb{F}$-nonanticipative stochastic processes $f^n = f^{\varepsilon_n}$, $g^n = g^{\varepsilon_n}$, and $x_t^n = x_0 + \int_0^t f_\tau^n \mathrm{d}\tau + \int_0^t g_\tau^n \mathrm{d}B_\tau$ for $0 \le t \le T$ such that conditions (i)–(iii) of Theorem 2.1 are satisfied. In particular, for every $m \ge 1$, $n \ge 1$, and $0 \le s \le t \le T$, we obtain

$$\begin{aligned} E|x^n(t) - x^n(s)|^{2m} &\le C_m^1 E\left[\left|\int_s^t f_\tau^n \mathrm{d}\tau\right|^{2m}\right] + C_m^2 E\left[\left|\int_s^t g_\tau^n \mathrm{d}B_\tau\right|^{2m}\right] \\ &\le C_m^1 T^m E\left(\int_s^t |f_\tau^n|^2 \mathrm{d}\tau\right)^m + C_m^2 E\left[\left|\int_s^t g_\tau^n \mathrm{d}B_\tau\right|^{2m}\right], \end{aligned}$$

where C_m^1 and C_m^2 are positive integers depending on $m \geq 1$. Let us observe that

$$E\left(\int_s^t |f_\tau^n|^2 \mathrm{d}\tau\right)^m \leq M^{2m}|t-s|^m \text{ and } E\left[\left|\int_s^t g_\tau^n \mathrm{d}B_\tau\right|^{2m}\right] \leq M^{2m}(2m-1)!!|t-s|^m.$$

Therefore,

$$E|x^n(t) - x^n(s)|^{2m} \leq \left[C_m^1 T^m + C_m^2(2m-1)!!\right] M^{2m} |t-s|^m$$

for every $0 \leq s \leq t \leq T$ and $n, m \geq 1$. In a similar way, we can verify that there exist positive numbers K and γ such that $E|x_0^n|^\gamma \leq K$. Then the sequence $(x^n)_{n=1}^\infty$ of continuous processes $x^n = (x_t^n)_{0\leq t\leq T}$ satisfies on the probability space $(\Omega, \mathcal{F}, P)$ the assumptions of Theorem 3.5 of Chap. 1. Furthermore, immediately from Theorem 2.1, it follows that $E[\text{dist}(x^n(\theta_n(t)), K(\theta_n(t)))] \leq \varepsilon_n$ and

$$E\left[l(x^n(s))\left(h(x^n(t)) - h(x^n(s)) - \int_s^t (\mathbb{L}_{f^n g^n}^{x^n} h)_\tau \mathrm{d}\tau\right)\right] = 0$$

for every $0 \leq s \leq t \leq T$, $l \in C_b(\mathbb{R}^d, \mathbb{R})$, and $h \in C_b^2(\mathbb{R}^d, \mathbb{R})$.

By virtue of Theorems 3.5 and 2.4 of Chap. 1, 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 d-dimensional continuous stochastic processes $\tilde{x}^{n_k}$ and $\tilde{x}$ on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ for $k = 1, 2, \ldots$ such that $P(x^{n_k})^{-1} = P(\tilde{x}^{n_k})^{-1}$ for $k = 1, 2, \ldots$ and $\sup_{0\leq t\leq T} |\tilde{x}^{n_k} - \tilde{x}| \to 0$ as $k \to \infty$. Let $\tilde{\mathcal{F}}_t^{n_k} = \bigcap_{\varepsilon>0} \sigma(\tilde{x}_u^{n_k} : u \leq t + \varepsilon)$ for $0 \leq t \leq T$ and let $\tilde{\mathbb{F}}_{n_k} = (\tilde{\mathcal{F}}_t^{n_k})_{0\leq t\leq T}$. For every $k \geq 1$, x^{n_k} and $\tilde{x}^{n_k}$ are continuous $\mathbb{F}$- and $\tilde{\mathbb{F}}_{n_k}$-adapted. Furthermore, immediately from (3), it follows that $\mathcal{M}_{FG}^{x^{n_k}} \neq \emptyset$ for every $k \geq 1$, which by Lemma 1.3 of Chap. 4, implies that $\mathcal{M}_{FG}^{\tilde{x}} \neq \emptyset$. This, by Theorem 1.3 of Chap. 4, implies the existence of an $\hat{\mathbb{F}}$-Brownian motion $\hat{B}$ on the standard extension $\hat{\mathcal{P}}_{\hat{\mathbb{F}}} = (\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{F}}, \hat{P})$ of the filtered probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{F}}, \tilde{P})$, with $\tilde{\mathbb{F}} = (\tilde{\mathcal{F}}_t)_{0\leq t\leq T}$ defined by $\tilde{\mathcal{F}}_t = \bigcap_{\varepsilon>0} \sigma(\tilde{x}(u) : u \leq t+\varepsilon)$, such that $(\hat{\mathcal{P}}_{\hat{\mathbb{F}}}, \hat{x}, \hat{B})$ is a weak solution of $\overline{SFI}(F, G)$ with $\hat{x}_t(\hat{\omega}) = \tilde{x}_t(\pi(\hat{\omega}))$ satisfying the initial condition $P\hat{x}_0^{-1} = P\tilde{x}_0^{-1}$, where $\pi : \hat{\Omega} \to \tilde{\Omega}$ is the $(\hat{\mathcal{F}}, \tilde{\mathcal{F}})$-measurable mapping described in the definition of the extension of $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{F}}, \tilde{P})$, because the standard extension $\hat{\mathcal{P}}_{\hat{\mathbb{F}}}$ is also an extension. Similarly as in the proof of Corollary 3.2 of Chap. 1, we obtain $P(e_s \circ x^{n_k})^{-1} = P(e_s \circ \tilde{x}^{n_k})^{-1}$ with $s = \theta_{n_k}(t)$ for $0 \leq t \leq T$. This, together with the inequality $E[\text{dist}(x^{n_k}(\theta_{n_k}(t)), K(\theta_{n_k}(t)))] \leq 1/2^{n_k}$ for $k \geq 1$ and properties of the sequence $(\tilde{x}^{n_k})_{k=1}^\infty$, implies that $E[\text{dist}(\tilde{x}_t, K(t))] = 0$. Similarly as in the proof of Theorem 1.3 of Chap. 4, by the definition of the process $\hat{x}$, it follows that $P\hat{x}^{-1} = P\tilde{x}^{-1}$, which implies that $P(e_t \circ \hat{x})^{-1} = P(e_t \circ \tilde{x})^{-1}$ for every $0 \leq t \leq T$. Therefore, $E[\text{dist}(\hat{x}_t, K(t))] = 0$ for every $0 \leq t \leq T$. Thus $\hat{x}_t \in K(t)$, $\hat{P}$-a.s. for $0 \leq t \leq T$. □

Remark 3.1. The results of Theorem 3.3 again hold if instead of (2.1), we assume that (2.2) is satisfied. It is also true if instead of (2.1), we assume that $[S^t_{\mathbb{F}}(F \circ x) \times S^t_{\mathbb{F}}(G \circ x)] \cap \mathcal{T}_K(t,x) \neq \emptyset$ for every $(t,x) \in Graph(\mathcal{K}^\varepsilon)$ and $\varepsilon \in (0,1)$. □

In a similar way as above, we obtain immediately from Theorem 2.3 the following existence theorem.

Theorem 3.4. *Let $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a complete filtered probability space with a continuous filtration $\mathbb{F} = (\mathcal{F}_t)_{0 \le t \le T}$ such that $\mathcal{F}_T = \mathcal{F}$. Assume that F satisfies conditions $(\mathcal{H}_1)$ and let $K : [0,T] \times \Omega \to \mathrm{Cl}(\mathbb{R}^d)$ be an $\mathbb{F}$-adapted set-valued process such that $\mathcal{K}(t) \neq \emptyset$ for every $0 \le t \le T$ and that the mapping $\mathcal{K} : [0,T] \ni t \to \mathcal{K}(t) \in \mathrm{Cl}(\mathbb{L}(\Omega, \mathcal{F}_T, \mathbb{R}^d))$ is continuous. If $\mathcal{P}_{\mathbb{F}}$, F, and K are such that (2.3) is satisfied for every $(t,x) \in Graph(\mathcal{K})$, then $BSDI(F,K)$ possesses a strong viable solution.*

Proof. Let $x_T \in \mathcal{K}(T)$ and $a \in (0,T)$ be fixed. Put $x^0_t = E[x_T|\mathcal{F}_t]$ a.s. for $a \le t \le T$ and let $f^0 = (f^0_t)_{a \le t \le T}$ be a measurable process on $\mathcal{P}_{\mathbb{F}}$ such that $f^0_t \in \mathrm{co}F(t, (x^0 \circ \theta_0)(t))$ a.s. for a.e. $a \le t \le T$, where $\theta_0(t) = T$ for $a \le t \le T$. Let $\phi_t = f^0_t$ a.s. for a.e. $a \le t \le T$. By virtue of Theorem 2.3, for $\varepsilon_1 = 1/2^{3/2}$ and the above process $\phi = (\phi_t)_{a \le t \le T}$, there exist a partition $a = t^1_{p_1} < t^1_{p_1 - 1} < \cdots < t^1_1 < t^1_0 = T$, a step function $\theta_1 : [a,T] \to [a,T]$, a step process $z^1 = (z^1_t)_{a \le t \le T}$, and a measurable process $f^1 = (f^1_t)_{a \le t \le T}$ on $\mathcal{P}_{\mathbb{F}}$ such that conditions (i)–(vi) of Theorem 2.3 are satisfied. In particular, $f^1_t \in \mathrm{co}F(t, (x^1 \circ \theta_1)(t))$, $|f^1_t - f^0_t| = \mathrm{dist}(f^0_t, \mathrm{co}F(t, (x^1 \circ \theta_1)(t)))$ a.s. for a.e. $a \le t \le T$ and $d(x^1(t), \mathcal{K}(t)) \le \varepsilon_1$ for $a \le t \le T$, because $d(x^1(t), \mathcal{K}(t)) \le |x^1(t) - x^1(\theta(t))| + d(x^1(\theta(t)), \mathcal{K}(\theta(t))) + D(\mathcal{K}(\theta(t)), \mathcal{K}(t)) \le \varepsilon_1$, where $x^1_t = E[x_T + \int_t^T f^0_\tau \mathrm{d}\tau|\mathcal{F}_t] + \int_t^T z^1_\tau \mathrm{d}\tau$ a.s. for $a \le t \le T$. In a similar way, for $\phi = (f^1_t)_{a \le t \le T}$ and $\varepsilon_2 = 1/2^3$, we can define a partition $a = t^2_{p_2} < t^2_{p_2 - 1} < \cdots < t^2_1 < t^2_0 = T$, a step function $\theta_2 : [a,T] \to [a,T]$, a step process $z^2 = (z^2_t)_{a \le t \le T}$, and a measurable process $f^2 = (f^2_t)_{a \le t \le T}$ such that $f^2_t \in \mathrm{co}F(t, (x^2 \circ \theta_2)(t))$, $|f^2_t - f^1_t| = \mathrm{dist}(f^1_t, \mathrm{co}F(t, (x^2 \circ \theta_2)(t)))$ a.s. for a.e. $a \le t \le T$ and $d(x^2(t), \mathcal{K}(t)) \le \varepsilon_2$ for $a \le t \le T$, where $x^2_t = E[x_T + \int_t^T f^1_\tau \mathrm{d}\tau|\mathcal{F}_t] + \int_t^T z^2_\tau \mathrm{d}\tau$ a.s. for $a \le t \le T$. Furthermore, for $i = 1, 2$, we have

$$E\left[\mathrm{dist}\left(x^i(s), E\left[x^i(t) + \int_s^t F(\tau, (x^i \circ \theta_i)(\tau))\mathrm{d}\tau|\mathcal{F}_s\right]\right)\right] \le \varepsilon_i$$

a.s. for $a \le s \le t \le T$. By the inductive procedure, for $\varepsilon_k = 1/2^{3k/2}$ and $\phi^k = (f^k_t)_{a \le t \le T}$, we can select for every $k \ge 1$, a partition $a = t^k_{p_k} < t^k_{p_k - 1} < \cdots < t^k_1 < t^k_0 = T$, a step function $\theta_k : [a,T] \to [a,T]$, a step process $z^k = (z^k_t)_{a \le t \le T}$, and a measurable process $f^k = (f^k_t)_{a \le t \le T}$ such that $f^k_t \in \mathrm{co}F(t, (x^k \circ \theta_k)(t))$, $|f^k_t - f^{k-1}_t| = \mathrm{dist}(f^k_t, \mathrm{co}F(t, (x^k \circ \theta_k)(t)))$ a.s. for a.e. $a \le t \le T$ and $d(x^k(t), \mathcal{K}(t)) \le \varepsilon_k$ for $a \le t \le T$, where

$$x^k_t = E[x_T + \int_t^T f^{k-1}_\tau \mathrm{d}\tau|\mathcal{F}_t] + \int_t^T z^k_\tau \mathrm{d}\tau$$

a.s. for $a \leq t \leq T$. Furthermore,

$$E\left[\text{dist}\left(x^k(s), E\left[x^k(t) + \int_s^t F(\tau, (x^k \circ \theta_k)(\tau))\mathrm{d}\tau | \mathcal{F}_s\right]\right)\right] \leq \varepsilon_k$$

for $a \leq s \leq t \leq T$. Of course, $x^k \in \mathcal{S}(\mathbb{F}, \mathbb{R}^m)$ for $k \geq 1$. By Corollary 3.2 of Chap. 3 and the continuity of the filtration $\mathbb{F}$, the process x^k is continuous for every $k \geq 1$. Furthermore, by the properties of the sequence $(f^k)_{k=1}^{\infty}$, one gets

$$\begin{aligned}|x^{k+1}(t) - x^k(t)| &\leq E\left[\int_t^T |f_\tau^k - f_\tau^{k-1}|^2 \mathrm{d}\tau | \mathcal{F}_t\right] + \int_t^T E|z_\tau^{k+1} - z_\tau^k|\mathrm{d}\tau \\ &\leq E\left[\int_t^T \text{dist}^2(f_\tau^{k-1} \text{ço}\, F(\tau, (x^k \circ \theta_k)(\tau)))\mathrm{d}\tau | \mathcal{F}_t\right] + 8T^2\varepsilon_k \\ &\leq \alpha\varepsilon_k + E\left[\int_t^T k(\tau) \sup_{\tau \leq s \leq T} |x^k(s) - x^{k-1}(s)|\mathrm{d}\tau | \mathcal{F}_t\right],\end{aligned}$$

a.s. for $a \leq t \leq T$, where $\alpha = 8T^2$. Therefore,

$$\begin{aligned}\sup_{t \leq u \leq T} |x^{k+1}(u) - x^k(u)| &\leq \alpha\varepsilon_k + \sup_{t \leq u \leq T} E\left[\int_u^T k(\tau) \sup_{\tau \leq s \leq T} |x^k(s) - x^{k-1}(s)|\mathrm{d}\tau | \mathcal{F}_u\right] \\ &\leq \alpha\varepsilon_k + \sup_{t \leq u \leq T} E\left[\int_t^T k(\tau) \sup_{\tau \leq s \leq T} |x^k(s) - x^{k-1}(s)|\mathrm{d}\tau | \mathcal{F}_u\right]\end{aligned}$$

a.s. for $a \leq t \leq T$ and $k = 1, 2, \ldots$. By Doob's inequality, we get

$$\begin{aligned}&E\left[\sup_{t \leq u \leq T} E[\int_t^T k(\tau) \sup_{\tau \leq s \leq T} |x^k(s) - x^{k-1}(s)|\mathrm{d}\tau | \mathcal{F}_u]\right]^2 \\ &\leq 4E\left[\int_t^T k(\tau) \sup_{\tau \leq s \leq T} |x^k(s) - x^{k-1}(s)|\mathrm{d}\tau]\right]^2\end{aligned}$$

for $a \leq t \leq T$. Therefore, for every $a \leq t \leq T$ and $k = 1, 2, \ldots$, we have

$$E[\sup_{t \leq u \leq T} |x^{k+1}(u) - x^k(u)|^2] \leq \alpha^2\varepsilon_k^2 + \beta \int_t^T k^2(\tau) E[\sup_{\tau \leq s \leq T} |x^k(s) - x^{k-1}(s)|^2]\mathrm{d}\tau,$$

where $\beta = 8T$. By the definitions of x^1 and x^0, we obtain $E[\sup_{t \leq u \leq T} |x^1(u) - x^0(u)|^2] \leq L$, where $L = T\int_0^T m^2(t)\mathrm{d}t$. Therefore,

$$E[\sup_{t \leq u \leq T} |x^2(u) - x^1(u)|^2] \leq 2\alpha^2\varepsilon_1^2 + L\beta \int_t^T k^2(\tau)\mathrm{d}\tau$$

for $a \le t \le T$. Hence it follows that

$$E[\sup_{t\le u\le T}|x^3(u)-x^2(u)|^2] \le 2\alpha\varepsilon_2^2+\alpha\beta\varepsilon_1^2\int_t^T k^2(\tau)\mathrm{d}\tau + L\beta^2\int_t^T k^2(\tau)\left(\int_\tau^T k^2(u)\mathrm{d}u\right)\mathrm{d}\tau$$
$$\le 2\alpha^2\varepsilon_2^2+\alpha^2\beta\varepsilon_1^2\int_t^T k^2(\tau)\mathrm{d}\tau + L\frac{\beta^2}{2!}\left(\int_t^T k^2(\tau)\mathrm{d}\tau\right)^2$$

for $a \le t \le T$. By the inductive procedure, for every $k = 1, 2, \ldots$ and $a \le t \le T$, we obtain

$$E[\sup_{t\le u\le T}|x^{n+1}(u)-x^n(u)|^2]$$
$$\le M\varepsilon_2^2\left[1+(8\beta)\int_t^T k^2(\tau)\mathrm{d}\tau+\frac{(8\beta)^2}{2!}\left(\int_t^T k^2(\tau)\mathrm{d}\tau\right)^2+\cdots+\frac{(8\beta)^n}{n!}\left(\int_t^T k^2(\tau)\mathrm{d}\tau\right)^n\right]$$
$$\le M\varepsilon_n^2 exp\left[8\beta\int_t^T k^2(\tau)\mathrm{d}\tau\right],$$

where $M = \max\{2\alpha^2, L\}$. Hence, by Chebyshev's inequality and the Borel–Cantelli lemma, it follows that the sequence $(x^k)_{k=1}^\infty$ of stochastic processes $(x^k(t))_{a\le t\le T}$ is for a.e. $\omega \in \Omega$ uniformly convergent in $[a, T]$ to a continuous process $(x(t))_{a\le t\le T}$. We can verify that the sequence $(f^k)_{k=1}^\infty$ is a Cauchy sequence of $\mathbb{L}([a, T]\times\Omega, \beta_T\otimes\mathcal{F}_T, \mathbb{R}^m)$. Indeed, for every $k = 0, 1, 2, \ldots$, one has

$$\int_0^a E[|f_\tau^{k+1}-f_\tau^k|]\mathrm{d}\tau$$
$$\le \int_0^a E[H(F(\tau,(x^k\circ\theta_k)(\tau))), F(\tau,(x^{k-1}\circ\theta_{k-1})(\tau))))]\mathrm{d}\tau$$
$$\le \int_0^a k(\tau)E[\sup_{0\le u\le\tau}|x^k(u)-x^{k-1}(u)|]\mathrm{d}\tau,$$

which by the properties of the sequence $(x^k)_{k=1}^\infty$, implies that $(f^k)_{k=1}^\infty$ is a Cauchy sequence. Then there is an $f \in \mathbb{L}([a, T]\times\Omega, \beta_T\otimes\mathcal{F}_T, \mathbb{R}^m)$ such that $|f^k - f| \to 0$ as $k \to \infty$. Let $y_t = E[x_T + \int_t^T f_\tau \mathrm{d}\tau|\mathcal{F}_t]$ a.s. for $a \le t \le T$. For every $k \ge 1$, we have

$$E[\sup_{a\le t\le T}|x(t)-y_t|] \le E[\sup_{a\le t\le T}|x(t)-x_t^k|]+E[\sup_{a\le t\le T}|x^k(t)-y_t|]$$
$$\le E[\sup_{a\le t\le T}|x(t)-x_t^k|]+E\left[\sup_{a\le t\le T}E[\int_t^T|f_\tau^k-f_\tau|\mathrm{d}\tau|\mathcal{F}_t]\right]+\int_t^T E|z_\tau^k|\mathrm{d}\tau$$
$$\le E[\sup_{a\le t\le T}|x(t)-x_t^k|]+E\left[E[\int_0^T|f_\tau^k-f_\tau|\mathrm{d}\tau|\mathcal{F}_t]\right]+T\varepsilon_k^2$$
$$\le E[\sup_{a\le t\le T}|x(t)-x_t^k|]+E\int_0^T|f_\tau^k-f_\tau|\mathrm{d}\tau+T\varepsilon_k^2,$$

which implies that $E[\sup_{a\le t\le T}|x(t)-y_t|]=0$. Then $x(t)=E[x_T+\int_t^T f_\tau \mathrm{d}\tau|\mathcal{F}_t]$ a.s. for $a\le t\le T$. Now, for every $a\le s\le t\le T$, we get

$$E\left[\operatorname{dist}\left(x(s),E\left[x(t)+\int_s^t F(\tau,x(\tau))\mathrm{d}\tau|\mathcal{F}_s\right]\right)\right]$$
$$\le E\left[|x(s)-x^k(s)|\right]+E\left[\operatorname{dist}\left(x^k(s),E\left[x^k(t)+\int_s^t F(\tau,x^k(\theta_k(\tau)))\mathrm{d}\tau|\mathcal{F}_s\right]\right)\right]$$
$$+E\left[H\left(E\left[\int_s^t F(\tau,x^k(\theta_k(\tau)))\mathrm{d}\tau|\mathcal{F}_s\right],E\left[\int_s^t F(\tau,x^k(\tau))\mathrm{d}\tau|\mathcal{F}_s\right]\right)\right]$$
$$+E\left[H\left(E\left[\int_s^t F(\tau,x^k(\tau))\mathrm{d}\tau|\mathcal{F}_s\right],E\left[\int_s^t F(\tau,x(\tau))\mathrm{d}\tau|\mathcal{F}_s\right]\right)\right]$$
$$\le ||x^k-x||+\varepsilon_k+E\int_a^T k(t)|x^k(\theta_k(t))-x^k(t)|\mathrm{d}t+E\int_a^T k(t)|x^k(t)-x(t)|\mathrm{d}t.$$

But

$$E[|x^k(\theta_k(t))-x^k(t)|]\le ||x^k-x||+E[\sup_{a\le t\le T}|x(\theta_k(t))-x^k(t)|]$$

for every $k\ge 1$ and $a\le t\le T$. Then

$$E\left[\operatorname{dist}\left(x(s),E\left[x(t)+\int_s^t F(\tau,x(\tau))\mathrm{d}\tau|\mathcal{F}_s\right]\right)\right]$$
$$\le\left(\int_0^T k(t)\mathrm{d}t\right)\left\{E[\sup_{a\le t\le T}|x(\theta_k(t))-x^k(t)|]+E[\sup_{a\le t\le T}|x(t)-x_t^k|]\right\}$$
$$+\|x^k-x\|+\varepsilon_k\le ||x^k-x||\left(1+\int_0^T k(t)\mathrm{d}t\right)+\varepsilon_k$$

for every $k\ge 1$ and $a\le s\le t\le T$. Thus

$$E\left[\operatorname{dist}\left(x(s),E\left[x(t)+\int_s^t F(\tau,x(\tau))\mathrm{d}\tau|\mathcal{F}_s\right]\right)\right]=0$$

for every $a\le s\le t\le T$. In a similar way, we also get that $d(x(t),\mathcal{K}(t))=0$ for every $a\le t\le T$. Then x is a strong solution of $BSDI(F,K)$ on the interval $[a,T]$ for every $a\in(0,T)$.

We can now extend the above solution to the whole interval $[0,T]$. Let us denote by Λ_x the set of all extensions of the above-obtained viable solution x of $BSDI(F,K)$. We have $\Lambda_x\neq\emptyset$, because we can repeat the above procedure for every interval $[\alpha,T]$ with $\alpha\in(0,a]$ and get a solution x^α of $BSDI(F,K)$ on the interval $[\alpha,T]$. The process $z=\mathbb{1}_{[\alpha,a]}x^\alpha+\mathbb{1}_{(a,T]}x$ is an extension of x to the interval

$[\alpha, T]$. Let us introduce in Λ_x the partial order relation $\preceq$ by setting $x \preceq z$ if and only if $a_z \le a_x$ and $x = z|_{[a_x,T]}$, where $a_x, a_z \in (0, a)$ are such that x and z are strong viable solutions for $BSDI(F, K)$ on $[a_x, T]$ and $[a_z, T]$, respectively, and $z|_{[a_x,T]}$ denotes the restriction of the solution z to the interval $[a_x, T]$. Let $\psi : [\alpha, T] \to \mathbb{R}^d$ be an extension of x to $[\alpha, T]$ with $\alpha \in (0, a]$ and denote by $P_x^\psi \subset \Lambda_x$ the set containing ψ and all its restrictions $\psi|_{[\beta,T]}$ for every $\beta \in (\alpha, a)$. It is clear that each completely ordered subset of Λ_x is of the form P_x^ψ determined by some extension ψ of x and contains its upper bound ψ. Then by the Kuratowski–Zorn lemma, there exists a maximal element γ of Λ_x. It has to be $a_\gamma = 0$, where $a_\gamma \in [0, T)$ is such that γ is a strong viable solution of $BSDI(F, K)$ on the interval $[a_\gamma, T]$. Indeed, if we had $a_\gamma > 0$, then we could repeat the above procedure and extend γ as a viable strong solution $\xi \in \Lambda_x$ of $BSFI(F, K)$ to the interval $[b, T]$ with $0 \le b < a_\gamma$. This would imply that $\gamma \preceq \xi$, a contradiction to the assumption that γ is a maximal element of Λ_x. Then x can be extended to the whole interval $[0, T]$. □

Remark 3.2. Theorem 3.4 is also true if $\mathcal{K}(t) = \{u \in \mathbb{L}(\Omega, \mathcal{F}_0, \mathbb{R}^d) : u \in K(t)\}$. In such a case, instead of (2.3), we can assume that $\liminf_{h\to 0+} \overline{D}(x + \int_{t-h}^t F(\tau, x)\mathrm{d}\tau, \mathcal{K}(t)) = 0$ for every $(t, x) \in Graph(\mathcal{K})$.

Proof. For every $(t, x) \in Graph(\mathcal{K})$, $f \in S(coF \circ x)$, and $u \in \mathcal{K}(t)$, we have

$$\begin{aligned}
E\left(\left|E[x + \int_{t-h}^t f_\tau \mathrm{d}\tau|\mathcal{F}_{t-h}] - u\right|\right) &= E\left(\left|E[x + \int_{t-h}^t f_\tau \mathrm{d}\tau|\mathcal{F}_{t-h}] - E[u|\mathcal{F}_{t-h}]\right|\right) \\
&\le E\left(E\left[\left|x + \int_{t-h}^t f_\tau \mathrm{d}\tau - u\right| \Big| \mathcal{F}_{t-h}\right]\right) \\
&= E\left|x + \int_{t-h}^t f_\tau \mathrm{d}\tau - u\right|.
\end{aligned}$$

Therefore, $d(E[x + \int_{t-h}^t f_\tau \mathrm{d}\tau|\mathcal{F}_{t-h}], \mathcal{K}(t)) \le d(x + \int_{t-h}^t f_\tau \mathrm{d}\tau, \mathcal{K}(t))$ for every $f \in S(coF \circ x)$. Then

$$\overline{D}\left[S(E[x + \int_{t-h}^t F(\tau, x)\mathrm{d}\tau|\mathcal{F}_{t-h}]), \mathcal{K}(t-h)\right] \le \overline{D}\left[x + \int_{t-h}^t F(\tau, x)\mathrm{d}\tau, \mathcal{K}(t-h)\right]$$

for every $(t, x) \in Graph(\mathcal{K})$. Thus, $\liminf_{h\to 0+} \overline{D}(x + \int_{t-h}^t F(\tau, x)\mathrm{d}\tau, \mathcal{K}(t-h)) = 0$ implies that (2.3) is satisfied. □

Remark 3.3. The results of the above existence theorems are also true if instead of continuity of the set-valued mapping $\mathcal{K}$, we assume that it is uniformly upper semicontinuous on $[0, T]$, i.e., that $\lim_{\delta\to 0} \sup_{0\le t\le T} \overline{D}(\mathcal{K}(t+\delta), \mathcal{K}(t)) = 0$. □

It can be verified that the requirement $X_t \in K(t)$ a.s. for $0 \le t \le T$ in the above viability problems is too strong to be satisfied for some stochastic differential equations. For example, the stochastic differential equation $dX_t = f(X_t) + dB_t$ with Lipschitz continuous and bounded function $f : \mathbb{R} \to \mathbb{R}$ does not have any solution $X = (X_t)_{0\le t\le T}$ with X_t belonging to a compact set $K \subset \mathbb{R}$ a.s. for every $0 \le t \le T$. This is a consequences of the following theorem.

Theorem 3.5. *Let $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ be a filtered probability space and $B = (B_t)_{t\geq 0}$ a real-valued $\mathbb{F}$-Brownian motion on $\mathcal{P}_{\mathbb{F}}$. Assume that $\xi = (\xi_t)_{0\leq t\leq T}$ is an Itô diffusion such that $d\xi_t = \alpha_t(\xi)dt + dB_t$, $\xi_0 = 0$ for $0 \leq t \leq T$. Then $P(\{\int_0^T \alpha_t^2(\xi)\mathrm{d}t < \infty\}) = 1$ and $P(\{\int_0^T \alpha_t^2(B)\mathrm{d}t < \infty\}) = 1$ if and only if ξ and B have the same distributions as C_T-random variables on $\mathcal{P}_{\mathbb{F}}$, where $C_T = C([0,T],\mathbb{R})$.*

Example 3.1. Let $f : \mathbb{R} \to \mathbb{R}$ be bounded and Lipschitz continuous. Let $\mathcal{P}_{\mathbb{F}}$ and B be as in Theorem 3.5. Put $\alpha_t(x) = f(e_t(x))$ for $x \in C_T$, where $C_T = C([0,T],\mathbb{R})$ and e_t is the evaluation mapping on C_T, i.e., $e_t(x) = x(t)$ for $x \in C_T$ and $0 \leq t \leq T$. Assume that K is a nonempty compact subset of $\mathbb{R}$ such that $0 \in K$ and consider the viable problem

$$\begin{cases} dX_t = f(X_t)dt + dB_t \quad a.s.\ for\ \ 0 \leq t \leq T, \\ X_t \in K \quad a.s.\ for\ \ t \in [0,T]. \end{cases}$$

Suppose there is a solution X, an Itô diffusion, of the above viability problem such that $X_0 = 0$. By the properties of f, we have $\int_0^T f^2(X_t)\mathrm{d}t < \infty$ and $\int_0^T f^2(B_t)\mathrm{d}t < \infty$ a.s. Therefore, by virtue of Theorem 3.4, for every $A \in \beta(C_T)$ with $PX^{-1}(A) = 1$, one has $PX^{-1}(A) = PB^{-1}(A)$. By the properties of the process X, one has $P(\{X_t \in K\}) = 1$. But $P(\{X_t \in K\}) = P(\{e_t(X) \in K\}) = PX^{-1}(e_t^{-1}(K))$, where e_t is the evolution mapping. Hence it follows that $1 = PX^{-1}(e_t^{-1}(K)) = PB^{-1}(e_t^{-1}(K)) = P(\{B_t \in K\}) < 1$, a contradiction. Then the problem (3) does not have any K-viable strong solution.

Remark 3.4. We can consider viability problems with weaker viable requirements of the form $P(\{X_t \in K(t)\}) \in (\varepsilon, 1)$ for $0 \leq t \leq T$ and $\varepsilon \in (0,1)$ sufficiently large. Solutions to such problems can be regarded as a type of approximations to viable solutions. □

4 Weak Compactness of Viable Solution Sets

Let us denote by $\mathcal{X}(F,G,K)$ the set of (equivalence classes of) all weak viable solutions of $SFI(F,G,K)$. We shall show that for every F, G, and K satisfying the assumptions of Theorem 3.3, the set $\mathcal{X}(F,G,K)$ is weakly compact, i.e., the set $\mathcal{X}^P(F,G,K)$ of distributions of all weak solutions of $SFI(F,G,K)$ is weakly compact subsets of the space $\mathcal{M}(C_T)$ of all probability measures on the Borel σ-algebra $\beta(C_T)$, where $C_T =: C([0,T],\mathbb{R}^d)$.

Theorem 4.1. *Assume that F and G are measurable, bounded, and convex-valued such that $F(t,\cdot)$ and $G(t,\cdot)$ are continuous for a.e. fixed $t \in [0,T]$. Let G be diagonally convex and $K : [0,T] \to \mathrm{Cl}(\mathbb{R}^d)$ continuous. If there exist a complete filtered probability space $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ with a filtration $\mathbb{F}$ satisfying the usual* *conditions and an m-dimensional $\mathbb{F}$-Brownian motion on $\mathcal{P}_{\mathbb{F}}$ such that (2.1) is satisfied for every $\varepsilon \in (0,1)$ and $(t,x) \in Graph(\mathcal{K}^{\varepsilon})$, then the set $\mathcal{X}(F,G,K)$ of all weak viable solutions $(\mathcal{P}_{\mathbb{F}}, x, B)$ of $\overline{SFI}(F,G,K)$ is weakly compact.*

Proof. By virtue of Theorem 3.3, the set $\mathcal{X}(F,G,K)$ is nonempty. Similarly as in the proof of Theorem 4.1 of Chap. 4, we can verify that $\mathcal{X}(F,G,K)$ is relatively weakly compact. We shall prove that it is a weakly closed subset of the space $\mathcal{M}(C_T)$. Let $(x^r)_{r=1}^{\infty}$ be a sequence of $\mathcal{X}(F,G,K)$ convergent in distribution. Then there exists a probability measure $\mathcal{P}$ on $\beta(C_T)$ such that $P(x^r)^{-1} \Rightarrow \mathcal{P}$ as $r \to \infty$. By virtue of Theorem 2.3 of Chap. 1, there are a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ and random variables $\tilde{x}^r : \tilde{\Omega} \to C_T$ and $\tilde{x} : \tilde{\Omega} \to C_T$ for $r = 1,2,\ldots$ such that $P(x^r)^{-1} = P(\tilde{x}^r)^{-1}$ for $r = 1,2,\ldots$, $\tilde{P}(\tilde{x})^{-1} = \mathcal{P}$ and $\lim_{r\to\infty}\sup_{0\le t\le T}|\tilde{x}_t^r - \tilde{x}_t| = 0$ with $(\tilde{P}.1)$. By Theorem 1.3 of Chap. 4, we have $\mathcal{M}_{FG}^{x_r} \neq \emptyset$ for every $r \geq 1$, which by Theorem 1.5 of Chap. 4, implies that $\mathcal{M}_{FG}^{\tilde{x}} \neq \emptyset$. Therefore, by Theorem 1.3 of Chap. 4, there exist a standard extension $\hat{\mathcal{P}}_{\hat{\mathbb{F}}} = (\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{F}}, \hat{P})$ of $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{F}}, \tilde{P})$ and an m-dimensional Brownian motion $\hat{\mathcal{P}}_{\hat{\mathbb{F}}}$ such that $(\hat{\mathcal{P}}_{\hat{\mathbb{F}}}, \hat{x}, \hat{B})$ is a weak solution of $\overline{SFI}(F,G,\mu)$ with an initial distribution μ equal to the probability distribution $P\tilde{x}_0^{-1}$. Similarly as in the proof of Theorem 3.3, this solution is defined by $\hat{x}(\hat{\omega}) = \tilde{x}(\pi(\hat{x}))$ for $\hat{\omega} \in \hat{\Omega}$. Similarly as in the proof of Theorem 4.1 of Chap. 4, we obtain $P(x^r)^{-1} \Rightarrow P(\hat{x})^{-1}$ as $r \to \infty$, which by the properties of the sequence $(\tilde{x}^r)_{n=1}^{\infty}$ implies that $P(\tilde{x}^r)^{-1} \Rightarrow P(\hat{x})^{-1}$ as $r \to \infty$. By the properties of the sequence $(x^r)_{r=1}^{\infty}$, we have $E^r[\text{dist}(x^r(t), K(t))] = 0$ for every $r \geq 1$, which implies that $\tilde{E}[\text{dist}(\tilde{x}^r(t), K(t))] = 0$ for every $r \geq 1$. Hence, by the continuity of the mapping $\text{dist}(\cdot, K(t))$ and properties of the sequence $(\tilde{x}^r)_{n=1}^{\infty}$, it follows that $\hat{E}[\text{dist}(\hat{x}_t, K(t))] = 0$. Thus $(\hat{\mathcal{P}}_{\hat{\mathbb{F}}}, \hat{x}, \hat{B})$ is a weak solution of $\overline{SFI}(F,G,\mu)$, with an appropriately chosen initial distribution μ, such that $x^r \Rightarrow \hat{x}$ and $\hat{x}_t \in K(t)$ with $(\hat{P}.1)$ for every $t \in [0,T]$. Then $(\hat{\mathcal{P}}_{\hat{\mathbb{F}}}, \hat{x}, \hat{B}) \in \mathcal{X}(F,G,K)$, and $\mathcal{X}(F,G,K)$ is weakly closed. □

Remark 4.1. The results of Theorem 4.1 continue to hold if instead of (2.1), we assume that $[S_{\mathbb{F}}^t(F \circ x) \times S_{\mathbb{F}}^t(G \circ x)] \cap \mathcal{T}_K(t,x) \neq \emptyset$ for every $(t,x) \in \mathcal{K}^{\varepsilon}$ and $\varepsilon \in (0,1)$. □

5 Notes and Remarks

The viability approach to optimal control problems is especially useful for problems with state constraints. There is a great number of papers dealing with viability problems for differential inclusions. The first results dealing with viability problems for differential inclusions were given by Aubin and Cellina in [5]. The first result extending to the stochastic case of Nagumo's viability theorem due to Aubin and Da Prato [7]. Most of the results concerning this topic have now been collected in the excellent book by Aubin [6]. Interesting generalizations of viability and invariance problems were given by Plaskacz [88]. A new approach to viability problems for stochastic differential equations was initiated by Aubin and Da Prato in [8] and [9] and by Millian in [79]. Later on, these results were extended by Aubin, Da Prato, and Frankowska [10, 12] in the case of stochastic inclusions written in differential form. Independently, viability problems for stochastic inclusions were also considered by Kisielewicz in [54] and Motyl in [85]. Viability theory provides geometric conditions that are equivalent to viability or invariance properties. Illustrations of viability approach to some issues in control theory and dynamical games with the problem of dynamic valuation and management of a portfolio, can be found in Aubin et al. [13]. The stochastic viability condition presented in Example 3.1 was constructed by M. Michta. The results contained in the present chapter are mainly based on methods applied in Aitalioubrahim and Sajid [3], Van Benoit and Ha [18], and Aubin and Da Prato [9]. The main results of this chapter dealing with the existence of viable strong and weak solutions of stochastic and backward stochastic inclusions and weak compactness with respect to convergence in the sense of distributions of viable weak solution sets are due to the author of this book.

Chapter 6
Partial Differential Inclusions

The present chapter is devoted to partial differential inclusions described by the semielliptic set-valued partial differential operators $\mathbb{L}_{FG}$ generated by given set-valued mappings F and G. Such inclusions will be investigated by stochastic methods. As in the theory of ordinary differential inclusions, the existence of solutions of such inclusions follows from continuous selections theorems and existence theorems for partial differential equations. Therefore, Sects. 2 and 3 are devoted to existence and representation theorems for elliptic and parabolic partial differential equations. Some selection theorems and existence and representation theorems for such partial differential inclusions follow. It will be proved that solutions of initial and boundary value problems for partial differential inclusions can be described by weak solutions of stochastic functional inclusions $SFI(F,G)$, as considered in Chap. 4.

1 Set-Valued Diffusion Generators

In the theory of Kolmogorov–Feller diffusion processes, diffusions are represented by their infinitesimal generators defined for continuous functions $f : \mathbb{R}\mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^d$ and $g : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^{d\times m}$ by setting

$$(\mathcal{A}_{fg}h)(s,x) = h_t'(s,x) + \sum_{i=1}^{d} f^i(s,x)h_{x_i}'(s,x) + \frac{1}{2}\sum_{i,j=1}^{d} \sigma^{ij}(s,x)h_{x_ix_j}''(s,x)$$

for every $h \in C_0^{1,2}(\mathbb{R}^{d+1})$ and $(s,x) \in \mathbb{R}^+ \times \mathbb{R}^d$ with $(\sigma^{ij})_{d\times d} = g \cdot g^*$. Here $C_0^{1,2}(\mathbb{R}^{d+1})$ denotes the space of all continuous functions $h : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}$ with compact support in $\mathbb{R}^{d+1}$ having continuous derivatives h_t', h_{x_i}' and $h_{x_ix_j}''$ for $i,j = 1,2,\ldots,d$. We extend this notion to the set-valued case and speak of set-valued diffusion generators.

M. Kisielewicz, *Stochastic Differential Inclusions and Applications*, Springer Optimization and Its Applications 80, DOI 10.1007/978-1-4614-6756-4_6,

Let $F : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ be given set-valued mappings and denote by $\mathbb{L}_{uv}$ a semielliptic partial differential operator defined on the space $C_0^2(\mathbb{R}^{d+1})$ by

$$(\mathbb{L}_{uv}h)(s,x) = \sum_{i=1}^{d} u_i h'_{x_i}(s,x) + \frac{1}{2}\sum_{i,j=1}^{d} v_{ij} h''_{x_i x_j}(s,x)$$

for $(s,x) \in \mathbb{R}^+ \times \mathbb{R}^d$, $u = (u_1,\dots,u_d)$, and $v = (v_{ij})_{d\times d}$. We define now on $C_0^2(\mathbb{R}^{d+1})$ a set-valued partial differential operator $\mathbb{L}_{FG}$ corresponding to F and G by setting

$$(\mathbb{L}_{FG}h)(s,x) = \{(\mathbb{L}_{uv}h)(s,x) : u \in F(s,x),\ v \in D(G)(s,x)\}$$

for $h \in C_0^2(\mathbb{R}^{d+1})$ and $(s,x) \in \mathbb{R}^+ \times \mathbb{R}^d$, where $D(G)(s,x) = \{g \cdot g^* : g \in G(s,x)\}$.

Let $\mathcal{C}(F)$ and $\mathcal{C}(G)$ denote the sets of all continuous selectors of F and G, respectively. Immediately from the above definitions, it follows that for every $(f,g) \in \mathcal{C}(F) \times \mathcal{C}(G)$ and $\tilde{h} \in \mathcal{D}_{FG} := \bigcup\{\mathcal{D}_{fg}(\mathbb{R}^{n+1}) : (f,g) \in \mathcal{C}(F) \times \mathcal{C}(G)\}$, one has $(\mathcal{A}_{fg}\tilde{h})(s,x) - \tilde{h}'_t(s,x) \in (\mathbb{L}_{FG}\tilde{h})(s,x)$ for $(s,x) \in [0,T] \times \mathbb{R}^d$, where $\mathcal{D}_{fg}(\mathbb{R}^{d+1})$ denotes the domain of the infinitesimal generator $\mathcal{A}_{fg}$ of a $(d+1)$-dimensional Itô diffusion $Y_{s,x}^{fg}$ defined for (f,g) as given above by Theorem 11.1 of Chap. 1. Let us observe that $\mathcal{D}_{FG} \neq \emptyset$, because by virtue of Theorem 10.1 of Chap. 1, one has $C_0^2(\mathbb{R}^{d+1}) \subset \mathcal{D}_{FG}$. The set-valued operator $\mathbb{L}_{FG}$ is called a semielliptic set-valued diffusion operator.

Corollary 1.1. *For every $h \in C_0^2(\mathbb{R}^{d+1})$ and all selectors f and g of F and G, respectively, the function $\mathbb{L}_{uv}h : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}$ with $u = f(t,x)$ and $v = (g \cdot g^*)(t,x)$ is a selector of a set-valued mapping $\mathbb{L}_{FG}h : \mathbb{R}^+ \times \mathbb{R}^d \ni (s,x) \to (\mathbb{L}_{FG}h)(s,x) \in \mathcal{P}(\mathbb{R})$. It is a measurable, continuous Carathéodory selector of $\mathbb{L}_{FG}h$ if f and g are measurable, continuous Carathéodory selectors of F and G, respectively.* □

Apart from the above-defined semielliptic set-valued operator $\mathbb{L}_{FG}$, we shall also consider the family $\mathcal{A}_{FG}$ of parabolic diffusion generators $\mathcal{A}_{fg}$ defined by

$$(\mathcal{A}_{FG}\tilde{h})(s,x) = \{(\mathcal{A}_{fg}\tilde{h})(s,x) : f \in \mathcal{C}(F),\ g \in \mathcal{C}(G)\}$$

for $\tilde{h} \in \mathcal{D}_{FG}$ and $(s,x) \in \mathbb{R}^+ \times \mathbb{R}^d$. Immediately from the relation $(\mathcal{A}_{fg}\tilde{h})(s,x) - \tilde{h}'_t(s,x) \in (\mathbb{L}_{FG}\tilde{h})(s,x)$, it follows that $(\mathcal{A}_{FG}\tilde{h})(s,x) - \tilde{h}'_t(s,x) \subset (\mathbb{L}_{FG}\tilde{h})(s,x)$ for every $\tilde{h} \in \mathcal{D}_{FG}$ and $(s,x) \in [0,T] \times \mathbb{R}^d$. In a similar way, we define a family $\mathcal{L}_{FG}$ of characteristic operators $\mathcal{L}_{fg}$ of the form

$$(\mathcal{L}_{FG}h)(s,x) = \{(\mathcal{L}_{fg}h)(s,x) : f \in \mathcal{C}(F), g \in \mathcal{C}(G)\}$$

for $(s,x) \in [0,T] \times \mathbb{R}^d$ and $h \in \mathcal{C}_{FG} := \bigcup\{\mathcal{C}_{fg}(\mathbb{R}^{d+1}) : f \in \mathcal{C}(F), g \in \mathcal{C}(G)\}$, where $\mathcal{C}_{fg}(\mathbb{R}^{d+1})$ denotes the domain of the characteristic operator of the diffusion $Y_{s,x}^{fg}$ defined above. In what follows, we shall need the following properties of the set-valued mapping $D(G)$ defined by $D(G)(t,x) = \{g \cdot g^* : g \in G(t,x)\}$ for $(t,x) \in \mathbb{R}^+ \times \mathbb{R}^d$.

Lemma 1.1. *For every set-valued mapping $G : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$, one has the following:*

(i) *If G is measurable, then $D(G)$ possesses a measurable selector.*
(ii) *If G is convex-valued and l.s.c., then $D(G)$ possesses a continuous selector.*
(iii) *If G is bounded, diagonally convex, and l.s.c., then $D(G)$ possesses a continuous selector.*
(iv) *If G is convex-valued, bounded, and Lipschitz continuous, then $D(G)$ possesses a Lipschitz continuous selector.*
(v) *If G is a convex-valued Carathéodory set-valued mapping, then $D(G)$ possesses a Carathéodory selector.*
(vi) *If G is a diagonally convex bounded Carathéodory set-valued mapping, then $D(G)$ possesses a Carathéodory selector.*
(vii) *If G is measurable and bounded, and σ is a measurable selector for $D(G)$, then there exists a measurable selector g of G such that $\sigma = g \cdot g^*$.*

Proof. Let $l : \mathbb{R}^{d\times m} \to \mathbb{R}^{d\times d}$ be defined by $l(u) = u \cdot u^*$ for $u \in \mathbb{R}^{d\times m}$. It is easy to see that l is continuous. Indeed, let $u_0 \in \mathbb{R}^{d\times m}$ and let $(u_n)_{n=1}^{\infty}$ be a sequence of $\mathbb{R}^{d\times m}$ converging in the norm topology of $\mathbb{R}^{d\times m}$ to u_0. There is $M > 0$ such that $\|u_0\| \le M$ and $\|u_n\| \le M$ for $n \ge 1$. For every $n \ge 1$, one has

$$\begin{aligned}\|l(u_n) - l(u_0)\| &= \|u_n \cdot u_n^* - u_0 \cdot u_0^*\| \le \|(u_n - u_0)\cdot u_n^*\| + \|u_0 \cdot (u_n^* - u_0^*)\| \\ &\le M(\|u_n - u_0\| + \|u_n^* - u_0^*\|) = 2M\|u_n - u_0\|.\end{aligned}$$

Then $\|l(u_n) - l(u_0)\| \to 0$ as $n \to \infty$.

(i) By the Kuratowski and Ryll-Nardzewski measurable selection theorem, there is a measurable selector g of G. Then $\sigma = l(g)$ is a measurable selector of $D(G)$ because $D(G) = l(G)$.
(ii) Similarly, if G is convex-valued, then by Michael's continuous selection theorem, there exists a continuous selector g of G, which by the continuity of l implies that $\sigma = l(g)$ is a continuous selector of $D(G)$.
(iii) By properties of G and the relation $D(G) = l(G)$, the multifunction $D(G)$ satisfies the conditions of Michael's theorem. Therefore, it has a continuous selector.
(iv) By virtue of Theorem 2.3 of Chap. 2, a multifunction G possesses a Lipschitz continuous selector g. Similarly as above, we obtain

$$\begin{aligned}\|\sigma(t,x) - \sigma(\bar{t},\bar{x})\| &= \|l(g(t,x)) - l(g(\bar{t},\bar{x}))\| \le 2M\|g(t,x) - g(\bar{t},\bar{x})\| \\ &\le 2ML(|t-\bar{t}| + |x-\bar{x}|)\end{aligned}$$

for $(t,x),(\bar{t},\bar{x}) \in \mathbb{R}^+ \times \mathbb{R}^d$, where $\sigma = l(g)$, $L > 0$, is a Lipschitz constant of G and $M > 0$ is such that $\|G(t,x)\| \le M$ for every $(t,x) \in \mathbb{R}^+ \times \mathbb{R}^d$. Then σ is a Lipschitz continuous selector for $D(G)$.

Conditions (v) and (vi) follow immediately from Theorem 2.7 of Chap. 2. Finally, if G and σ satisfy (vii), then $\sigma(t,x) \in l(G(t,x))$ for $(t,x) \in \mathbb{R}^+ \times \mathbb{R}^d$. Therefore, the existence of a measurable selector g of G such that $\sigma = g \cdot g^*$ follows immediately from Theorem 2.5 of Chap. 2. □

Similarly as above, for a given set-valued mapping $G : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$, by $\mathcal{C}(D(G))$ we denote the set of all continuous selectors of $D(G)$. In what follows, we shall consider set-valued mappings G satisfying the following condition $(\mathcal{P})$:

$(\mathcal{P})$ *For every $\sigma = (\sigma_{ij})_{d\times d} \in \mathcal{C}(D(G))$ there are a nonempty set $Q \subset \mathbb{R}^+ \times \mathbb{R}^d$ and a positive number α such that $\sum_{i,j}^d \sigma_{ij}(t,x)\xi_i\xi_j \ge \alpha|\xi|^2$ for $(t,x) \in Q$ and $\xi \in \mathbb{R}^d$.*

We say that a pair (F,G) of set-valued mappings generates a uniformly semielliptic diffusion operator $\mathbb{L}_{FG}$ on $Q \subset \mathbb{R}^+ \times \mathbb{R}^d$ if $\mathcal{C}(F)$ and $\mathcal{C}(G)$ are nonempty and G satisfies condition $(\mathcal{P})$. If $Q = \mathbb{R}^+ \times \mathbb{R}^d$, we simply say that a pair (F,G) generates a uniformly semielliptic diffusion operator. It is clear that if (F,G) generates a uniformly semielliptic diffusion operator, then $(\mathcal{A}_{FG}h)(s,x)$ is nonempty for every $h \in \mathcal{D}_{FG}$ and $(s,x) \in [0,T] \times \mathbb{R}^d$.

Corollary 1.2. *If a pair (F,G) of set-valued mappings generates a uniformly semielliptic diffusion operator $\mathbb{L}_{FG}$ on $Q \subset \mathbb{R}^+ \times \mathbb{R}^d$, then for every $g \in \mathcal{C}(G)$, the symmetric matrix $l(g)$ is continuous and uniformly positive definite, i.e., $\sup_{(t,x)\in Q} \inf_{i\le i\le d} \lambda_i(g)(t,x) > 0$, where $\lambda_1(g)(t,x),\ldots,\lambda_d(g)(t,x)$ denote the eigenvalues of $l(g(t,x))$ for fixed $(t,x) \in Q$.* □

Corollary 1.3. *If a pair (F,G) of bounded set-valued mappings generates a uniformly semielliptic diffusion operator $\mathbb{L}_{FG}$, then for every $(s,x) \in \mathbb{R}^+ \times \mathbb{R}^d$, there exists a $(d+1)$-dimensional Itô diffusion $Y = (Y_t)_{t\ge 0}$ defined by $Y_t = (s+t, X_{s+t})$ on the filtered probability space $\mathcal{P}_{\mathbb{F}} = (\Omega,\mathcal{F},\mathbb{F},P)$ such that its infinitesimal generator $\mathcal{A}_Y$ satisfies $(\mathcal{A}_Y h)(t,x) - h_t'(t,x) \in (\mathbb{L}_{FG}h)(t,x)$ for $h \in C_0^{1,2}(\mathbb{R}^{d+1})$ and $(t,x) \in [s,\infty) \times \mathbb{R}^d$, where $X = (X_t)_{t\ge s}$ is a weak solution of $SFI(F,G)$ on $\mathcal{P}_{\mathbb{F}}$ such that $X_s = x$ a.s.*

Proof. By the properties of the set-valued mappings F and G, we have $\mathcal{C}(F) \neq \emptyset$ and $\mathcal{C}(G) \neq \emptyset$, and for every $g \in \mathcal{C}(G)$, the matrix-value function $\sigma = g \cdot g^*$ is continuous and uniformly positive definite. Then every pair $(f,g) \in \mathcal{C}(F) \times \mathcal{C}(G)$ satisfies the assumptions of Theorem 11.1 of Chap. 1. Therefore, by virtue of this theorem, there is a $(d+1)$-dimensional Itô diffusion $Y = (Y_t)_{t\ge 0}$ on the filtered probability space $\mathcal{P}_{\mathbb{F}}$ defined by $Y_t = (s+t, X_{s+t})$, where $X = (X_t)_{t\ge s}$ is a weak solution of the stochastic differential equation $SDE(f,g)$ on $\mathcal{P}_{\mathbb{F}}$ such that $X_s = x$ a.s. It is clear that $X = (X_t)_{t\ge s}$ is a weak solution of the stochastic functional inclusion $SFI(F,G)$ on $\mathcal{P}_{\mathbb{F}}$ such that $X_s = x$ a.s. Furthermore, we have

$$(\mathcal{A}_Y h)(t,x) = h_t'(t,x) + \sum_{i=1}^{d} f^i(t,x) h_{x_i}'(t,x) + \frac{1}{2}\sum_{i,j=1}^{d} \sigma^{ij}(t,x) h_{x_i x_j}''(t,x)$$

for every $h \in C_0^{1,2}(\mathbb{R}^{d+1})$ and $(t,x) \in [s,\infty) \times \mathbb{R}^d$. Together with the definition of $\mathbb{L}_{FG}$, it follows that $(\mathcal{A}_Y h)(t,x) - h_t'(t,x) \in (\mathbb{L}_{FG} h)(t,x)$ for $h \in C_0^{1,2}(\mathbb{R}^{d+1})$ and $(t,x) \in [s,\infty) \times \mathbb{R}^d$. □

Remark 1.1. If a pair (F,G) of bounded set-valued mappings generates a uniformly semielliptic diffusion operator $\mathbb{L}_{FG}$, then for every $(f,g) \in \mathcal{C}(F) \times \mathcal{C}(G)$, the infinitesimal generator $\mathcal{A}_Y$ of the diffusion $Y_{s,x}$ defined in Corollary 1.3 will be still denoted by $\mathcal{A}_{fg}$. It satisfies the equality

$$(\mathcal{A}_{fg} h)(s,t) = \lim_{t \to 0} \frac{E^{s,x}[h(Y_{s,x}(t))] - h(s,x)}{t}$$

for every $(s,x) \in \mathbb{R}^+ \times \mathbb{R}^d$ and $h \in \mathcal{D}_{fg}$, where $E^{s,x}$ denotes the mean value operator taken with respect to the probability law $Q^{s,x}$ of $Y_{s,x}$. □

Recall that intuitively, the probability law $Q^{s,x}$ of $Y_{s,x}$ gives the distribution of $(Y_{s,x}(t))_{0\le t\le T}$. To express this mathematically, we let $\mathcal{M}_{s,x}$ be the σ-algebra on Ω generated by the random variables $\Omega \ni \omega \to Y_{s,x}(t)(\omega) \in \mathbb{R}^{d+1}$ with $t \in [s,T]$ and define on $\mathcal{M}_{s,x}$ a probability measure $Q^{s,x}$ such that

$$Q^{s,x}[Y_{s,x}(t_1) \in A_1, \ldots, Y_{s,x}(t_k) \in A_k] = P[Y_{s,x}(t_1) \in A_1, \ldots, Y_{s,x}(t_k) \in A_k]$$

for $0 \le t_i < \infty$, $A_i \in \beta(\mathbb{R}^{d+1})$, and $1 \le i \le k$ with $k \ge 1$.

Remark 1.2. If a pair (F,G) of bounded set-valued mappings generates a uniformly semielliptic diffusion operator $\mathbb{L}_{FG}$, then for all $(f,g) \in \mathcal{C}(F) \times \mathcal{C}(G)$, the $(d+1)$-dimensional Itô diffusions $Y = (Y_{0,x}(t))_{t\ge 0}$ and $Y = (Y_{s,x}(t))_{t\ge 0}$ defined above for $(0,x) \in \mathbb{R}^+ \times \mathbb{R}^d$ and $(s,x) \in \mathbb{R}^+ \times \mathbb{R}^d$, respectively, have the same distributions. □

In what follows, we shall consider set-valued mappings' $F : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ that satisfy some of the following conditions $(\mathcal{A})$:

(i) F and G are bounded and convex-valued.
(ii) F and G are l.s.c.
(iii) G is diagonally convex.
(iv) G satisfies $(\mathcal{P})$ and $\mathcal{C}(D(G)) = l(\mathcal{C}(G))$, i.e., for every $\sigma \in \mathcal{C}(D(G))$, there is $g \in \mathcal{C}(G)$ such that $\sigma = g \cdot g^*$.
(v) F and G are continuous.
(iv') G satisfies $(\mathcal{P})$ and is such that $\mathcal{C}(D(G)) = \mathrm{cl}_C[l(\mathcal{C}(G))]$, where cl_C denotes the closure in the topology of the uniform convergence of continuous

$d\times d$-matrix-valued functions on $\mathbb{R}^+\times\mathbb{R}^d$, i.e., for every $\sigma\in\mathcal{C}(D(G))$, there is a sequence $(g^n)_{n=1}^{\infty}$ of $\in\mathcal{C}(G)$ such that the sequence $\{l(g^n)\}_{n=1}^{\infty}$ converges uniformly to σ.

2 Continuous Selections of Set-Valued Diffusion Operators

Let $F:\mathbb{R}^+\times\mathbb{R}^d\to\mathrm{Cl}(\mathbb{R}^d)$, $G:\mathbb{R}^+\times\mathbb{R}^d\to\mathrm{Cl}(\mathbb{R}^{d\times m})$, $h\in C_0^2(\mathbb{R}^{d+1})$ and a continuous function $u:\mathbb{R}^+\times\mathbb{R}^d\to\mathbb{R}$ be such that $u(t,x)\in(\mathbb{L}_{FG}h)(t,x)$ for $(t,x)\in\mathbb{R}^+\times\mathbb{R}^d$. We are interested in the existence for every $\varepsilon>0$ of continuous selectors $f_\varepsilon\in\mathcal{C}(F)$ and $g_\varepsilon\in\mathcal{C}(G)$ such that $|u(t,x)-(\mathbb{L}_{f_\varepsilon g_\varepsilon}h)(t,x)|\le\varepsilon$ for $(t,x)\in\mathbb{R}^+\times\mathbb{R}^d$, where

$$(\mathbb{L}_{f_\varepsilon g_\varepsilon}h)(t,x)=\sum_{i=1}^d f_\varepsilon^i(t,x)h'_{x_i}(t,x)+\frac{1}{2}\sum_{i,j=1}^d(g_\varepsilon\cdot g_\varepsilon^*)_{ij}(t,x)h''_{x_ix_j}(t,x).$$

Immediately from Theorem 2.2 of Chap. 2, it follows that if F and G satisfy conditions (i)–(iii) of $(\mathcal{A})$, then for every $\varepsilon>0$ and $h\in C_0^2(\mathbb{R}^{d+1})$, there are $f_\varepsilon\in\mathcal{C}(F)$ and $\sigma_\varepsilon\in\mathcal{C}(D(G))$ such that

$$\left|u(t,x)-\left(\sum_{i=1}^d f_\varepsilon^i(t,x)h'_{x_i}(t,x)+\frac{1}{2}\sum_{i,j=1}^d\sigma_\varepsilon^{ij}(t,x)h''_{x_ix_j}(t,x)\right)\right|\le\varepsilon$$

for $(t,x)\in\mathbb{R}^+\times\mathbb{R}^d$. To get the desired result, we have to show that for a given $\sigma\in\mathcal{C}(D)$, there exist a continuous selector $g\in\mathcal{C}(G)$ or a sequence $(g_n)_{n=1}^{\infty}$ of continuous selectors of G such that $\sigma(t,x)=(g\cdot g^*)(t,x)$ or $(g_n\cdot g_n^*)(t,x)\to\sigma(t,x)$ uniformly with respect to $(t,x)\in\mathbb{R}^+\times\mathbb{R}^d$ as $n\to\infty$. Such problems concerning measurable selectors have been affirmatively solved in Lemma 1.1. For continuous selectors, the above problem is much more complicated, although in the case of single-valued mappings $D(G)=\{\sigma\}$ with a positive definite and continuous matrix function $\sigma:\mathbb{R}^+\times\mathbb{R}^d\to\mathbb{R}^{d\times d}$, there exists a continuous mapping $g:\mathbb{R}^+\times\mathbb{R}^d\to\mathbb{R}^{d\times m}$ such that $\sigma=g\cdot g^*$. If σ belongs to the space $C^{1,2}(\mathbb{R}^+\times\mathbb{R}^d,\mathbb{R}^{d\times d})$, then one can also prove the existence of $g\in C^{1,2}(\mathbb{R}^+\times\mathbb{R}^d,\mathbb{R}^{d\times m})$ such that $\sigma=g\cdot g^*$. To consider the general problem concerning continuous selectors, let us introduce in $\mathbb{R}^{d\times m}$ an equivalence relation R by setting xRy if and only if $l(y)=l(x)$, where similarly as above, we put $l(u)=u\cdot u^*$ for $u\in\mathbb{R}^{d\times m}$. Let $\mathcal{X}=\mathbb{R}^{d\times m}$ and let $\tilde{\mathcal{X}}=\mathcal{X}/R$ be the R-quotient space. Let $q:\mathcal{X}\to\tilde{\mathcal{X}}$ be the quotient mapping defined in the usual way by setting $\mathcal{X}\ni x\to q(x)=[x]\in\tilde{\mathcal{X}}$, where $[x]=\{z\in\mathcal{X}:zRx\}$. Denote by $\mathcal{T}_{\mathcal{X}}$ the norm topology in $\mathcal{X}$, and let $\tilde{\mathcal{T}}_l$ be the natural topology in $\tilde{\mathcal{X}}$ defined by $\tilde{\mathcal{T}}_l=\{V\subset\tilde{\mathcal{X}}:q^{-1}(V)\in\mathcal{T}_{\mathcal{X}}\}$. It is clear that q is $(\mathcal{T}_{\mathcal{X}},\tilde{\mathcal{T}}_l)$-continuous. Let us introduce in $\mathcal{X}$ the topology $\mathcal{T}_l=\{q^{-1}(V):V\in\tilde{\mathcal{T}}_l\}$. We have $\mathcal{T}_l\subset\mathcal{T}_{\mathcal{X}}$.

Lemma 2.1. *Given the set-valued mapping* $G : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$, *let* $f : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^{d\times m}$ *be continuous and such that* $l(f(t,x)) \in l(G(t,x))$ *for* $(t,x) \in \mathbb{R}^+ \times \mathbb{R}^d$. *There exists a* $(\mathcal{T}_{\mathbb{R}^{n+1}}, \mathcal{T}_l)$*-continuous selector* g *of* G *such that* $l(f(t,x)) = l(g(t,x))$ *for* $(t,x) \in \mathbb{R}^+ \times \mathbb{R}^d$, *where* $\mathcal{T}_{\mathbb{R}^{d+1}}$ *denotes the norm topology in* $\mathbb{R}^{d+1}$.

Proof. For every $(t,x) \in \mathbb{R}^+ \times \mathbb{R}^d$, we can select $u_x^t \in G(t,x)$ such that $l(f(t,x)) = l(u_x^t)$. Put $g(t,x) = u_x^t$ for $t \in [0,T]$ and $x \in \mathbb{R}^d$. We have $q(g(t,x)) = q(f(t,x))$. By the $(\mathcal{T}_{\mathbb{R}^{d+1}}, \tilde{\mathcal{T}}_l)$-continuity of $q \circ f$ for every $V \in \tilde{\mathcal{T}}_l$, we have $(q \circ f)^{-1}(V) \in \mathcal{T}_{\mathbb{R}^{d+1}}$. Then $g^{-1}(q^{-1}(V)) = f^{-1}(q^{-1}(V)) \in \mathcal{T}_{\mathbb{R}^{d+1}}$ for every $V \in \tilde{\mathcal{T}}_l$. Therefore, for every $U \in \mathcal{T}_l$, we have $g^{-1}(U) = f^{-1}(q^{-1}(V)) \in \mathcal{T}_{\mathbb{R}^{d+1}}$, because by the definition of $\mathcal{T}_l$, for every $U \in \mathcal{T}_l$, there is $V \in \tilde{\mathcal{T}}_l$ such that $U = q^{-1}(V)$. □

In what follows, we shall deal with right triangular matrices. Recall that a matrix $v \in \mathbb{R}^{d\times d}$ with elements v_{ij} is said to be right triangular if $v_{ii} \neq 0$ for $i = 1, 2, \ldots, n$ and all elements of v lying below its main diagonal are equal to zero. It can be proved (see [80], pp. 81–82) that for every symmetric matrix $\sigma \in \mathbb{R}^{d\times d}$ of the rank r such that its minors d_k are nonzero for $k = 1, 2, \ldots, r$, there is a right triangular matrix $v = (v_{ij})_{d\times d}$ such that $\sigma = v \cdot v^*$ and elements v_{ij} of such a matrix v are defined by

$$v_{ij} = \begin{cases} \frac{1}{\sqrt{d_j d_{j-1}}}\, \sigma \begin{pmatrix} 1\,2\,\ldots\,j-1\,i \\ 1\,2\,\ldots\,j-1\,j \end{pmatrix}; & j = 1,2,\ldots,r,\ i = j, j+1,\ldots,d, \\ 0 & ;\ j = r+1, r+2, \ldots, d, \end{cases} \tag{2.1}$$

where $\sigma\begin{pmatrix} i_1 i_2 \ldots i_k \\ j_1 j_2 \ldots j_k \end{pmatrix}$ denotes the kth-order minor consisting of elements of σ lying in the intersection of the k rows with indices $i_1, \ldots, i_k$ and the k columns with indices $j_1, \ldots, j_k$ and $d_p = \sigma\begin{pmatrix} 1\,2\ldots p \\ 1\,2\ldots p \end{pmatrix}$ for $p = 1, 2, \ldots, d$.

We shall now show that if $G : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times d})$ is such that $\det(u_x^t) \neq 0$ for every $(t, x, u_x^t) \in \mathrm{Graph}(G)$ and the set-valued mapping $D(G) : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times d})$, defined by $D(G)(t,x) = l(G(t,x))$ for $t \in \mathbb{R}^+$ and $x \in \mathbb{R}^d$ has a continuous selector σ, then there is a $(\mathcal{T}_{\mathbb{R}^{n+1}}, \mathcal{T}_l)$-continuous selector g of G such that $\sigma(t,x) = l(g(t,x))$ for $t \in \mathbb{R}^+$ and $x \in \mathbb{R}^d$. The result will follow from the properties of positive definite symmetric matrices (see [80], pp. 81, 153).

Lemma 2.2. *Let* $G : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ *be such that* $\det(u_x^t) \neq 0$ *for every* $(t, x, u_x^t) \in \mathrm{Graph}(G)$ *and such that a set-valued mapping* $D(G) = l(G)$ *has a continuous selector* σ. *There exists a* $(\mathcal{T}_{\mathbb{R}^{d+1}}, \mathcal{T}_l)$*-continuous selector* g *of* G *such that* $\sigma(t,x) = l(g(t,x))$ *for* $t \in \mathbb{R}^+$ *and* $x \in \mathbb{R}^d$.

Proof. For every $t \in \mathbb{R}^+$ and $x \in \mathbb{R}^d$, there is $u_x^t \in G(t,x)$ such that $\sigma(t,x) = u_x^t \cdot (u_x^t)^*$ and $\det(u_x^t) \neq 0$. Then (see [80], p. 153) $\sigma(t,x)$ is symmetric and positive definite for every $t \in \mathbb{R}^+$ and $x \in \mathbb{R}^d$. Therefore, for every $t \in \mathbb{R}^+$ and $x \in \mathbb{R}^d$, there is a right triangular matrix $v(t,x) = (v_{ij}(t,x))_{d\times d}$ such that $\sigma(t,x) = v(t,x)\cdot v^*(t,x)$ and such that all its elements $v_{ij}(t,x)$ are defined by (2.1). By the continuity of σ, all its minors are continuous, too. Therefore, by (2.1), all elements v_{ij} of v are continuous on $\mathbb{R}^+ \times \mathbb{R}^d$. Then $v : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$ is a continuous matrix such that $\sigma(t,x) = l(v(t,x))$ for $t \in \mathbb{R}^+$ and $x \in \mathbb{R}^d$. By the definitions of $D(G)$ and σ, we get $l(v(t,x)) \in l(G(t,x))$ for every $t \in \mathbb{R}^+$ and $x \in \mathbb{R}^d$. Therefore, by virtue of Lemma 2.1, there is a $(\mathcal{T}_{\mathbb{R}^{d+1}}, \mathcal{T}_l)$-continuous selector g for G such that $\sigma(t,x) = l(g(t,x))$ for $t \in \mathbb{R}^+$ and $x \in \mathbb{R}^d$. □

Lemma 2.3. *Let $G : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ be l.s.c., diagonally convex, and such that for every $(t,x,u_x^t) \in \mathrm{Graph}(G)$, u_x^t is a right triangular matrix. Then $\mathcal{C}(D(G)) \neq \emptyset$ and for every $\sigma \in \mathcal{C}(D(G))$, there is $g \in \mathcal{C}(G)$ such that $\sigma = l(g)$.*

Proof. By (iii) of Lemma 1.1, there is $\sigma \in \mathcal{C}(D(G))$. Hence, similarly as in the proof of Lemma 2.2, the existence of a continuous right triangular matrix function $v : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$ such that $\sigma = l(v)$ follows. By the properties of G, for every $(t,x) \in \mathbb{R}^+ \times \mathbb{R}^d$, there is a right triangular matrix $u_x^t \in G(t,x)$ such that $\sigma(t,x) = l(u_x^t)$. Hence it follows that for every $(t,x) \in \mathbb{R}^+ \times \mathbb{R}^d$, all elements of matrices $v(t,x)$ and u_x^t are defined by the same formulas (2.1) by the elements of the matrix $\sigma(t,x)$. Therefore, for every $(t,x) \in \mathbb{R}^+ \times \mathbb{R}^d$, one has $v(t,x) = u_x^t \in G(t,x)$. Taking now $g = v$, we get a continuous selector of G satisfying $\sigma = l(g)$. Let us observe that we also have $\sigma = l(-g)$, but $-g$ does not have to be a selector of G. □

In further applications, we are interested in the existence of continuous selectors $g \in \mathcal{C}(G)$ such that the matrix function $l(g)$ is uniformly positive definite. Such selectors can be obtained immediately from Lemma 2.3. Let $\boldsymbol{l}(u_x^t)$ denote the quadratic form on $\mathbb{R}^d$ with matrix $l(u_x^t)$ for every $(t,x,u_x^t) \in \mathrm{Graph}(G)$.

Lemma 2.4. *Let $G : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ be l.s.c., diagonally convex, and such that u_x^t is a right triangular matrix for every $(t,x,u_x^t) \in \mathrm{Graph}(G)$. If furthermore, there is $L > 0$ such that $\boldsymbol{l}(u_x^t)(u,u) > L\|u\|^2$ for every $(t,x,u_x^t) \in \mathrm{Graph}(G)$ and $u \in \mathbb{R}^d$ with $\|u\| \neq 0$, then there exists $g \in \mathcal{C}(G)$ such that $l(g)$ is uniformly positive definite.*

Proof. By (iii) of Lemma 1.1, there is a continuous selector σ of $D(G)$. For every $(t,x,u_x^t) \in \mathrm{Graph}(G)$, we have

$$\lambda = \sup_{t\geq 0} \sup_{x\in\mathbb{R}^d} \min_{u\in\mathbb{R}^d, u\neq 0} [\boldsymbol{l}(u_x^t)(u,u)/\|u\|^2] = \sup_{t\geq 0} \sup_{x\in\mathbb{R}^d} \min_{1\leq i\leq d} \lambda_i(l(u_x^t)),$$

where $\lambda_i(l(u_x^t))$ denotes for $i = 1,2,\ldots,d$ the eigenvalues of the symmetric matrix $l(u_x^t)$. Immediately from the inequality $\boldsymbol{l}(u_x^t)(u,u) > L\|u\|^2$, it follows that $\lambda \geq L > 0$. Then the matrix function $\mathbb{R}^+ \times \mathbb{R}^d \ni (t,x) \to l(u_x^t) \in \mathbb{R}^{d\times d}$ is uniformly positive definite. Similarly as in the proof of Lemma 2.3, we can verify that there exists a continuous right triangular matrix $v : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$ such that $v(t,x) \in G(t,x)$ and $l(v(t,x)) = l(u_x^t)$ for every $(t,x) \in \mathbb{R}^+ \times \mathbb{R}^d$. Taking $g = v$, we obtain a selector $g \in \mathcal{C}(G)$ such that $l(g)$ is uniformly positive definite. □

Corollary 2.1. *Suppose* $G : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ *is l.s.c., diagonally convex, and such that for every* $(t,x,u_x^t) \in \mathrm{Graph}(G)$, u_x^t *is a right triangular matrix. Assume that there is* $L > 0$ *such that* $\boldsymbol{l}(u_x^t)(u,u) > L\|u\|^2$ *for every* $(t,x,u_x^t) \in \mathrm{Graph}(G)$ *and* $u \in \mathbb{R}^d$ *with* $\|u\| \neq 0$. *For every convex-valued l.s.c. set-valued mapping* $F : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$, *the pair* (F,G) *generates a uniformly semielliptic set-valued diffusion operator* $\mathbb{L}_{FG}$.

Proof. By (iii) of Lemma 1.1 and Michael's continuous selection theorem, we have $\mathcal{C}(F) \neq \emptyset$, $\mathcal{C}(G) \neq \emptyset$ and $\mathcal{C}(D(G)) \neq \emptyset$. By virtue of Lemma 2.3, for every $\sigma \in \mathcal{C}(D(G))$, there is $g \in \mathcal{C}(G)$ such that $\sigma = l(g)$. Finally, similarly as in the proof of Lemma 2.4, we obtain that for every $g \in \mathcal{C}(G)$, the matrix function $l(g)$ is uniformly positive definite. Then every $\sigma \in \mathcal{C}(D(G)) = l(\mathcal{C}(G))$ is uniformly positive definite. Thus a pair (F,G) generates a uniformly semielliptic set-valued diffusion operator $\mathbb{L}_{F,G}$. □

Corollary 2.2. *For every symmetric uniformly positive definite matrix function* $g : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$, *the matrix function* $\sigma = l(g)$ *is uniformly positive definite.*

Proof. Let $\lambda = \sup_{t\geq 0}\sup_{x\in\mathbb{R}^d}\min_{1\leq i\leq d}\lambda_i(g)(t,x)$ for $(t,x) \in \mathbb{R}^+ \times \mathbb{R}^d$, where $\lambda_i(g)(t,x)$ denote for every $i = 1,2,\ldots,d$ and $(t,x) \in \mathbb{R}^+ \times \mathbb{R}^d$, the eigenvalues of the symmetric matrix $g(t,x)$. For every $(t,x) \in \mathbb{R}^+ \times \mathbb{R}^d$, one has $\det(g(t,x) - \lambda_i(g)(t,x)E) = 0$, where E is the $d \times d$ unit matrix. Hence it follows that

$$\begin{aligned}\det\left(\sigma(t,x) - \lambda_i^2(g)(t,x)E\right) &= \det[l(g)(t,x) - \lambda_i^2(g)(t,x)E]\\ &= \det\left[(g(t,x) - \lambda_i(g)(t,x)E)\cdot(g(t,x) + \lambda_i(g)(t,x)E)\right]\\ &= \det[g(t,x) - \lambda_i(g)(t,x)E]\cdot\det[g(t,x) + \lambda_i(g)(t,x)E] = 0.\end{aligned}$$

Therefore, $\gamma_i(t,x) =: \lambda_i^2(t,x)$ is for every $i = 1,\ldots,d$ an eigenvalue of the symmetric matrix $\sigma(t,x)$ for $i = 1,2,\ldots,d$ and $(t,x) \in \mathbb{R}^+ \times \mathbb{R}^d$ such that $0 < \lambda^2 = \sup_{t\geq 0}\sup_{x\in\mathbb{R}^d}\min_{1\leq i\leq d}\gamma_i(t,x)$. □

Now we can prove some continuous selection theorems dealing with the problem presented above.

Theorem 2.1. *Let* $F : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ *and* $G : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times d})$ *be l.s.c., bounded, convex-valued, and diagonally convex, respectively set-valued, mappings. Assume that* $\det(u_x^t) \neq 0$ *for every* $(t,x,u_x^t) \in \mathrm{Graph}(G)$ *and let* $\upsilon \in C^{1,2}(\mathbb{R}^+ \times \mathbb{R}^d, \mathbb{R})$ *be such that* $\upsilon_t'(t,x) \in (\mathbb{L}_{FG}\upsilon(t,\cdot))(t,x)$ *for every* $(t,x) \in \mathbb{R}^+ \times \mathbb{R}^d$. *For every* $k = 1,2,\ldots$, *there are continuous and* $(\mathcal{T}_{\mathbb{R}^{d+1}}, \mathcal{T}_l)$-*continuous selectors* f_k *and* g_k *of* F *and* G, *respectively, such that* $|\upsilon_t'(t,x) - (\mathbb{L}_{f_k g_k}\upsilon(t,\cdot))(t,x)| \leq 1/k$ *for* $t \in \mathbb{R}^+$ *and* $x \in \mathbb{R}^d$.

Proof. Let us observe that $F \times D(G) : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d \times \mathbb{R}^{d\times d})$ is l.s.c., bounded, and convex-valued. Put $X = \mathbb{R}^+ \times \mathbb{R}^d$ and $Y = \mathbb{R}^d \times \mathbb{R}^{d\times d}$. Define on $X \times Y$ a function $\lambda[(t,x),(u,\sigma)] = \gamma[\upsilon(t,x),u,\sigma]$ for $(t,x) \in X$ and $(u,\sigma) \in Y$, where $\gamma[\upsilon(t,x),u,\sigma] = \langle \nabla_x \upsilon(t,x), u\rangle + (1/2)tr[\partial_{xx}\upsilon(t,x)\cdot\sigma]$, with $\nabla_x \upsilon(t,x) = [\upsilon'_{x_1}(t,x),\ldots,\upsilon'_{x_d}(t,x)]$ and $\partial_{xx}\upsilon(t,x) = [\upsilon''_{x_i x_j}(t,x)]_{d\times d}$. It is clear that λ is continuous on $X \times Y$ such that $\lambda[(t,x),\cdot\,]$ is affine. By the properties of the function υ, we have $\upsilon'_t(t,x) \in \lambda[(t,x),(F\times D(G))(t,x)]$ for every $(t,x) \in X$. Therefore, by virtue of Theorem 2.2 of Chap. 2, for every $k = 1,2,\ldots$, there exists a continuous selector (f_k,σ_k) of $F \times D(G)$ such that $|\upsilon'_t(t,x) - \lambda[(t,x),(f_k,\sigma_k)(t,x)]| \le 1/k$ for $(t,x) \in X$. We have $f_k \in \mathcal{C}(F)$ and $\sigma_k \in \mathcal{C}(D(G))$ for $k = 1,2,\ldots$. By virtue of Lemma 2.2, for every $k = 1,2,\ldots$, there exists a $(\mathcal{T}_{\mathbb{R}^{d+1}},\mathcal{T}_l)$-continuous selector g_k of G such that $\sigma_k = l(g_k)$, which together with the above properties of the pair (f_k,σ_k) proves the theorem. □

Theorem 2.2. *Let $F : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times d})$ be l.s.c., bounded, convex-valued, and diagonally convex, respectively set-valued, mappings. Assume that G is such that for every $(t,x,u_x^t) \in \mathrm{Graph}(G)$, u_x^t is a right triangular matrix and let $\upsilon \in C^{1,2}(\mathbb{R}^+ \times \mathbb{R}^d,\mathbb{R})$ be such that $\upsilon'_t(t,x) \in (\mathbb{L}_{FG}\upsilon(t,\cdot))(t,x)$ for $t \in \mathbb{R}^+$ and $x \in \mathbb{R}^d$. For every $k = 1,2,\ldots$, there are continuous selectors f_k and g_k of F and G, respectively, such that $|\upsilon'_t(t,x) - (\mathbb{L}_{f_k g_k}\upsilon(t,\cdot))(t,x)| \le 1/k$ for $t \in \mathbb{R}^+$ and $x \in \mathbb{R}^d$.*

Proof. Similarly as in the proof of Theorem 2.1, the result follows immediately from Theorem 2.2 of Chap. 2 and Lemma 2.3. □

Immediately from Theorem 2.2 of Chap. 2 and Lemma 2.4, we obtain the following selection theorems.

Theorem 2.3. *Let $F : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times d})$ be l.s.c., bounded, convex-valued, and diagonally convex, respectively set-valued, mappings. Assume that G is such that for every $(t,x,u_x^t) \in \mathrm{Graph}(G)$, u_x^t is a down triangular matrix and there is $L > 0$ such that $\boldsymbol{l}(u_x^t)(u,u) > L\|u\|^2$ and $u \in \mathbb{R}^d$ with $\|u\| \neq 0$. Let $\upsilon \in C^{1,2}(\mathbb{R}^+ \times \mathbb{R}^d,\mathbb{R})$ be such that $\upsilon'_t(t,x) \in (\mathbb{L}_{FG}\upsilon(t,\cdot))(t,x)$ for $(t,x) \in \mathbb{R}^+ \times \mathbb{R}^d$. For every $k = 1,2,\ldots$, there are $f_k \in \mathcal{C}(F)$ and $g_k \in \mathcal{C}(G)$ such that $l(g_k)$ is uniformly positive definite and $|\upsilon'_t(t,x) - (\mathbb{L}_{f_k g_k}\upsilon(t,\cdot))(t,x)| \le 1/k$ for $k = 1,2,\ldots$ and $(t,x) \in \mathbb{R}^+ \times \mathbb{R}^d$.* □

Theorem 2.4. *Let $F : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times d})$ be l.s.c., bounded, convex-valued, and diagonally convex, respectively set-valued, mappings. Assume that G is such that for every $(t,x,u_x^t) \in \mathrm{Graph}(G)$, u_x^t is a right triangular matrix such that for every continuous selector φ of G, there is a number $L_\varphi > 0$ such that $\boldsymbol{l}(u_x^t)(u,u) > L_\varphi\|u\|^2$ for every $(t,x,u_x^t) \in \mathrm{Graph}(G)$ and $u \in \mathbb{R}^d$ with $\|u\| \neq 0$. Let $\upsilon \in C^{1,2}(\mathbb{R}^+ \times \mathbb{R}^d,\mathbb{R})$ be such that $\upsilon'_t(t,x) \in (\mathbb{L}_{FG}\upsilon(t,\cdot))(t,x)$ for every $t \in \mathbb{R}^+$ and $x \in \mathbb{R}^d$. For every $k = 1,2,\ldots$, there are $f_k \in \mathcal{C}(F)$ and $g_k \in \mathcal{C}(G)$ such that $l(g_k)$ is uniformly positive definite and $|\upsilon'_t(t,x) - (\mathbb{L}_{f_k g_k}\upsilon(t,\cdot))(t,x)| \le 1/k$ for $k = 1,2,\ldots$ and $(t,x) \in \mathbb{R}^+ \times \mathbb{R}^d$.* □

Theorem 2.5. *Let* $F : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ *and* $G : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times d})$ *be l.s.c., bounded, convex-valued, and diagonally convex, respectively set-valued, mappings satisfying (iv) of conditions* $(\mathcal{A})$. *If* $\upsilon \in C^{1,2}(\mathbb{R}^+ \times \mathbb{R}^d, \mathbb{R})$ *is such that* $\upsilon_t'(t,x) \in (\mathbb{L}_{FG}\upsilon(t,\cdot))(t,x)$ *for every* $t \in \mathbb{R}^+$ *and* $x \in \mathbb{R}^d$, *then for every* $k = 1,2,\ldots$, *there are continuous selectors* f_k *and* g_k *of* F *and* G, *respectively, such that* $l(g_k)$ *is uniformly positive definite and* $|\upsilon_t'(t,x) - (\mathbb{L}_{f_k g_k}\upsilon(t,\cdot))(t,x)| \le \frac{1}{k}$ *for* $k = 1,2,\ldots$ *and* $(t,x) \in \mathbb{R}^+ \times \mathbb{R}^d$. □

In some optimal control problems, we have to deal with condition (iv′) of $(\mathcal{A})$ instead of (iv). This follows from the following example.

Example 2.1. Let $g : \mathbb{R}^+ \times \mathbb{R}^d \times \mathcal{U} \to \mathbb{R}^{d\times d}$ be continuous and bounded, and put $G(t,x) = \{g(t,x,u) : u \in \mathcal{U}\}$ for every $(t,x) \in \mathbb{R}^+ \times \mathbb{R}^d$, where $\mathcal{U}$ is a nonempty compact convex subset of $\mathbb{R}^m$. Assume that g is such that for every $(t,x) \in \mathbb{R}^+ \times \mathbb{R}^d$, the mapping $\lambda(t,x,\cdot) : \mathcal{U} \to \mathbb{R}^{d\times d}$ defined by $\lambda(t,x,u) = g(t,x,u)\cdot g^*(t,x,u)$ for $u \in \mathcal{U}$ is affine. Then G does not satisfy all conditions of (iv) mentioned above. It is l.s.c. and diagonally convex, and for every continuous selector $\sigma : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$ of $D(G)$, we have $\sigma(t,x) \in \lambda(t,x,\mathcal{U})$ for every $(t,x) \in \mathbb{R}^+ \times \mathbb{R}^d$. In the general case, no selector $\gamma : \mathbb{R}^+ \times \mathbb{R}^d \to \mathcal{U}$ of $\mathcal{U}$ exists such that $\sigma = \lambda(t,x,\gamma)$. Therefore, to such a function λ we can apply Theorem 2.2 of Chap. 2 and get for every $k \ge 1$ a continuous selector $u_k : \mathbb{R}^+ \times \mathbb{R}^d \to \mathcal{U}$ of $\mathcal{U}$ such that $\sup_{(t,x)\in\mathbb{R}^+\times\mathbb{R}^d} \|\sigma(t,x) - \lambda(t,x,u_k(t,x))\| \le 1/k$ for $(t,x) \in \mathbb{R}^+ \times \mathbb{R}^d$. Taking $\gamma_k(t,x) = g(t,x,u_k(t,x))$ for $k \ge 1$, we obtain a sequence $(\gamma_k)_{k=1}^\infty$ of selectors of G such that $\sup_{(t,x)\in\mathbb{R}^+\times\mathbb{R}^d} \|\sigma(t,x) - l(\gamma_k)(t,x)\| \to 0$ as $k \to \infty$, because $l(\gamma_k)(t,x) = g(t,x,u_k(t,x))\cdot g^*(t,x,u_k(t,x)) = \lambda(t,x,u_k(t,x))$.

We now extend Theorem 2.5 to G satisfying (iv′) of conditions $(\mathcal{A})$.

Theorem 2.6. *Let* $F : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ *and* $G : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times d})$ *be l.s.c., bounded, convex-valued, and diagonally convex, respectively set-valued, mappings and suppose G is such that condition (iv′) of* $(\mathcal{A})$ *is satisfied. If* $\upsilon \in C_0^{1,2}(\mathbb{R}^+ \times \mathbb{R}^d, \mathbb{R})$ *is such that* $\upsilon_t'(t,x) \in (\mathbb{L}_{FG}\upsilon(t,\cdot))(t,x)$ *for every* $t \in \mathbb{R}^+$ *and* $x \in \mathbb{R}^d$, *then for every* $k = 1,2,\ldots$ *and* $m = 1,2,\ldots$, *there are continuous selectors* f_k *and* g_k^m *of* F *and* G, *respectively, such that* $l(g_k^m)$ *is uniformly positive definite for* $k,m = 1,2,\ldots$ *and* $\lim_{m\to\infty}\sup_{(t,x)\in\mathbb{R}^+\times\mathbb{R}^d} |\upsilon_t'(t,x) - (\mathbb{L}_{f_k g_k^m}\upsilon(t,\cdot))(t,x)| \le 1/k$ *for* $k = 1,2,\ldots$.

Proof. Similarly as above, by virtue of Theorem 2.2 of Chap. 2, for every $k = 1,2,\ldots$, there are $f_k \in \mathcal{C}(F)$ and $\sigma_k \in \mathcal{C}(D(G))$ such that $\sup_{(t,x)\in\mathbb{R}^+\times\mathbb{R}^d} |\upsilon_t'(t,x) - \lambda[(t,x),(f_k,\sigma_k(t,x))]| \le 1/k$, where λ is as in the proof of Theorem 2.1. By (iv′), for every $k \ge 1$ there is a sequence $(g_k^m)_{m=1}^\infty$ of $\mathcal{C}(G)$ such that $\sup_{(t,x)\in\mathbb{R}^+\times\mathbb{R}^d} |l(g_k^m(t,x)) - \sigma_k(t,x)| \to 0$ as $m \to \infty$. For every $k \ge 1$ and $m \ge 1$, one has

$$\sup_{(t,x)\in\mathbb{R}^+\times\mathbb{R}^d} |\lambda[(t,x),(f_k,\sigma_k)(t,x)] - \lambda[(t,x),(f_k,\sigma_k^m)(t,x)]|$$

$$\le M\cdot d^2 \sup_{(t,x)\in\mathbb{R}^+\times\mathbb{R}^d} \|\sigma_k(t,x) - \sigma_k^m(t,x)\|,$$

where $\sigma_k^m(t,x) = l(g_k^m(t,x))$ and $M > 0$ is such that $M \geq \sup_{(t,x)\in\mathbb{R}^+\times\mathbb{R}^d} |\upsilon''_{x_i x_j}(t,x)|$ for $i, j = 1, 2, \ldots, d$. Let us observe that such $M > 0$ exists, because υ has compact support. Therefore, for every $k \geq 1$ and $m \geq 1$, we get

$$\sup_{(t,x)\in\mathbb{R}^+\times\mathbb{R}^d} |\upsilon'_t(t,x) - (\mathbb{L}_{f_k g_k^m}\upsilon(t,\cdot))(t,x)| \leq \sup_{(t,x)\in\mathbb{R}^+\times\mathbb{R}^d} |\upsilon'_t(t,x) - \lambda[(t,x),(f_k,\sigma_k)(t,x)]|$$

$$+ \sup_{(t,x)\in\mathbb{R}^+\times\mathbb{R}^d} |\lambda[(t,x),(f_k,\sigma_k)(t,x)] - \lambda[(t,x),(f_k,\sigma_k^m)(t,x)]|$$

$$\leq \frac{1}{k} + M\cdot d^2 \sup_{(t,x)\in\mathbb{R}^+\times\mathbb{R}^d} \|\sigma_k(t,x) - \sigma_k^m(t,x)\|.$$

Then

$$\lim_{m\to\infty} \sup_{(t,x)\in\mathbb{R}^+\times\mathbb{R}^d} |\upsilon'_t(t,x) - (\mathbb{L}_{f_k g_k^m}\upsilon(t,\cdot))(t,x)| \leq \frac{1}{k}$$

for every $k = 1, 2, \ldots$. □

3 Initial and Boundary Value Problems for Semielliptic Partial Differential Equations

Let $F : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ generate a uniformly semielliptic set-valued operator $\mathbb{L}_{FG}$ on $Q \subset \mathbb{R}^+ \times \mathbb{R}^d$ and let $f \in \mathcal{C}(F)$ and $g \in \mathcal{C}(G)$. Assume that Q is a bounded domain in $\mathbb{R}^+\times\mathbb{R}^d$ lying in the strip $(0,T)\times\mathbb{R}^d$ for a given $T > 0$. Put $\bar{B} = \bar{Q} \cap [\{t = 0\} \times \mathbb{R}^d]$ and $\bar{B}_T = \bar{Q} \cap [\{t = T\} \times \mathbb{R}^d]$ and assume that $\bar{B}$ and $\bar{B}_T$ are nonempty. Let $B_T = \mathrm{Int}(\bar{B}_T)$ and $B = \mathrm{Int}(\bar{B})$. Denote by S_0 the boundary of Q lying in the strip $(0,T)\times\mathbb{R}^d$ and let $S = S_0 \setminus B_T$. The set $\partial Q = B \cup S$ is a parabolic boundary of Q. Let $c : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}$, $\varphi : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}$, $\psi : \mathbb{R}^d \to \mathbb{R}$, and $\gamma : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}$ be given.

Given functions $f, g, c, \varphi, \psi, \gamma$ as above and a bounded domain $Q \subset (0,T) \times \mathbb{R}^d$, by the first initial–boundary value problem generated by $\mathbb{L}_{fg}$, we mean the problem consisting in finding a function $u \in C(\mathbb{R}^+\times\mathbb{R}^d, \mathbb{R})$ such that the following conditions are satisfied:

$$\begin{cases} u'_t(t,x) = \left(\mathbb{L}_{fg}u(t,\cdot)\right)(t,x) + c(t,x)u(t,x) - \varphi(t,x), \\ \text{for } (t,x) \in Q \cup B_T, \\ u(0,x) = \psi(x) \text{ for } x \in B, \\ u(t,s) = \gamma(t,x) \text{ for } (t,x) \in S. \end{cases} \tag{3.1}$$

Remark 3.1. The last two conditions in (3.1) are called initial and boundary conditions, respectively. □

A function $w_{\mathcal{R}}(P) \in C_0^2(\overline{Q})$ with $\mathcal{R} \in \bar{B} \cup S$ is called a barrier at the point $\mathcal{R}$ corresponding to (f, g, c) if $w_{\mathcal{R}}(P) > 0$ for $P \in \bar{Q} \setminus \{\mathcal{R}\}$, $w_{\mathcal{R}}(\mathcal{R}) = 0$ and $\left(\mathbb{L}_{fg} w_{\mathcal{R}}(t, \cdot)\right)(t, x) + c(t, x) w_{\mathcal{R}}(t, x) + 1 \leq \frac{\partial}{\partial t}[w_{\mathcal{R}}(t, x)]$ for $(t, x) \in Q \cup B_T$.

The following existence theorem can be proved.

Theorem 3.1. *Assume that a pair (F, G) of multifunctions generates a uniformly semielliptic set-valued operator $\mathbb{L}_{FG}$ on a bounded domain $Q \subset (0, T) \times \mathbb{R}^d$ and let $f \in \mathcal{C}(F)$ and $g \in \mathcal{C}(G)$. Assume that f, g, c and φ are uniformly Hölder continuous in $\bar{Q}$ and let ψ and γ be continuous on $\bar{B}$ and $\bar{S}$, respectively. If furthermore, there exists a barrier corresponding to (f, g, c) at every point of S, then there exists a unique solution of the first initial–boundary problem (3.1).* □

Remark 3.2. It can be verified that if $Q = (0, T) \times B$ with a bounded domain B in $\mathbb{R}^d$ is such that there exists a closed ball $K \subset \mathbb{R}^d$ with center $\bar{x}$ such that $K \cap B = \emptyset$ and $K \cap \bar{B} = \{x_0\}$, then there exists a barrier corresponding to (f, g, c) at each point (t_0, x_0) of S $(0 < t_0 \leq T)$, namely $w_{\mathcal{R}}(t, x) = k e^{\alpha t}(1/\mathcal{R}_0^p - 1/\mathcal{R}^p)$, where $\alpha \geq c(t, x)$, $\mathcal{R}_0 = |x_0 - \bar{x}|$, $\mathcal{R} = [|x - \bar{x}|^2 + |t - t_0|^2]^{1/2}$, and k, p are suitable positive numbers. □

Corollary 3.1. *If $Q = (0, T) \times B$ is as in Remark 3.2, then for every f, g, c, φ, and γ satisfying the assumptions of Theorem 3.1, there exists a unique solution of (3.1).* □

Given functions f, g, c, φ, ψ, γ, and $T > 0$ as above, by the Cauchy problem generated by $\mathbb{L}_{fg}$ we mean the problem consisting in finding a function $u \in C(\mathbb{R}^+ \times \mathbb{R}^d, \mathbb{R})$ such that the following conditions are satisfied:

$$\begin{cases} u_t'(t, x) = \left(\mathbb{L}_{fg} u(t, \cdot)\right)(t, x) + c(t, x) u(t, x) - \varphi(t, x), \\ \text{for } \ (t, x) \in (0, T] \times \mathbb{R}^d, \\ u(0, x) = \psi(x) \ \text{ for } \ x \in \mathbb{R}^d. \end{cases} \tag{3.2}$$

The following existence theorem can be proved.

Theorem 3.2. *Assume that a pair (F, G) of multifunctions generates a uniformly semielliptic set-valued operator $\mathbb{L}_{FG}$ on $[0, T] \times \mathbb{R}^d$ and let $f \in \mathcal{C}(F)$ and $g \in \mathcal{C}(G)$ be bounded. Furthermore, assume that f and g are Hölder continuous in x uniformly with respect to $(t, x) \in [0, T] \times \mathbb{R}^d$. Let φ be continuous and Hölder continuous in x uniformly with respect to (t, x) on a compact subset of $[0, T] \times \mathbb{R}^d$ and let ψ be continuous on $\mathbb{R}^d$. If furthermore, there are positive numbers α and A such that $\max(|\varphi(t, x)|, |\psi(x)|) \leq A \cdot \exp\left(\alpha |x|^2\right)$ for $(t, x) \in [0, T] \times \mathbb{R}^d$, then there is $\bar{c} > 0$ such that the Cauchy problem (3.2) possesses a solution in the strip $[0, T^*] \times \mathbb{R}^d$, where $T^* = \min\{T, \bar{c}/\alpha\}$.* □

4 Stochastic Representation of Solutions of Partial Differential Equations

If the assumptions of Theorems 3.1 and 3.2 are satisfied, then solutions of problems (3.1) and (3.2) can in some special cases be represented by solutions of stochastic differential equations $SDE(f,g)$. We shall consider the above problem in the case that f and g satisfy only assumptions that guarantee the existence and uniqueness in law of weak solutions of $SDE(f,g)$ with a given initial distribution. We shall still assume that a pair (F,G) of multifunctions generates a uniformly semielliptic set-valued operator $\mathbb{L}_{FG}$ on $[0,T]\times\mathbb{R}^d$ and $(f,g)\in\mathcal{C}(F)\times\mathcal{C}(G)$.

By virtue of Corollary 1.3, for every $(s,x)\in\mathbb{R}^+\times\mathbb{R}^d$, there is a $(d+1)$-dimensional Itô diffusion $Y^{fg}_{s,x}=(Y^{fg}_{s,x}(t))_{t\geq 0}$ defined by $Y^{fg}_{s,x}(t)=(s+t,X^{fg}_{s,x}(s+t))$ on a filtered probability space $\mathcal{P}_{\mathbb{F}}=(\Omega,\mathcal{F},\mathbb{F},P)$. Its infinitesimal generator $\mathcal{A}_{fg}$ satisfies $(\mathcal{A}_{fg}h)(t,x)-h'_t(t,x)\in(\mathbb{L}_{FG}h)(t,x)$ for $h\in C_0^{1,2}(\mathbb{R}^{d+1})$ and $(t,x)\in[s,\infty)\times\mathbb{R}^d$, where $X^{fg}_{s,x}=(X^{fg}_{s,x}(t))_{t\geq s}$ is a weak solution of a stochastic differential equation $SDE(f,g)$ with initial condition $X^{fg}_{s,x}(s)=x$ a.s.

Immediately from results presented in Sect. 12 of Chap. 1, we obtain the following existence and representation theorems.

Theorem 4.1. *Assume that conditions* (i)–(iv) *of* $(\mathcal{A})$ *are satisfied,* $T>0$*, and let* $c\in C([0,T]\times\mathbb{R}^d,\mathbb{R})$ *be bounded. For every* $(f,g)\in\mathcal{C}(F)\times\mathcal{C}(G)$ *and* $(s,x)\in[0,T)\times\mathbb{R}^d$*, there is a unique in law weak solution* $X^{fg}_{s,x}$ *of* $SDE(f,g)$ *satisfying* $X^{fg}_{s,x}(s)=x$ *a.s. such that the function* v *defined by*

$$v(t,s,x)=E^{s,x}\left[\exp\left(-\int_0^t c(Y^{fg}_{s,x}(\tau))\mathrm{d}\tau\right)\tilde{h}(Y^{fg}_{s,x}(t))\right]$$

for $\tilde{h}\in C^{1,2}(\mathbb{R}^{d+1})$*,* $(s,x)\in[0,T)\times\mathbb{R}^d$*, and* $t\in[0,T-s]$ *satisfies*

$$\begin{cases} v'_t(t,s,x)=\left(\mathcal{A}_{fg}v(t,\cdot)\right)(s,x)-c(s,x)v(t,s,x) \\ \text{for } \ (s,x)\in[0,T)\times\mathbb{R}^d \ \text{ and } t\in[0,T-s], \\ v(0,s,x)=\tilde{h}(s,x) \ \text{ for } \ (s,x)\in[0,T)\times\mathbb{R}^d. \end{cases} \qquad \square$$

Theorem 4.2. *Assume that conditions* (i)–(iv) *of* $(\mathcal{A})$ *are satisfied,* $T>0$*, and* $c\in C([0,T]\times\mathbb{R}^d,\mathbb{R})$ *is bounded. For every* $(f,g)\in\mathcal{C}(F)\times\mathcal{C}(G)$ *and* $x\in\mathbb{R}^d$*, there is a unique in law weak solution* $X^{fg}_{0,x}$ *of* $SDE(f,g)$ *satisfying* $X^{fg}_{0,x}(0)=x$ *a.s., such that the function* v *defined by*

$$v(t,x)=E^{x}\left[\exp\left(-\int_0^t c(Y^{fg}_{0,x}(\tau))\mathrm{d}\tau\right)(h\circ\pi)(Y^{fg}_{0,x}(t))\right]$$

for $h\in C_0^2(\mathbb{R}^d)$ *and* $(t,x)\in[0,T]\times\mathbb{R}^d$ *satisfies*

$$\begin{cases} v_t'(t,x) = \left(\mathcal{A}_{fg}v(t,\cdot)\right)(t,x) - c(t,x)v(t,x) \\ \text{for } \ (t,x) \in (0,T] \times \mathbb{R}^d, \\ v(0,x) = h(x) \ \text{ for } \ x \in \mathbb{R}^d. \end{cases}$$

□

Theorem 4.3. *Assume that conditions* (i)–(iv) *of* $(\mathcal{A})$ *are satisfied,* $T > 0$, $c \in C([0,T] \times \mathbb{R}^d$ *and* $v \in C^{1,1,2}([0,T] \times [0,T] \times \mathbb{R}^d, \mathbb{R})$ *is bounded and such that*

$$\begin{cases} v_t'(t,s,x) = \left(\mathcal{A}_{fg}v(t,\cdot)\right)(s,x) - c(s,x)v(t,s,x) \\ \text{for } \ (s,x) \in [0,T) \times \mathbb{R}^d \ \text{ and } \ t \in [0,T-s] \\ v(0,s,x) = \tilde{h}(s,x) \ \text{ for } \ (s,x) \in [0,T) \times \mathbb{R}^d. \end{cases}$$

for $(f,g) \in \mathcal{C}(F) \times \mathcal{C}(G)$ *and* $\tilde{h} \in C_0^{1,2}(\mathbb{R}^{d+1})$. *For every* $(s,x) \in [0,T) \times \mathbb{R}^d$, *there is a unique in law weak solution* $X_{s,x}^{fg}$ *of* $SDE(f,g)$ *satisfying initial condition* $X_{s,x}^{fg}(s) = x$ *such that*

$$v(t,s,x) = E^{s,x}\left[\exp\left(-\int_s^{s+t} c(\tau, X_{s,x}^{fg}(\tau))\mathrm{d}\tau\right)\tilde{h}(s+t, X_{s,x}^{fg}(s+t))\right]$$

for $(s,x) \in [0,T) \times \mathbb{R}^d$ *and* $t \in [0,T-s]$. □

Remark 4.1. If the assumptions of Theorem 4.1 are satisfied, $w \in C^{1,2}([0,T] \times \mathbb{R}^d, \mathbb{R})$ and $u \in C([0,T] \times \mathbb{R}^d, \mathbb{R})$ are bounded and such that the function v defined by $v(s,x) = w(T-s,x)$ for $0 \le s < T$ satisfies

$$\begin{cases} v_s'(s,x) + (\mathcal{A}_{fg}v)(s,x) = -u(s,x) \\ \text{for } \ (s,x) \in [0,T) \times \mathbb{R}^d \\ v(0,x) = \tilde{h}(T,x) \ \text{ for } \ x \in \mathbb{R}^d, \end{cases}$$

then for every $(s,x) \in [0,T) \times \mathbb{R}^d$, there exists a unique in law weak solution $X_{s,x}^{fg}$ of $SDE(f,g)$ with an initial condition $X_{s,x}^{fg}(s) = x$ a.s. such that

$$w(T-s,x) = E^{s,x}\left[\tilde{h}(T, X_{s,x}^{fg}(T))\right] + E^{s,x}\left[\int_s^T u(\tau, X_{s,x}^{fg}(\tau))\mathrm{d}\tau\right].$$

□

We shall consider now some generalized Dirichlet–Poisson problems with partial differential operators generated by $(f,g) \in \mathcal{C}(F) \times \mathcal{C}(G)$ with F and G satisfying conditions (i)–(iv) of $(\mathcal{A})$. Similarly as in Chap. 1, we obtain the following results.

Theorem 4.4. *Assume that conditions* (i)–(iv) *of* $(\mathcal{A})$ *are satisfied,* $T > 0$, *and* D *is a bounded domain in* $\mathbb{R}^d$. *Let* $u \in C([0,T] \times \mathbb{R}^d, \mathbb{R})$, $\Phi \in C((0,T) \times \partial D, \mathbb{R})$ *be bounded and* $(f,g) \in \mathcal{C}(F) \times \mathcal{C}(G)$. *If* $v \in C_0^{1,2}(\mathbb{R}^{d+1})$ *is bounded such that*

$$\begin{cases} u(t,x) = \left(\mathcal{A}_{fg}v\right)(t,x) \ \text{ for } \ (t,x) \in [0,T) \times D, \\ \lim_{D \ni x \to y} v(t,x) = \Phi(t,y) \ \text{ for } \ (t,y) \in (0,T) \times \partial D, \end{cases} \tag{4.1}$$

then for every $(s,x) \in [0,T) \times D$*, there exists a unique in law weak solution* $X_{s,x}^{fg}$ *of* $SDE(f,g)$ *with initial condition* $X_{s,x}^{fg}(s) = x$ *such that*

$$v(s,x) = E^{s,x}\left[\Phi(\tau_D, X_{s,x}^{fg}(\tau_D))\right] - E^{s,x}\left[\int_s^{\tau_D} u(t, X_{s,x}^{fg}(t))\mathrm{d}t\right]$$

for $(s,x) \in [0,T) \times D$*, where* $\tau_D = \inf\{r \in (s,T] : X_{s,x}^{fg}(r) \notin D\}$.

Proof. By virtue of Remark 11.1 of Chap. 1, for every $(s,x) \in [0,T) \times D$, there is a unique in law weak solution $X_{s,x}^{fg}$ of $SDE(f,g)$ with initial condition $X_{s,x}^{fg}(s) = x$ a.s. such that the process $Y_{s,x}^{fg} = (s+t, X_{s,x}^{fg}(s+t))_{0 \le t \le T-s}$ is an Itô diffusion with the infinitesimal generator $\mathcal{A}_{fg}$ defined above. Let $\mathcal{U}_k = (0,T) \times D_k$ and $\tau_{\mathcal{U}_k}^s = \inf\{r \in (0, T-s] : Y_{s,x}^{fg}(t) \notin \mathcal{U}_k\}$, where $(D_k)_{k=1}^\infty$ is an increasing sequence of open sets D_k such that $D_k \subset D$ and $D = \bigcup_{k=1}^\infty D_k$. It can be verified that $\tau_{\mathcal{U}_k}^s = \tau_k^s - s$, where $\tau_k^s = \inf\{r \in (s,T] : X_{s,x}^{fg}(r) \notin D_k\}$. By Dynkin's formula, for every $k = 1, 2, \ldots$ we get

$$E^{s,x}\left[v(Y_{s,x}^{fg}(\tau_{\mathcal{U}_k}^s))\right] = v(s,x) + E^{s,x}\left[\int_0^{\tau_{\mathcal{U}_k}^s} (\mathcal{A}_{fg}v)(Y_{s,x}^{fg}(t))\mathrm{d}t\right]$$

for every $(s,x) \in [0,T) \times D$. By (4.1), we have $u(s+t, X_{s,x}^{fg}(s+t)) = (\mathcal{A}_{fg}v)(Y_{s,x}^{fg}(t))$ for $(s,x) \in [0,T] \times D$ and $t \in [0, T-s]$. Hence, by the definition of $Y_{s,x}^{fg}$ and the equality $\tau_{\mathcal{U}_k}^s = \tau_k^s - s$, for every $k = 1, 2, \ldots$, we obtain

$$v(s,x) = E^{s,x}\left[v(\tau_k^s, X_{s,x}^{fg}(\tau_k^s))\right] - E^{s,x}\left[\int_s^{\tau_k^s} u(t, X_{s,x}^{fg}(t))\mathrm{d}t\right].$$

On the other hand, by (4.1) and the boundedness of the functions Φ and u, we get

$$\lim_{k\to\infty} E^{s,x}[v(\tau_k^s, X_{s,x}^{fg}(\tau_k^s)] = E^{s,x}[\Phi(\tau_D, X_{s,x}^{fg}(\tau_D))]$$

and

$$\lim_{k\to\infty} E^{s,x}\left[\int_s^{\tau_k^s} u(t, X_{s,x}^{fg}(t))\mathrm{d}t\right] = E^{s,x}\left[\int_s^{\tau_D} u(t, X_{s,x}^{fg}(t))\mathrm{d}t\right]$$

for $(s, x) \in [0, T) \times D$. Then

$$v(s, x) = E^{s,x}\left[\Phi(\tau_D, X^{fg}_{s,x}(\tau_D))\right] - E^{s,x}\left[\int_s^{\tau_D} u(t, X^{fg}_{s,x}(t))\mathrm{d}t\right]$$

for $(s, x) \in [0, T) \times D$. □

Theorem 4.5. *Assume that conditions* (i)–(iv) *of* $(\mathcal{A})$ *are satisfied,* $T > 0$*, and* D *is a bounded domain in* $\mathbb{R}^d$*. Let* $c, u \in C([0, T] \times \mathbb{R}^d, \mathbb{R})$ *and* $\Phi \in C((0, T) \times \partial D, \mathbb{R})$ *be bounded and* $(f, g) \in \mathcal{C}(F) \times \mathcal{C}(G)$*. If* $v \in C_0^{1,2}(\mathbb{R}^{d+1})$ *is bounded such that*

$$\begin{cases} u(t, x) = \left(\mathcal{A}_{fg} v\right)(t, x) - c(t, x)v(t, x) \ \text{ for } \ (t, x) \in (0, T], \times D \\ \lim_{D \ni x \to y} v(t, x) = \Phi(t, y) \ \text{ for } \ (t, y) \in (0, T) \times \partial D, \end{cases}$$

then for every $(s, x) \in [0, T) \times D$*, there exists a unique in law weak solution* $X^{fg}_{s,x}$ *of* $SDE(f, g)$ *with initial condition* $X^{fg}_{s,x}(s) = x$ *a.s such that*

$$\begin{aligned} v(s, x) = {} & E^{s,x}\left[\Phi(\tau_D, X^{fg}_{s,x}(\tau_D)) \exp\left(-\int_s^{\tau_D} c(t, X^{fg}_{s,x}(t))\mathrm{d}t\right)\right] \\ & - E^{s,x}\left\{\int_s^{\tau_D}\left[u(r, X^{fg}_{s,x}(r)) \exp\left(-\int_s^{s+r} c(r, X^{fg}_{s,x}(t))\mathrm{d}t\right)\right]\mathrm{d}r\right\} \end{aligned}$$

for $(s, x) \in [0, T) \times D$*, where* $\tau_D = \inf\{r \in (s, T] : X^{fg}_{s,x}(r) \notin D\}$.

Proof. Similarly as in the proof of Theorem 4.4, we can verify that for every $(f, g) \in \mathcal{C}(F) \times \mathcal{C}(G)$ and $(s, x) \in [0, T) \times D$, there is a unique in law weak solution $X^{fg}_{s,x}$ of the stochastic differential equation $SDE(f, g)$ satisfying the initial condition $X^{fg}_{s,x}(s) = x$ a.s. such that the process $Y^{fg}_{s,x} = ((s + t, X^{fg}_{s,x}(s + t)))_{0 \le t \le T-s}$ is an Itô diffusion with the infinitesimal generator $\mathcal{A}_{fg}$ defined above. Let $\mathcal{U} = (0, T) \times D$ and $\tau^s_{\mathcal{U}} = \inf\{r \in (0, T - s] : Y^{fg}_{s,x}(t) \notin \mathcal{U}\}$. It can be verified that $\tau^s_{\mathcal{U}} = \tau_D - s$. Fix $(s, x, z) \in [0, T] \times \mathbb{R} \times \mathbb{R}^d$ and define $Z^s_t = z + \int_0^t c(Y^{fg}_{s,x}(\tau))\mathrm{d}\tau$ and $H^s_t = (Y^{fg}_{s,x}(t), Z^s_t)$. It can be verified that $(H^s_t)_{0 \le t \le T-s}$ is an Itô diffusion with the infinitesimal generator $(\mathcal{A}_H \psi)(s, x, z) = (\mathcal{A}_{fg}\psi)(s, x, z) + \psi'_z(s, x, z)c(s, x)$ for $\psi \in C_0^{1,2}([0, T] \times \mathbb{R}^{n+1})$. Hence by Dynkin's formula, it follows that

$$E^{s,x,z}\left[\psi(H^s_{\tau^s_{\mathcal{U}} \wedge \tau_R})\right] = \psi(s, x, z) + E^{s,x,z}\left[\int_0^{\tau^s_{\mathcal{U}} \wedge \tau_R} (\mathbb{L}_H \psi)(H^s_r)\mathrm{d}r\right],$$

where $\tau_R = \inf\{t \in (0, T - s] : |H^s_t| \ge R\}$. Taking $\psi(s, x, z) = e^{-z}v(s, x)$, we get

$$E^{s,x,z}\left[\psi(H^s_{\tau^s_{\mathcal{U}} \wedge \tau_R})\right] = E^{s,x,z}\left[\exp\left(-\int_0^{\tau^s_{\mathcal{U}} \wedge \tau_R} c(Y^{fg}_{s,x}(r))\mathrm{d}r\right) v(Y^{fg}_{s,x}(\tau^s_{\mathcal{U}} \wedge \tau_R))\right]$$

and

$$(\mathcal{A}_H\psi)(H_r^s) = \exp\left(-\int_0^r c(Y_{s,x}^{fg}(\tau))\mathrm{d}\tau\right)\left[(\mathcal{A}_{fg}v)(Y_{s,x}^{fg}(r)) - c(Y_{s,x}^{fg}(r))v(Y_{s,x}^{fg}(r))\right].$$

From this and (4.5), it follows that

$$\begin{aligned} e^{-z}v(s,x) = E^{s,x,z}&\left[\exp\left(-\int_0^{\tau_{\mathcal{U}}^s\wedge\tau_R} c(Y_{s,x}^{fg}(r))\mathrm{d}r\right)v(Y_{s,x}^{fg}(\tau_{\mathcal{U}}^s\wedge\tau_R))\right]\\ &- E^{s,x,z}\left[\int_0^{\tau_{\mathcal{U}}^s\wedge\tau_R}\exp\left(-\int_0^r c(Y_{s,x}^{fg}(\tau))\mathrm{d}\tau\right)u(Y_{s,x}^{fg}(r))\mathrm{d}r\right]. \end{aligned}$$

Taking $z = 0$ and passing to the limit $R \to \infty$, one obtains

$$\begin{aligned} v(s,x) = E^{s,x}&\left[\exp\left(-\int_0^{\tau_{\mathcal{U}}^s} c(Y_{s,x}^{fg}(\tau))\mathrm{d}\tau\right)\Phi(Y_{s,x}^{fg}(\tau_{\mathcal{U}}^s))\right]\\ &- E^{s,x}\left[\int_0^{\tau_{\mathcal{U}}^s}\exp\left(-\int_0^r c(Y_{s,x}^{fg}(\tau))\mathrm{d}\tau\right)u(Y_{s,x}^{fg}(r))\mathrm{d}r\right], \end{aligned}$$

because $\lim_{R\to\infty} v(Y_{s,x}^{fg}(\tau_{\mathcal{U}}^s\wedge\tau_R)) = \Phi(Y_{s,x}^{fg}(\tau_{\mathcal{U}}^s))$. From this and the equality $Y_{s,x}^{fg}(t) = (s+t, X_{s,x}^{fg}(s+t))$, it follows that

$$\begin{aligned} v(s,x) = E^{s,x}&\left[\exp\left(-\int_s^{s+\tau_{\mathcal{U}}^s} c(,\tau,X_{s,x}^{fg}(\tau))\mathrm{d}\tau\right)\Phi(s+\tau_{\mathcal{U}}^s, X_{s,x}^{fg}(s+\tau_{\mathcal{U}}^s))\right]\\ &- E^{s,x}\left[\int_s^{s+\tau_{\mathcal{U}}^s}\exp\left(-\int_s^{s+r} c(\tau,X_{s,x}^{fg}(\tau))\mathrm{d}\tau\right)u(r,X_{s,x}^{fg}(r))\mathrm{d}r\right]. \end{aligned}$$

Therefore,

$$\begin{aligned} v(s,x) = E^{s,x}&\left[\Phi(\tau_D, X_{s,x}^{fg}(\tau_D))\exp\left(-\int_s^{\tau_D} c(t,X_{s,x}^{fg}(t))\mathrm{d}t\right)\right]\\ &- E^{s,x}\left\{\int_s^{\tau_D}\left[u(r,X_{s,x}^{fg}(r))\exp\left(-\int_s^{s+r} c(r,X_{s,x}^{fg}(t))\mathrm{d}t\right)\right]\mathrm{d}r\right\} \end{aligned}$$

for $(s,x) \in [0,T)\times D$, because $s+\tau_{\mathcal{U}}^s = \tau_D$. □

Corollary 4.1. *If the assumptions of Theorem 4.5 are satisfied and* $v \in C_0^{1,2}(\mathbb{R}^{d+1})$ *is bounded such that*

$$\begin{cases} -v_t'(t,x) = \left(\mathcal{A}_{fg}v\right)(t,x) - c(t,x)v(t,x) & \text{for } (t,x) \in [0,T) \times D, \\ \lim_{D \ni x \to y} v(t,x) = \Phi(t,y) & \text{for } (t,y) \in (0,T) \times \partial D, \end{cases}$$

then for every $(s,x) \in [0,T) \times D$*, there is a unique in law weak solution* $X_{s,x}^{fg}$ *of* $SDE(f,g)$ *with initial condition* $X_{s,x}^{fg}(s) = x$ *a.s. such that*

$$v(s,x) = E^{s,x}\left[\Phi(\tau_D, X_{s,x}^{fg}(\tau_D)) \exp\left(-\int_s^{\tau_D} c(t, X_{s,x}^{fg}(t))\mathrm{d}t\right)\right]$$

for $(s,x) \in [0,T) \times D$. □

5 Existence of Solutions of the Stochastic Dirichlet–Poisson Problem

The question of the existence of a solution to the Dirichlet and Poisson problems is much more complicated then the boundary values problems presented above. For example, a natural candidate for a solution to the simple Dirichlet problem

$$\begin{cases} \left(\mathcal{A}_{fg}v\right)(t,x) = 0 & \text{for } (t,x) \in (0,T) \times D, \\ \lim_{D \ni x \to y} v(t,x) = \Phi(t,y) & \text{for } (t,y) \in (0,T] \times \partial D, \end{cases} \tag{5.1}$$

seems to be the function defined by $w(s,x) = E^{s,x}[\Phi(\tau_D, X_{s,x}^{fg}(\tau_D))]$, where similarly as above, $X_{s,x}^{fg}$ is a weak solution of $SDE(f,g)$ with initial condition $X_{s,x}^{fg}(s) = x$ a.s. Unfortunately, in the general case, such a function w need not be in $C^{1,2}((0,T) \times D, \mathbb{R})$. In fact, it need not even be continuous.

Example 5.1. Let $X(t) = (X_1(t), X_2(t))$ be a solution of the equation $\mathrm{d}X(t) = (1,0)\mathrm{d}t$, so that $X(t) = X(0) + (1,0)t \in \mathbb{R}^2$ for $t \geq 0$. Let $D = [(0,1) \times (0,1)] \cup [(0,2) \times (0,1/2)]$ and let Φ be a continuous function on ∂D such that $\Phi = 1$ on $\{1\} \times [1/2,1]$ and $\Phi = 0$ on $(\{2\} \times [0,1/2]) \cup (\{0\} \times [0,1])$ (see Fig. 6.1). Then $w(s,x) = E^{s,x}[\Phi(X_{\tau_D})] = 1$ if $x \in (1/2,1)$ and $w(s,x) = 0$ if $x \in (0,1/2)$. So w is not continuous. Moreover, $\lim_{t \to 0+} w(t,x) = 1 \neq \Phi(0,x)$ if $1/2 < x < 1$. Then the second condition of (5.1) does not hold.

However, the function w defined above will solve the Dirichlet problem in a weaker stochastic sense:

$$\begin{cases} \left(\mathcal{L}_{fg}v\right)(t,x) = 0 & \text{for } (t,x) \in [0,T) \times D, \\ \lim_{t \to \tau_D} v(Y_{s,x}^{fg}(t)) = \Phi(Y_{s,x}^{fg}(\tau_D)) & \text{for } (s,x) \in (0,T) \times D, \end{cases} \tag{5.2}$$

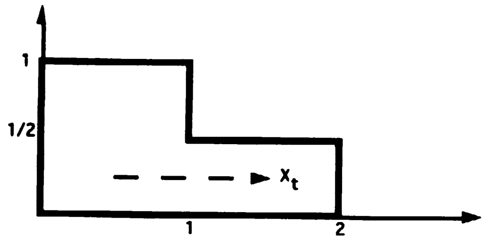

Fig. 6.1 Domain of the function Φ

where $\mathcal{L}_{fg}$ is the characteristic operator of the diffusion $Y_{s,x}^{fg}$ defined above. We shall consider now the problem of the existence of solutions of the stochastic Dirichlet–Poisson problem of the form (5.2). It will be convenient to divide it into two problems: stochastic Dirichlet and stochastic Poisson problems. Assume $(f,g) \in \mathcal{C}(F) \times \mathcal{C}(G)$ with F and G such that the pair (F,G) generates the uniformly parabolic diffusion set-valued operator $\mathbb{L}_{FG}$.

Let φ be a locally bounded measurable function on $(0,T) \times D$ with $T > 0$ and a bounded domain $D \subset \mathbb{R}^d$. A function φ is said to be $Y_{s,x}^{fg}$-harmonic in $\Lambda = (0,T) \times D$ if $\varphi(s,x) = E^{s,x}[\varphi(Y_{s,x}^{fg}(\tau_{\mathcal{U}}))]$ for all $(s,x) \in \Lambda$ and all open sets $\mathcal{U}$ such that $\mathrm{cl}\{\mathcal{U}\} \subset \Lambda$, where $\tau_{\mathcal{U}}$ is the first exit time of $Y_{s,x}^{fg}$ from the set $\mathcal{U}$. We get the following basic results.

Lemma 5.1. *If φ is $Y_{s,x}^{fg}$-harmonic in $\Lambda = (0,T) \times D$, then $\mathcal{L}_{fg}\varphi = 0$ in Λ. Conversely, if $\varphi \in C^{1,2}(\Lambda)$ is such that $\mathcal{L}_{fg}\varphi = 0$ in Λ, then φ is $Y_{s,x}^{fg}$-harmonic in Λ.*

Proof. By the definition of the characteristic operator $\mathcal{L}_{fg}$ of $Y_{s,x}^{fg}$, we have

$$(\mathcal{L}_{fg}\varphi)(s,x) = \lim_{\mathcal{U}\to(s,x)} \frac{E^{s,x}[\varphi(Y_{s,x}^{fg}(\tau_{\mathcal{U}}))] - \varphi(s,x)}{E^{s,x}[\tau_{\mathcal{U}}]}$$

for $(s,x) \in \Lambda$. If φ is $Y_{s,x}^{fg}$-harmonic in Λ, then $\mathcal{L}_{fg}\varphi = 0$ in Λ. Conversely, if $\varphi \in C^{1,2}(\Lambda)$ is such that $(\mathcal{L}_{fg}\varphi)(s,x) = 0$ for $(s,x) \in \Lambda$, then by Theorem 10.2 of Chap. 1 and Remark 11.1 of Chap. 1, we have $(\mathcal{L}_{fg}\varphi)(s,x) = \varphi_t'(s,x) + (\mathbb{L}_{fg}\varphi(t,\cdot))(s,x) = (\mathcal{A}_{fg}\varphi)(s,x)$, where $\mathcal{A}_{fg}$ is an infinitesimal generator of the Itô diffusion $Y_{s,x}^{fg}$. By Dynkin's formula, it follows that

$$E^{s,x}[\varphi(Y_{s,x}^{fg}(\tau_{\mathcal{U}}))] = \lim_{k\to\infty} E^{s,x}[\varphi(Y_{s,x}^{fg}(\tau_{\mathcal{U}} \wedge k))] = \varphi(s,x) + E^{s,x}\left[\int_0^{\tau_{\mathcal{U}}\wedge k} (\mathcal{A}_{fg}\varphi)(s,x)\mathrm{d}s\right]$$

for $(s,x) \in \Lambda$. Hence, by the equalities $(\mathcal{L}_{fg}\varphi)(s,x) = 0$ and $(\mathcal{L}_{fg}\varphi)(s,x) = (\mathcal{A}_{fg}\varphi)(s,x)$, for $(s,x) \in \Lambda$, we obtain $E^{s,x}[\varphi(Y_{s,x}^{fg}(\tau_{\mathcal{U}}))] = \varphi(s,x)$ for $(s,x) \in \Lambda$. Then φ is $Y_{s,x}^{fg}$-harmonic. □

Lemma 5.2. *Let Φ be a bounded measurable function on $(0,T)\times\partial D$ and $u(s,x) = E^{s,x}[\Phi(Y^{fg}_{s,x}(\tau_D))]$ for $(s,x)\in(0,T)\times D$. Then u is $Y^{fg}_{s,x}$-harmonic in $\Lambda = (0,T)\times D$.*

Proof. By the mean value property of the diffusion $Y^{fg}_{s,x}$, we get

$$u(s,x) = \int_{\partial V} u(y)Q^{s,x}(\{Y^{fg}_{s,x}\in \mathrm{d}y\}) = E^{s,x}[u(Y^{fg}_{s,x}(\tau_V))]$$

for every open set $V\subset D$ such that $\mathrm{cl}\{V\}\subset D$. Then u is $Y^{fg}_{s,x}$-harmonic. □

Given a bounded domain $D\subset\mathbb{R}^d$ and a function Φ on $(0,T)\times\partial D$, the problem consisting in finding a function $u:(0,T)\times D\to\mathbb{R}$ such that it is $Y^{fg}_{s,x}$-harmonic in $\Lambda=(0,T)\times D$ and $\lim_{t\to\tau_D} u(Y^{fg}_{s,x}(t)) = \Phi(Y^{fg}_{s,x}(\tau_D))$, $Q^{s,x}$-a.s. for $(s,x)\in\Lambda$, where $Q^{s,x}$ is a law of $Y^{fg}_{s,x}$, is called the stochastic Dirichlet problem generated by the diffusion process $Y^{fg}_{s,x}$.

Theorem 5.1. *Let $T>0$ and let $D\subset\mathbb{R}^d$ be a bounded domain. Assume $(f,g)\in\mathcal{C}(F)\times\mathcal{C}(G)$ with F and G such that the pair (F,G) generates a uniformly semielliptic set-valued diffusion operator $\mathbb{L}_{FG}$. Let Φ be a bounded measurable function on $(0,T)\times\partial D$. For every $(s,x)\in\Lambda=(0,T)\times D$, there is a unique in law weak solution $X^{fg}_{s,x}$ of $SDE(f,g)$ with initial condition $X^{fg}_{s,x}(s)=x$ a.s. such that the function $u(s,x)=E^{s,x}[\Phi(\tau_D, X^{fg}_{s,x}(\tau_D))]$ defined for $(s,x)\in\Lambda$ is a solution of the stochastic Dirichlet problem*

$$\begin{cases} \left(\mathcal{L}_{fg}v\right)(t,x) = 0 \quad \text{for } (t,x)\in[0,T)\times D, \\ \lim_{t\to\tau_D} v(t,x) = \Phi(\tau_D, X^{fg}_{s,x}(\tau_D)) \quad \text{for } (s,x)\in(0,T)\times\partial D, \end{cases} \tag{5.3}$$

where $\tau_D = \inf\{r\in(s,T): X^{fg}_{s,x}(r)\notin D\}$.

Proof. By virtue of Remark 11.1 of Chap. 1 and Theorem 11.1 of Chap. 1, for every $(s,x)\in\Lambda=(0,T)\times D$, there is a unique in law weak solution $X^{fg}_{s,x}$ of $SDE(f,g)$ with initial condition $X^{fg}_{s,x}(s)=x$ a.s. such that the process $Y^{fg}_{s,x} = (s+t,(X^{fg}_{s,x}(s+t)))_{0\le t\le T-s}$ is an Itô diffusion with characteristic operator $\mathcal{L}_{fg}$. By virtue of Lemma 5.2, the function $u(s,x)=E^{s,x}[\Phi(Y^{fg}_{s,x}(\tau_D - s))]$ defined for $(s,x)\in\Lambda$ is $Y^{fg}_{s,x}$-harmonic in Λ. Then by Lemma 5.1, one has $(\mathcal{L}_{fg}u)(s,x)=0$ for $(s,x)\in\Lambda$. Fix $(s,x)\in\Lambda$ and let $\mathcal{U}_k=(0,T)\times D_k$, where $(D_k)_{k=1}^\infty$ is an increasing sequence of open sets $D_k\subset\mathbb{R}^d$ such that $D_k\subset D$ and $D=\bigcup_{k=1}^\infty D_k$. Put $\tau_{\mathcal{U}_k}=\inf\{r\in(0,T-s]: Y^{fg}_{s,x}(r)\notin\mathcal{U}_k\}$ and $\tau_k=\inf\{r\in(s,T]: X^{fg}_{s,x}(r)\notin D_k\}$. It can be verified that $\tau_{\mathcal{U}_k}=\tau_k - s$ for $k=1,2,\ldots$. Let $\tau=\inf\{r\in(s,T]: X^{fg}_{s,x}(r)\notin D\}$. By virtue of Theorem 9.6 of Chap. 1, we get

$$\begin{aligned} u(Y^{fg}_{s,x}(\tau_{\mathcal{U}_k})) &= E^{\tau_k, X^{fg}_{s,x}(\tau_k)}[\Phi(\tau, X^{fg}_{s,x}(\tau))] \\ &= E^{s,x}[\theta_{\tau_k}(\Phi(\tau, X^{fg}_{s,x}(\tau)))|\mathcal{F}_{\tau_k}] = E^{s,x}[\Phi(\tau, X^{fg}_{s,x}(\tau))|\mathcal{F}_{\tau_k}]. \end{aligned}$$

Let us note that the process $(M_k)_{k\geq 1}$ defined by $M_k = E^{s,x}[\Phi(\tau, X_{s,x}^{fg}(\tau))|\mathcal{F}_{\tau_k}]$ is a bounded discrete-time martingale with respect to the discrete filtration $(\mathcal{F}_{\tau_k})_{k\geq 1}$. By virtue of Remark 4.1 of Chap. 1, we get

$$\lim_{k\to\infty} u(Y_{s,x}^{fg}(\tau_{\mathcal{U}_k}))=\lim_{k\to\infty} E^{s,x}[\Phi(\tau, X_{s,x}^{fg}(\tau))|\mathcal{F}_{\tau_k}]=E[\Phi(\tau, X_{s,x}^{fg}(\tau))|\mathcal{F}_\infty]=\Phi(\tau, X_{s,x}^{fg}(\tau))$$

a.s. and in the $\mathbb{L}^2(\Omega, \mathcal{F}_\infty, Q^{s,x}, \mathbb{R})$-topology, where $\mathcal{F}_\infty = \sigma(\{\mathcal{F}_{\tau_1}, \mathcal{F}_{\tau_2}, \dots\})$. Moreover, a sequence $(N^k)_{k=1}^\infty$ of stochastic processes $N_t^k = u(Y_{s,x}^{fg}(\tau_{\mathcal{U}_k} \vee (t \wedge \tau_{\mathcal{U}_{k+1}}))) - u(Y_{s,x}^{fg}(\tau_{\mathcal{U}_k}))$ is a martingale with respect to a filtration $\mathcal{G}_k = (\mathcal{G}_t^k)_{t\geq 0}$ of the form $\mathcal{G}_t^k = \mathcal{F}_{\tau_{\mathcal{U}_k}\vee(t\wedge\tau_{\mathcal{U}_{k+1}})}$ for $t \geq 0$. Therefore, by Doob's martingale inequality, we get

$$Q^{s,x}\left(\left\{\sup_{\tau_{\mathcal{U}_k}\leq r\leq\tau_{\mathcal{U}_{k+1}}} |u(Y_{s,x}^{fg}(\tau_{\mathcal{U}_{k+1}})) - u(Y_{s,x}^{fg}(\tau_{\mathcal{U}_k}))| > \varepsilon\right\}\right)$$
$$\leq \frac{1}{\varepsilon^2} E^{s,x}[|u(Y_{s,x}^{fg}(\tau_{\mathcal{U}_{k+1}})) - u(Y_{s,x}^{fg}(\tau_{\mathcal{U}_k}))|^2] \to 0$$

for every $\varepsilon > 0$ as $k \to \infty$. From this and the equality $\lim_{k\to\infty} u(Y_{s,x}^{fg}(\tau_{\mathcal{U}_k})) = \Phi(\tau, X_{s,x}^{fg}(\tau)))$ a.s., it follows that

$$\lim_{k\to\infty} u(s + \tau_{\mathcal{U}_k}, X_{s,x}^{fg}(s + \tau_{\mathcal{U}_k})) = \lim_{k\to\infty} u(\tau_k, X_{s,x}^{fg}(\tau_{\tau_k})) = \Phi(\tau, X_{s,x}^{fg}(\tau))),$$

which is equivalent to $\lim_{t\to\tau_D} u(t, X_{s,x}^{fg}(t)) = \Phi(\tau_D, X_{s,x}^{fg}(\tau_D))$. □

Theorem 5.2. *Let $T > 0$ and let $D \subset \mathbb{R}^d$ be a bounded domain. Assume $(f, g) \in \mathcal{C}(F) \times \mathcal{C}(G)$ with F and G such that the pair (F, G) generates the uniformly semielliptic set-valued diffusion operator $\mathbb{L}_{FG}$. Let Φ be a bounded measurable function on $(0, T) \times \partial D$. If $\varphi : (0, T) \times D \to \mathbb{R}$ is bounded, $Y_{s,x}^{fg}$-harmonic in $(0, T)\times D$ such that $\lim_{t\to\tau_D} \varphi(t, X_{s,x}^{fg}(t)) = \Phi(\tau_D, X_{s,x}^{fg}(\tau_D))$, $Q^{s,x}$-a.s. for $(s, x) \in (0, T) \times D$, where $X_{s,x}^{fg}$ is the unique in law weak solution of $SDE(f, g)$ with initial condition $X_{s,x}^{fg}(s) = x$ a.s. such that $Y_{s,x}^{fg}(t) = (s+t, X_{s,x}^{fg}(s+t))$ for $t \in [0, T-s]$, then $\varphi(s, x) = E^{s,x}[\Phi(\tau_D, X_{s,x}^{fg}(\tau_D))]|$ for $(s, x) \in (0, T) \times D$.*

Proof. Let $\tau_{\mathcal{U}_k}$, τ_k, and τ be as in the proof of Theorem 5.1. Since φ is $Y_{s,x}^{fg}$-harmonic in $(0, T) \times D$, it follows immediately from the definition that $\varphi(s, x) = E^{s,x}[\Phi(Y_{s,x}^{fg}(\tau_{\mathcal{U}_k}))]$ for every $k = 1, 2, \dots$. Hence, similarly as in the proof of Theorem 5.1, by the properties of φ, it follows that $\varphi(s, x) = \lim_{k\to\infty} E^{s,x}[\varphi(s + \tau_{\mathcal{U}_k}, X_{s,x}^{fg}(s + \tau_{\mathcal{U}_k}))] = \lim_{k\to\infty} E^{s,x}[\varphi(\tau_k, X_{s,x}^{fg}(\tau_k))] = \Phi(\tau, X_{s,x}^{fg}(\tau))]$ for every $(s, x) \in (0, T) \times D$. □

Let $T > 0$ and let $D \subset \mathbb{R}^d$ be a bounded domain. Given a continuous function $\psi : (0, T) \times D \to \mathbb{R}$ and $(f, g) \in \mathcal{C}(F) \times \mathcal{C}(G)$ with F and G such that the pair

(F, G) generates the uniformly semielliptic set-valued diffusion operator $\mathbb{L}_{FG}$, the problem consisting in finding a function $v : (0, T) \times D \to \mathbb{R}$ such that

$$\begin{cases} \left(\mathcal{L}_{fg} v\right)(t, x) = -\psi(s, x) \quad \text{for } (t, x) \in [0, T) \times D \\ \lim_{t \to \tau_D} v(t, X_{s,x}^{fg}(t)) = 0, \quad Q^{s,x} - a.s. \quad \text{for } (s, x) \in (0, T) \times D, \end{cases} \tag{5.4}$$

where $\mathcal{L}_{fg}$ is the characteristic operator of the Itô diffusion $Y_{s,x}^{fg}$ defined by the unique in law solution $X_{s,x}^{fg}$ of $SDE(f, g)$ with initial condition $X_{s,x}^{fg}(s) = x$ a.s., is said to be the stochastic Poisson problem generated by $Y_{s,x}^{fg}$.

Theorem 5.3. *Let $T > 0$ and let $D \subset \mathbb{R}^d$ be a bounded domain. Assume that $(f, g) \in \mathcal{C}(F) \times \mathcal{C}(G)$ with F and G such that the pair (F, G) generates the uniformly semielliptic set-valued diffusion operator $\mathbb{L}_{FG}$. Let $X_{s,x}^{fg}$ be the unique in law weak solution of $SDE(f, g)$ with initial condition $X_{s,x}^{fg}(s) = x$ a.s. and let $\varphi : (0, T) \times D \to \mathbb{R}$ be a continuous function such that $E^{s,x}\left[\int_0^{\tau_D} |\varphi(t, X_{s,x}^{fg}(t))| \mathrm{d}t\right] < \infty$ for every $(s, x) \in (0, T) \times D$. Then the function $v : (0, T) \times D \to \mathbb{R}$ defined by $v(s, x) = E^{s,x}\left[\int_0^{\tau_D} \varphi(t, X_{s,x}^{fg}(t)) \mathrm{d}t\right]$ is a solution of the stochastic Poisson problem (5.4), where $\tau_D = \inf\{r \in (s, T) : X_{s,x}^{fg}(r) \neq D\}$.*

Proof. Fix $(s, x) \in (0, T) \times D$ and choose an open set $\mathcal{U} \subset (0, T) \times D$ such that $(s, x) \in \mathcal{U}$. Put $\tau_{\mathcal{U}} = \inf\{r \in (s, T - s) : Y_{s,x}^{fg}(r) \notin \mathcal{U}\}$ and $\eta = \int_0^{\tau_D} \varphi(t, X_{s,x}^{fg}(t)) \mathrm{d}t$, where $(Y_{s,x}^{fg}(t))_{0 \le t \le T-s}$ is the Itô diffusion defined in Theorem 11.1 of Chap. 1. By virtue of Theorem 9.6 of Chap. 1, we get

$$\begin{aligned} \frac{E^{s,x}[v(Y_{s,x}^{fg}(t))] - v(s, x)}{E^{s,x}[\tau_{\mathcal{U}}]} &= \frac{1}{E^{s,x}[\tau_{\mathcal{U}}]} \left(E^{s,x}[E^{Y_{s,x}^{fg}}[\eta]] - E^{s,x}[\eta]\right) \\ &= \frac{1}{E^{s,x}[\tau_{\mathcal{U}}]} \left(E^{s,x}[E^{s,x}[\theta_{\tau_{\mathcal{U}}} \eta | \mathcal{F}_{\tau_{\mathcal{U}}}] - E^{s,x}[\eta]\right) = \frac{1}{E^{s,x}[\tau_{\mathcal{U}}]} \left(E^{s,x}[\theta_{\tau_u} \eta - \eta]\right). \end{aligned}$$

Let $(\eta_k)_{k=1}^{\infty}$ be a sequence of Riemann approximating sums for η of the form $\eta_k = \sum_i^{n_k} \varphi(t_i, X_{s,x}^{fg}(t_i)) \mathbb{I}_{\{t_i < \tau_D\}} \Delta t_i$ for $k = 1, 2, \ldots$. Since $\theta_t \eta = \sum_i^{n_k} \varphi(t_i + t, X_{s,x}^{fg}(t_i + t)) \mathbb{I}_{\{t_i + t < \tau_D\}} \Delta t_i$ for $k = 1, 2, \ldots$ and $\eta_k \to \eta$ as $k \to \infty$, it follows that $\theta_{\tau_{\mathcal{U}}} \eta = \int_{\tau_{\mathcal{U}}}^{\tau_D} \varphi(t, X_{s,x}^{fg}(t)) \mathrm{d}t$. Therefore,

$$\frac{E^{s,x}[v(Y_{s,x}^{fg}(\tau_{\mathcal{U}}))] - v(s, x)}{E^{s,x}[\tau_{\mathcal{U}}]} = -\frac{1}{E^{s,x}[\tau_{\mathcal{U}}]} E^{s,x}\left[\int_0^{\tau_{\mathcal{U}}} \varphi(t, X_{s,x}^{f,g}(t)) \mathrm{d}t\right].$$

By the continuity of φ, it follows that

$$\lim_{\mathcal{U} \to (s,x)} \frac{E^{s,x}[v(Y_{s,x}^{fg}(\tau_{\mathcal{U}}))] - v(s, x)}{E^{s,x}[\tau_{\mathcal{U}}]} = -\varphi(s, x).$$

Then $(\mathcal{L}_{fg}v)(s,x) = -\varphi(s,x)$ for $(s,x) \in (0,T) \times D$. Let D_k and τ_k be as in the proof of Theorem 5.1 and put

$$H(s,x) = E^{s,x}\left[\int_0^{\tau_D} |\varphi(t, X_{s,x}^{fg}(t))|\mathrm{d}t\right].$$

Similarly as above, we get

$$\begin{aligned} E^{s,x}[H(\tau_k \wedge t, X_{s,x}^{f,g}(\tau_k \wedge t))] &= E^{s,x}\left[E^{s,x}\left[\int_{\tau_k \wedge t}^{\tau_D} |\varphi(u, X_{s,x}^{fg}(u))|\mathrm{d}u|\mathcal{F}_{\tau_k \wedge t}\right]\right] \\ &= E^{s,x}\left[\int_{\tau_k \wedge t}^{\tau_D} |\varphi(u, X_{s,x}^{fg}(u))|\mathrm{d}u\right]. \end{aligned}$$

Passing to the limit in the above equality with $k \to \infty$ and $t \to \tau_D$, we obtain $\lim_{t \to \tau_D, k \to \infty} E^{s,x}[H(\tau_k \wedge t, X_{s,x}^{fg}(\tau_k \wedge t))] = 0$, which implies that $\lim_{t \to \tau_D} v(t, X_{s,x}^{fg}(t)) = 0$ $Q^{s,x}$-a.s. for $(s,x) \in (0,T) \times D$. □

Immediately from Theorems 5.1 and 5.3 we obtain the following existence and representation theorems for the stochastic Dirichlet–Poisson problem generated by an Itô diffusion $Y_{s,x}^{fg}$.

Theorem 5.4. *Let $(f,g) \in \mathcal{C}(F) \times \mathcal{C}(G)$ with F and G such that the pair (F,G) generates a uniformly semielliptic set-valued diffusion operator $\mathbb{L}_{FG}$. Assume $T > 0$ and that $D \subset \mathbb{R}^d$ is a bounded domain, and let $X_{s,x}^{fg}$ be the unique in law weak solution of $SDE(f,g)$ with initial condition $X_{s,x}^{fg}(s) = x$ a.s. for $(s,x) \in (0,T) \times D$. Assume that Φ is continuous and ψ is continuous and bounded on $(0,T) \times \partial D$ and $(0,T) \times D$, respectively, such that $E^{s,x}\left[\int_0^{\tau_D} |\psi(t, X_{s,x}^{fg}(t))|\mathrm{d}t\right] < \infty$ for every $(s,x) \in (0,T) \times D$, where $\tau_D = \inf\{r \in (s,T] : X_{s,x}^{fg}(r) \notin D\}$. The function $v : (0,T) \times D \to \mathbb{R}$ defined by*

$$v(s,x) = E^{s,x}[\Phi(\tau_D, X_{s,x}^{fg}(\tau_D))] + E^{s,x}\left[\int_0^{\tau_D} \psi(t, X_{s,x}^{fg}(t)\mathrm{d}t\right]$$

for $(s,x) \in (0,T) \times D$ is a solution of the stochastic Dirichlet–Poisson problem

$$\begin{cases} \left(\mathcal{L}_{fg}v\right)(t,x) = -\psi(s,x) \quad \text{for } (t,x) \in [0,T) \times D, \\ \lim_{t \to \tau_D} v(t, X_{s,x}^{fg}(t)) = \Phi(\tau_D, X_{s,x}^{fg}(\tau_D)), \quad Q^{s,x} - a.s. \quad \text{for } (s,x) \in (0,T) \times D. \end{cases} \tag{5.5}$$

□

Theorem 5.5. *Let $(f,g) \in \mathcal{C}(F) \times \mathcal{C}(G)$ with F and G such that the pair (F,G) generates a uniformly semielliptic set-valued operator diffusion $\mathbb{L}_{FG}$. Assume $T > 0$ and that $D \subset \mathbb{R}^d$ is a bounded domain; let $X_{s,x}^{fg}$ be the unique in law* *weak solution of* $SDE(f,g)$ *with initial condition* $X_{s,x}^{fg}(s) = x$ *a.s. for* $(s,x) \in (0,T) \times D$*, and let* Φ *be continuous and* ψ *continuous and bounded on* $(0,T) \times \partial D$ *and* $(0,T) \times D$*, respectively, such that* $E^{s,x}\left[\int_0^{\tau_D} |\psi(t, X_{s,x}^{fg}(t))| \mathrm{d}t\right] < \infty$ *for every* $(s,x) \in (0,T) \times D$*, where* $\tau_D = \inf\{r \in (s,T] : X_{s,x}^{fg}(r) \notin D\}$*. If* $v \in C^{1,2}((0,T) \times D, \mathbb{R})$ *is such that* $|v(s,x)| \le C\left(1 + E^{s,x}\left[\int_0^{\tau_D} |\psi(t, X_{s,x}^{fg}(t))| \mathrm{d}t\right]\right)$ *for* $(s,x) \in (0,T) \times D$ *and* v *satisfies (5.5), then*

$$v(s,x) = E^{s,x}[\Phi(\tau_D, X_{s,x}^{fg}(\tau_D))] + E^{s,x}\left[\int_0^{\tau_D} \psi(t, X_{s,x}^{fg}(t)) \mathrm{d}t\right]$$

for $(s,x) \in (0,T) \times D$. □

6 Existence Theorems for Partial Differential Inclusions

Let $F : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d \times m})$ generate a uniformly semielliptic set-valued operator $\mathbb{L}_{FG}$ on $C_0^2(\mathbb{R}^{d+1})$ and let $\mathcal{A}_{FG}$ be a family of uniformly parabolic diffusion generators on $\mathcal{D}_{FG}$ defined by (1). Similarly as in Sect. 3, assume that Q is a bounded domain in $\mathbb{R}^+ \times \mathbb{R}^d$ lying in the strip $(0,T) \times \mathbb{R}^d$ for a given $T > 0$. Put $\bar{B} = \bar{Q} \cap [\{t = 0\} \times \mathbb{R}^d]$ and $\bar{B}_T = \bar{Q} \cap [\{t = T\} \times \mathbb{R}^d]$ and assume that $\bar{B}$ and $\bar{B}_T$ are nonempty. Let $B_T = \mathrm{Int}(\bar{B}_T)$ and $B = \mathrm{Int}(\bar{B})$. Denote by S_0 the boundary of Q lying in the strip $(0,T) \times \mathbb{R}^d$ and let $S = S_0 \setminus B_T$. The set $\partial Q = B \cup S$ is a parabolic boundary of Q. Let $c : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}$, $\varphi : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}$, $\psi : \mathbb{R}^d \to \mathbb{R}$, and $\gamma : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}$ be given. The following results for partial differential inclusions follow immediately from the above existence and representation theorems for partial differential equations.

Theorem 6.1. *Assume that* F *and* G *are convex-valued and uniformly Hölder continuous such that the pair* (F,G) *generates a uniformly semielliptic set-valued operator* $\mathbb{L}_{FG}$ *on a bounded domain* $Q \subset (0,T) \times \mathbb{R}^d$*. Let* c *and* φ *be uniformly Hölder continuous in* $\bar{Q}$ *and assume that* ψ *and* γ *are continuous on* $\bar{B}$ *and* $\bar{S}$*, respectively. If furthermore, there exists a barrier corresponding to* (f,g,c) *for every pair* $(f,g) \in \mathcal{C}(F) \times \mathcal{C}(G)$ *at every point of* S*, then the initial–boundary value problem*

$$\begin{cases} u_t'(t,x) \in (\mathbb{L}_{FG} u(t,\cdot))(t,x) + c(t,x)u(t,x) - \varphi(t,x) \\ \text{for } \ (t,x) \in Q \cup B_T, \\ u(0,x) = \psi(x) \ \text{ for } \ x \in B, \\ u(t,s) = \gamma(t,x) \ \text{ for } \ (t,x) \in S, \end{cases} \tag{6.1}$$

possesses at least one solution.

Proof. By virtue of Corollary 2.2 of Chap. 2, it follows that the functions $f : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^d$ and $g : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^{d\times m}$ defined by $f(t,x) = s(F(t,x))$ and $g(t,x) = s(G(t,x))$ for $(t,x) \in \mathbb{R}^+ \times \mathbb{R}^d$, where s is a Steiner point map defined by (2.1) of Chap. 1 are continuous selectors of F and G that satisfy the conditions of Theorem 3.1. Therefore, by virtue of this theorem, there exists a unique solution v of the first initial–boundary problem (3.1) that, by the properties of the selectors f and g and the definition of $\mathbb{L}_{FG}$, satisfies $(\mathbb{L}_{fg}v(t,\cdot))(t,x) \in (\mathbb{L}_{FG}v(t,\cdot))(t,x)$ for $(t,x) \in Q \cup B_T$. Therefore, v is a solution of (6.1). □

Remark 6.1. Immediately from Corollary 3.1, it follows that if $Q = (0,T)\times D$ is as in Remark 3.2, then for F, G, c, φ, and γ satisfying the conditions of Theorem 6.1, the initial–boundary valued problem (6.1) possesses at least one solution. □

Similarly as above, we also obtain the following existence theorem for Cauchy problems for partial differential inclusions.

Theorem 6.2. *Assume that F and G are bounded, convex-valued, Hölder continuous in x uniformly with respect to $(t,x) \in [0,T] \times \mathbb{R}^d$, and such that (F,G) generates a uniformly semielliptic diffusion set-valued operator $\mathbb{L}_{FG}$ on $\mathcal{D}_{FG}$. Furthermore, let φ be continuous and Hölder continuous in x uniformly with respect to (t,x) on compact subsets of $[0,T] \times \mathbb{R}^d$ and let ψ be continuous on $\mathbb{R}^d$. If there are positive numbers α and A such that* $\max(|\varphi(t,x)|, |\psi(x)|) \le A \cdot \exp\left(\alpha|x|^2\right)$ *for $(t,x) \in [0,T] \times \mathbb{R}^d$, then there is $\bar{c} > 0$ such that the Cauchy problem*

$$\begin{cases} u'_t(t,x) \in (\mathbb{L}_{FG}u(t,\cdot))(t,x) + c(t,x)u(t,x) - \varphi(t,x) \ \text{ for } \ (t,x) \in (0,T] \times \mathbb{R}^d, \\ u(0,x) = \psi(x) \ \text{ for } \ x \in \mathbb{R}^d, \end{cases} \quad]$$

possesses a solution in the strip $[0,T^*] \times \mathbb{R}^d$, where $T^* = \min\{T, \bar{c}/\alpha\}$.

Proof. The result follows immediately from Theorem 3.2 for f and g defined similarly as in the proof of Theorem 6.1 by setting $f(t,x) = s(F(t,x))$ and $g(t,x) = s(G(t,x))$ for $(t,x) \in \mathbb{R}^+ \times \mathbb{R}^d$. □

Theorem 6.3. *Assume that conditions (i)–(iv) of $(\mathcal{A})$ are satisfied, $T > 0$, $\tilde{h} \in C_0^{1,2}(\mathbb{R}^{d+1})$, and let $c \in C([0,T] \times \mathbb{R}^d, \mathbb{R})$ be bounded. For every $(s,x) \in [0,T) \times \mathbb{R}^d$, there is a weak solution $(\mathcal{P}_{\mathbb{F}}, X_{s,x}^{fg}, B)$ of $SFI(F,G)$ with initial condition $x_s = x$ a.s. such that the function*

$$v(t,s,x) = E^{s,x}\left[\exp\left(-\int_s^{s+t} c(\tau, X_{s,x}(\tau))\mathrm{d}\tau\right)\tilde{h}(s+t, X_{s,x}(s+t))\right]$$

satisfies

$$\begin{cases} v'_t(t,s,x) \in (\mathcal{A}_{FG}v(t,\cdot))(s,x) - c(s,x)v(t,s,x) \\ \text{for } \ (s,x) \in [0,T) \times \mathbb{R}^d \ \text{ and } \ t \in [0,T-s] \\ v(0,s,x) = \tilde{h}(s,x) \ \text{ for } \ (s,x) \in [0,T) \times \mathbb{R}^d, \end{cases} \tag{6.2}$$

where $\mathcal{A}_{FG}$ is the set-valued parabolic diffusion generator defined above.

Proof. By virtue of Michael's continuous selection theorem, there are $f \in \mathcal{C}(F)$ and $g \in \mathcal{C}(G)$. By Theorems 9.3, 9.4, and 11.1 of Chap. 1, there is a unique in law weak solution $(\mathcal{P}_{\mathbb{F}}, X_{s,x}^{fg}, B)$ of $SDE(f, g)$ with initial condition $x_s = x$ a.s. such that the process $Y_{s,x}^{fg} = (s+t, X_{s,x}^{fg}(s+t))_{0\le t\le T-s}$ is an Itô diffusion with infinitesimal generator $\mathcal{A}_{fg} \in \mathcal{A}_{FG}$. By virtue of Theorem 12.1 of Chap. 1, the function

$$v(t,s,x) = E^{s,x}\left[\exp\left(-\int_0^t c(Y_{s,x}^{fg}(\tau))\mathrm{d}\tau\right)\tilde{h}(Y_{s,x}^{fg}(t))\right]$$

defined for $\tilde{h} \in C_0^{1,2}(\mathbb{R}^{d+1})$, $(s,x) \in [0,T)\times\mathbb{R}^d$ and $t \in [0, T-s]$ satisfies equation 12.1 of Chap. 1. But $(\mathcal{A}_{fg}v(t,\cdot))(s,x) \in (\mathcal{A}_{FG}v(t,\cdot))(s,x)$ for $(s,x) \in [0,T)\times\mathbb{R}^d$ and $t \in [0, T-s]$. Therefore, conditions (6.2) are satisfied. It is easy to see that

$$v(t,s,x) = E^{s,x}\left[\exp\left(-\int_s^{s+t} c(\tau, X_{s,x}(\tau))\mathrm{d}\tau\right)\tilde{h}(s+t, X_{s,x}(s+t))\right]$$

for $(s,x) \in [0,T)\times\mathbb{R}^d$. □

Theorem 6.4. *Assume that conditions (i), (iii), (iv′), and (v) of* $(\mathcal{A})$ *are satisfied,* $T > 0$, *and let* $\tilde{h} \in C_0^{1,2}(\mathbb{R}^{d+1})$. *Suppose* $c \in C([0,T]\times\mathbb{R}^d, \mathbb{R})$ *and* $v \in C^{1,1,2}([0,T]\times[0,T]\times\mathbb{R}^d, \mathbb{R})$ *are bounded such that*

$$\begin{cases} v_t'(t,s,x) - v_s'(t,s,x) \in (\mathbb{L}_{FG}v(t,\cdot))(s,x) - c(s,x)v(t,s,x) \\ \text{for } \ (s,x) \in [0,T)\times\mathbb{R}^d \ \text{ and } \ t \in [0, T-s], \\ v(0,s,x) = \tilde{h}(s,x) \ \text{ for } \ (s,x) \in [0,T)\times\mathbb{R}^d. \end{cases} \tag{6.3}$$

For every $(s,x) \in [0,T)\times\mathbb{R}^d$, *there exists* $\hat{X}_{s,x} \in \mathcal{X}_{s,x}(F,G)$ *defined on the probability space* $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P})$ *such that*

$$v(t,s,x) = \hat{E}\left[\exp\left(-\int_s^{s+t} c(\tau, \hat{X}_{s,x}(\tau))\mathrm{d}\tau\right)\tilde{h}(s+t, \hat{X}_{s,x}(s+t))\right]$$

for $(s,x) \in [0,T)\times\mathbb{R}^d$ *and* $t \in [0, T-s]$.

Proof. By virtue of Theorem 2.6, for every $k,m \ge 1$, there are $f_k \in \mathcal{C}(F)$ and $g_k^m \in \mathcal{C}(G)$ such that $\sigma_k^m = g_k^m\cdot(g_k^m)^*$ is uniformly positive definite and

$$\lim_{m\to\infty} |v_t'(t,s,x) - v_s'(t,s,x) - [(\mathbb{L}_{f_k g_k^m}v(t,\cdot))(s,x) - c(s,x)v(t,s,x)]| \le 1/k \tag{6.4}$$

uniformly with respect to $(s,x) \in [0,T)\times\mathbb{R}^d$ and $t \in [0, T-s]$, where $\mathbb{L}_{f_k g_k^m}$ is defined by $\mathbb{L}_{uv}$ for $u = f_k(t,x)$ and $v = g_k^m(t,x)\cdot(g_k^m(t,x))^*$. But $(\mathcal{A}_{f_k g_k^m}v(t,\cdot))(s,x) = v_s'(t,s,x) + (\mathbb{L}_{f_k g_k^m}v(t,\cdot))(s,x)$. Then inequality (6.4) can be written in the form

$$\lim_{m\to\infty} |v_t'(t,s,x) - [(\mathcal{A}_{f_k g_k^m}v(t,\cdot))(s,x) - c(s,x)v(t,s,x)]| \le 1/k \tag{6.5}$$

for $k \geq 1$ uniformly with respect to $(s, x) \in [0, T) \times \mathbb{R}^d$ and $t \in [0, T - s]$. By virtue of Theorems 9.3, 9.4, and 11.1 of Chap. 1, for every $(s, x) \in [0, T) \times \mathbb{R}^d$ and $k, m = 1, 2, \ldots$, there is a unique in law weak solution $(\mathcal{P}_k^m, X_k^m, B_k^m)$ of the stochastic differential equation $x_t = f_k(t, x_t)\mathrm{d}t + g_k^m(t, x_t)\mathrm{d}B_t$ with $\mathcal{P}_k^m = (\Omega_k^m, \mathcal{F}_k^m, \mathbb{F}_k^m, P_k^m)$ and initial condition $x_s = x$ a.s. such that the process $Y_k^m = (s + t, X_k^m(s + t))_{0 \leq t \leq T - s}$ is an Itô diffusion with infinitesimal generator $\mathcal{A}_{f_k g_k^m} \in \mathcal{A}_{FG}$. We have $X_k^m \in \mathcal{X}_{s,x}(F, G)$ for $k, m \geq 1$, where $\mathcal{X}_{s,x}(F, G)$ denotes the set of all weak solutions (equivalence classes of) of $SFI(F, G)$ with initial condition $x_s = x$ a.s. By the weak compactness of the set $\mathcal{X}_{s,x}(F, G)$, for every fixed $k \geq 1$ there are an increasing subsequence $(m_r)_{r=1}^\infty$ of $(m)_{m=1}^\infty$, a probability space $\tilde{\mathcal{P}}_k = (\tilde{\Omega}_k, \tilde{\mathcal{F}}_k, \tilde{P}_k)$, and continuous processes $\tilde{X}_k^{m_r}$, $\tilde{X}_k$ on $\tilde{\mathcal{P}}_k$ such that $P(X_k^{m_r})^{-1} = P(\tilde{X}_k^{m_r})^{-1}$ for $k, r = 1, 2, \ldots$ and such that $\sup_{0 \leq t \leq T} |\tilde{X}_k^{m_r}(t) - \tilde{X}_k(t)| \to 0$ $\tilde{P}_k$-a.s. for $k \geq 1$ as $r \to \infty$. By Theorem 9.4 of Chap. 1, it follows that $\tilde{X}_k^{m_r}$ is a weak solution of the stochastic differential equation $\mathrm{d}x_t = f_k(t, x_t)\mathrm{d}t + g_k^{m_r}(t, x_t)\mathrm{d}B_t$ with initial condition $x_s = x$ a.s. Then $\tilde{X}_k \in \mathcal{X}_{s,x}(F, G)$ for every $k \geq 1$. Let us observe that by virtue of Lemma 10.1 of Chap. 1 and Remark 10.4 of Chap. 1, for fixed $k \geq 1$, every $r \geq 1$, and every bounded domain $D_\lambda = \{z \in \mathbb{R}^d : |z| < \lambda\}$ with $\lambda > 0$, one has $E_k^{m_r}[\tau_k^{m_r}] < \infty$, where $\tau_k^{m_r} = \inf\{t > s : X_k^{m_r}(t) \notin D_\lambda\}$ and $E_k^{m_r}$ denotes the mean value operator with respect to the probability measure $P_k^{m_r}$. By virtue of Theorem 5.2 of Chap. 4, we have $P(\tau_k^{m_r})^{-1} = P(\tilde{\tau}_k^{m_r})^{-1}$, where $\tilde{\tau}_k^{m_r} = \inf\{t > s : \tilde{X}_k^{m_r}(t) \notin D_\lambda\}$. Hence, by Chebyshev's inequality, it follows that

$$\tilde{P}_k(\{\tilde{\tau}_k^{m_r} > 2^l\}) = P_k^{m_r}(\{\tau_k^{m_r} > 2^l\}) \leq 2^{-l} E_k^{m_r}[\tau_k^{m_r}]$$

for every $l = 1, 2, \ldots$. Therefore $\sum_{l=1}^\infty \tilde{P}_k(\{\tilde{\tau}_k^{m_r} > 2^l\}) < \infty$, which by the Borel–Cantelli lemma, implies $\tilde{P}_k(\bigcap_{l=1}^\infty \bigcup_{s=l}^\infty \{\tilde{\tau}_k^{m_r} > 2^l\}) = 0$. Then there is a subsequence $(l_s)_{s=1}^\infty$ of $(l)_{l=1}^\infty$ such that $\tilde{P}_k(\{\tilde{\tau}_k^{m_r} > 2^{l_s}\}) = 0$ for every $k, r, s \geq 1$. Therefore, $\tilde{\tau}_k^{m_r} \leq 2^{l_s}$ $\tilde{P}_k$-a.s. for fixed $k, s \geq 1$ and every $r \geq 1$. Hence, by virtue of Theorem 5.1 of Chap. 4, it follows that $\lim_{r \to \infty} |\tilde{\tau}_k^{m_r} - \tilde{\tau}_k| = 0$ a.s., where $\tilde{\tau}_k = \inf\{t > s : \tilde{X}_k(t) \notin D\}$. In a similar way, we can verify that there exist a subsequence $(k_r)_{r=1}^\infty$ of a sequence $(k)_{k=1}^\infty$ and continuous stochastic processes $\hat{X}_{k_r}$ and $\hat{X}$ on the probability space $\hat{\mathcal{P}} = (\hat{\Omega}, \hat{\mathcal{F}}, \hat{P})$ such that $P\hat{X}_{k_r}^{-1} = P\tilde{X}_{k_r}^{-1}$, $\sup_{0 \leq t \leq T} |\hat{X}_{k_r}(t) - \hat{X}(t)| \to 0$, and $\lim_{r \to \infty} |\hat{\tau}_{k_r}^\lambda - \hat{\tau}_\lambda| = 0$ $\hat{P}$-a.s. as $r \to \infty$, where $\hat{\tau}_{k_r}^\lambda$ and $\hat{\tau}_\lambda$ denote the first exit times of $\hat{X}_{k_r}$ and $\hat{X}$ from D_λ, respectively. Put $Y_k^{m_r}(t) = (s + t, X_k^{m_r}(s + t))_{0 \leq t \leq T - s}$, $\tilde{Y}_k^{m_r}(t) = (s + t, \tilde{X}_k^{m_r}(s + t))_{0 \leq t \leq T - s}$, $\tilde{Y}_k(t) = (s + t, \tilde{X}_k(s + t))_{0 \leq t \leq T - s}$, $\tilde{Y}(t) = (s + t, \tilde{X}(s + t))_{0 \leq t \leq T - s}$ for $k, r \geq 1$, $\hat{Y}_{k_r}(t) = (s + t, \hat{X}_{k_r}(s + t))_{0 \leq t \leq T - s}$, and $\hat{Y}(t) = (s + t, \hat{X}(s + t))_{0 \leq t \leq T - s}$. Let $K_\lambda = (0, T) \times D_\lambda$ and $\sigma_k^{m_r}$, $\tilde{\sigma}_k^{m_r}$, $\tilde{\sigma}_k$, $\hat{\sigma}_{k_r}^\lambda$, and $\hat{\sigma}_\lambda$ denote the first exit times from K_λ of $Y_k^{m_r}$, $\tilde{Y}_k^{m_r}$, $\tilde{Y}_k$, $\hat{Y}_{k_r}$, and $\hat{Y}$, respectively. Similarly, by $\tau_k^{m_r}$ and $\tilde{\tau}_k$, we denote the first exit times from D_λ of $X_k^{m_r}$ and $\tilde{X}_k$, respectively. We have $\sigma_k^{m_r} = \tau_k^{m_r} - s$, $\tilde{\sigma}_k = \tilde{\tau}_k - s$, $\hat{\sigma}_{k_r}^\lambda = \hat{\tau}_{k_r}^\lambda - s$, and $\hat{\sigma}_\lambda = \hat{\tau}_\lambda - s$. By virtue of Theorem 5.2 of Chap. 4 and Corollary 5.1 of Chap. 4, we have

$$E_k^{m_r}\left[\exp\left(-\int_0^{t\wedge\sigma_k^{m_r}} c(Y_k^{m_r}(\tau))\mathrm{d}\tau\right) v(t_0 - t\wedge\sigma_k^{m_r}, Y_k^{m_r}(t\wedge\sigma_k^{m_r}))\right]$$
$$= \tilde{E}_k\left[\exp\left(-\int_0^{t\wedge\tilde{\sigma}_k^{m_r}} c(\tilde{Y}_k^{m_r}(\tau))\mathrm{d}\tau\right) v(t_0 - t\wedge\tilde{\sigma}_k^{m_r}, \tilde{Y}_k^{m_r}(t\wedge\tilde{\sigma}_k^{m_r}))\right]$$

for fixed $t_0 \in (0, T - s]$, $0 \le t < t_0$, and $s \in [0, T)$. By Itô's formula, we obtain

$$E_k^{m_r}\left[\exp\left(-\int_0^{t\wedge\sigma_k^{m_r}} c(Y_k^{m_r}(\tau))\mathrm{d}\tau\right) v(t_0 - t\wedge\sigma_k^{m_r}, Y_k^{m_r}(t\wedge\sigma_k^{m_r}))\right] - v(t_0, s, x)$$
$$= E_k^{m_r}\int_0^{t\wedge\sigma_k^{m_r}} \exp\left(-\int_0^{\tau} c(Y_k^{m_r}(\tau))\mathrm{d}\tau\right)[(\mathcal{A}_{f_{n_k} g_k^{m_r}} v(t_0-\tau,\cdot))(Y_k^{m_r}(\tau))$$
$$- v_t'(t_0-\tau, Y_k^{m_r}(\tau)) - c(Y_k^{m_r}(\tau))v(t_0-\tau, Y_k^{m_r}(\tau))]\mathrm{d}\tau.$$

Similarly as above, we also get

$$E_k^{m_r}\int_0^{t\wedge\sigma_k^{m_r}} \exp\left(-\int_0^{\tau} c(Y_k^{m_r}(\tau))\mathrm{d}\tau\right)[\mathcal{A}_{f_{n_k} g_k^{m_r}} v(t_0-\tau, Y_{s,x}^{m_r}(\tau))$$
$$- v_t'(t_0-\tau, Y_k^{m_r}(\tau)) - c(Y_k^{m_r}(\tau))v(t_0-\tau, Y_k^{m_r}(\tau))]\mathrm{d}\tau$$
$$= \tilde{E}_k\int_0^{t\wedge\tilde{\sigma}_k^{m_r}} \exp\left(-\int_0^{\tau} c(\tilde{Y}_k^{m_r}(\tau))\mathrm{d}\tau\right)[\mathcal{A}_{f_{n_k} g_k^{m_r}} v(t_0-\tau, \tilde{Y}_k^{m_r}(\tau))$$
$$- v_t'(t_0-\tau, \tilde{Y}_k^{m_r}(\tau)) - c(\tilde{Y}_k^{m_r}(\tau))v(t_0-\tau, \tilde{Y}_k^{m_r}(\tau))]\mathrm{d}\tau.$$

Therefore,

$$\tilde{E}_k\left[\exp\left(-\int_0^{t\wedge\tilde{\sigma}_k^{m_r}} c(\tilde{Y}_k^{m_r}(\tau))\mathrm{d}\tau\right) v(t_0 - t\wedge\tilde{\sigma}_r^{k}, \tilde{Y}_k^{m_r}(t\wedge\tilde{\sigma}_r^{k}))\right] - v(t_0, s, x)$$
$$= \tilde{E}_k\int_0^{t\wedge\tilde{\sigma}^{m_r k}} \exp\left(-\int_0^{\tau} c(\tilde{Y}_k^{m_r}(\tau))\mathrm{d}\tau\right)[\mathcal{A}_{f_k g_k^{m_r}} v(t_0-\tau, \tilde{Y}_k^{m_r}(\tau))$$
$$- v_t'(t_0-\tau, \tilde{Y}_k^{m_r}(\tau)) - c(\tilde{Y}_k^{m_r}(\tau))v(t_0-\tau, \tilde{Y}_k^{m_r}(\tau))]\mathrm{d}\tau$$

for fixed $k = 1, 2, \ldots,\ s \in [0, T)$, $0 \le t < t_0$, and every $r \ge 1$. From this together with (6.5), it follows that

$$\lim_{r\to\infty}\left|\tilde{E}_k\left[\exp\left(-\int_0^{t\wedge\tilde{\sigma}_k^{m_r}} c(\tilde{Y}_k^{m_r}(\tau))\mathrm{d}\tau\right) v(t_0 - t\wedge\tilde{\sigma}_k^{m_r}, \tilde{Y}_k^{m_r}(t\wedge\tilde{\sigma}_k^{m_r}))\right]\right.$$
$$\left. - v(t_0, s, x)\right|$$

$$\leq \lim_{r\to\infty} \sup_{x\in} \mathbb{R}^d \tilde{E}_k \int_0^{t\wedge\tilde{\sigma}_k^{mr}} \exp\left(-\int_0^{\tau} c(s+\tau,x)\mathrm{d}\tau\right) |(\mathcal{A}_{f_k g_k^{mr}} v(t_0-\tau,\cdot))(s+t,x)$$
$$- v_t'(t_0-\tau,s+t,x) - c(s+t,x)v(t_0-\tau,s+t,x)|\mathrm{d}\tau$$
$$\leq \frac{1}{k}\sup_{x\in\mathbb{R}^d}\int_0^T \exp\left(-\int_s^{s+\tau} c(u,x)\mathrm{d}u\right)\mathrm{d}\tau \leq \frac{M}{k}$$

for $k{=}1,2,\dots$ and $s\in[0,T)$, where $M \geq \sup_{x\in\mathbb{R}^d}\int_0^T \exp\left(-\int_s^{s+\tau} c(u,x)\mathrm{d}u\right)\mathrm{d}\tau$. Similarly, we obtain $\lim_{r\to\infty}\tilde{E}_k|\tilde{\sigma}_k^{m_r}-\tilde{\sigma}_k| = 0$ for every $k\geq 1$. From this and the properties of the sequence $(\tilde{Y}_k^{m_r})_{r=1}^{\infty}$, one gets

$$\left|\tilde{E}_k\left[\exp\left(-\int_0^{t\wedge\tilde{\sigma}_k} c(\tilde{Y}_k(\tau))\mathrm{d}\tau\right) v(t_0-t\wedge\tilde{\sigma}_k,\tilde{Y}_k(t\wedge\tilde{\sigma}_k))\right] - v(t_0-\tau,s,x)\right| \leq \frac{M}{k}$$

for every $k\geq 1$. Similarly as above, by the properties of processes $\tilde{Y}_{k_r}$ and $\hat{Y}_{k_r}$, it follows that

$$\left|\hat{E}\left[\exp\left(-\int_0^{t\wedge\hat{\sigma}_{k_r}^{\lambda}} c(\hat{Y}_{k_r}(\tau))\mathrm{d}\tau\right) v(t_0-t\wedge\hat{\sigma}_{k_r}^{\lambda},\hat{Y}_{k_r}(t\wedge\hat{\sigma}_{k_r}^{\lambda}))\right] - v(t_0-\tau,s,x)\right| \leq \frac{M}{k_r}$$

for every $r\geq 1$, which implies

$$\hat{E}\left[\exp\left(-\int_0^{t\wedge\hat{\sigma}_{\lambda}} c(\hat{Y}(\tau))\mathrm{d}\tau\right) v(t_0-t\wedge\hat{\sigma}_{\lambda},\hat{Y}(t\wedge\hat{\sigma}_{\lambda}))\right] = v(t_0-\tau,s,x)$$

for $(s,x)\in[0,T]\times\mathbb{R}^d$, $t\in[0,t_0)$, and $\lambda>0$. Let us observe that

$$\lim_{\lambda\to\infty}\int_{\{\hat{\sigma}_\lambda\leq t\}} v(t-\hat{\sigma}_\lambda,\hat{Y}(\hat{\sigma}_\lambda))\mathrm{d}\hat{P} = 0$$

and

$$v(t_0,s,x) = \int_{\{\hat{\sigma}_\lambda\leq t\}}\left[\exp\left(-\int_0^{\hat{\sigma}_\lambda} c(\hat{Y}(\tau))\mathrm{d}\tau\right) v(t_0-\hat{\sigma}_\lambda,\hat{Y}(\hat{\sigma}_\lambda))\right]\mathrm{d}\hat{P}$$
$$+\int_{\{\hat{\sigma}_\lambda> t\}}\left[\exp\left(-\int_0^{t} c(\hat{Y}(\tau))\mathrm{d}\tau\right) v(t_0-t,\hat{Y}(t))\right]\mathrm{d}\hat{P}\,.$$

Therefore,

$$\hat{E}\left[\exp\left(-\int_0^{t} c(\hat{Y}(\tau))\mathrm{d}\tau\right) v(t_0-t,\hat{Y}(t))\right] = v(t_0,s,x)$$

for $(s,x) \in [0,T]\times\mathbb{R}^d$, $t \in [0,t_0)$. Passing to the limit in the last equality as $t_0 \to t$, we get

$$v(t,s,x) = \hat{E}\left[\exp\left(-\int_0^t c(\hat{Y}(\tau)\mathrm{d}\tau\right) v(0,\hat{Y}(t))\right],$$

which by virtue of (6.3), can be written in the form

$$v(t,s,x) = \hat{E}\left[\exp\left(-\int_s^{s+t} c(\tau,\hat{X}_{s,x}(\tau))\mathrm{d}\tau\right)\tilde{h}(s+t,\hat{X}_{s,x}(s+t))\right]$$

for $(s,x) \in [0,T)\times\mathbb{R}^d$ and $t \in [0,T-s]$. □

Theorem 6.5. *Assume that conditions (i), (iii), (iv′), and (v) of* $(\mathcal{A})$ *are satisfied. Let* $T > 0$*, let* D *be a bounded domain in* $\mathbb{R}^d$*, and let* $\Phi \in C((0,T)\times\partial D,\mathbb{R})$*. Assume that* $u \in C((0,T)\times D,\mathbb{R})$ *is bounded. If* $v \in C_0^{1,2}(\mathbb{R}^{d+1})$ *is bounded such that*

$$\begin{cases} u(t,x) - v_t'(t,x) \in (\mathbb{L}_{FG}v)(t,x) \ \text{ for } \ (t,x) \in (0,T)\times D, \\ \lim_{D\ni x\to y} v(t,x) = \Phi(t,y) \ \text{ for } (t,y) \in (0,T]\times\partial D, \end{cases} \tag{6.6}$$

then for every $(s,x) \in [0,T)\times D$*, there exists* $\hat{X}_{s,x} \in \mathcal{X}_{s,x}(F,G)$ *defined on the probability space* $(\hat{\Omega},\hat{\mathcal{F}},\hat{P})$ *such that*

$$v(s,x) = \hat{E}[\Phi(\hat{\tau}_D,\hat{X}_{s,x}(\hat{\tau}_D))] - \hat{E}\left[\int_s^{\hat{\tau}_D} u(t,\hat{X}_{s,x}(t))\mathrm{d}t\right]$$

for $(s,x) \in [0,T)\times D$*, where* $\hat{\tau}_D = \inf\{r \in (s,T] : \hat{X}_{s,x}(r) \notin D\}$.

Proof. By virtue of Theorem 2.6, for every $k,m \geq 1$, there are $f_k \in \mathcal{C}(F)$ and $g_k^m \in \mathcal{C}(G)$ such that $\sigma_k^m = g_k^m\cdot(g_k^m)^*$ is uniformly positive definite and

$$\lim_{m\to\infty} |u(t,x) - v_t'(t,x) - (\mathbb{L}_{f_k g_k^m} v(t,\cdot))(t,x)| \leq 1/k \tag{6.7}$$

uniformly with respect to $(t,x) \in (0,T)\times D$. Similarly as in the proof of Theorem 6.4, we can verify that for every $(s,x) \in [0,T)\times D$ and $k,m \geq 1$, there is a unique in law weak solution $(\mathcal{P}_k^m, X_k^m, B_k^m)$ of the stochastic differential equation $\mathrm{d}x_t = f_k(t,x_t)\mathrm{d}t + g_k^m(t,x_t)\mathrm{d}B_t$ with initial condition $x_s = x$ a.s. such that the process $Y_{s,x}^k = (s+t, X_{s,x}^k(s+t))_{0\leq t\leq T-s}$ is an Itô diffusion with infinitesimal generator $\mathcal{A}_{f_k g_k^m}$ satisfying $(\mathcal{A}_{f_k g_k^m} v)(t,x) = v_t'(t,x) + (\mathbb{L}_{f_k g_k^m} v)(t,x)$ for $(t,x) \in [0,T]\times D$ and $v \in C_0^{1,2}(\mathbb{R}^{d+1})$. Then (6.7) can be written as the inequality

$$\lim_{m\to\infty} |u(t,x) - (\mathcal{A}_{f_k g_k^m} v)(t,x)| \leq 1/k, \tag{6.8}$$

which has to be satisfied uniformly with respect to $(t,x) \in (0,T)\times D$. By the weak compactness of the set $\mathcal{X}_{s,x}(F,G)$ of (equivalence classes of) weak solutions of $SFI(F,G)$ with initial condition $x_s = x$ a.s., we can select an increasing subsequence $(m_r)_{r=1}^{\infty}$ of the sequence $(m)_{m=1}^{\infty}$, a probability space $\tilde{\mathcal{P}}_k = (\tilde{\Omega}_k, \tilde{\mathcal{F}}_k, \tilde{P}_k)$, and continuous processes $\tilde{X}_k^{m_r}$, $\tilde{X}_k$ on $\tilde{\mathcal{P}}_k$ such that $P(X_k^{m_r})^{-1} = P(\tilde{X}_k^{m_r})^{-1}$ for $k, r \geq 1$ and such that $\lim_{r\to\infty} \sup_{s\leq t\leq T} |\tilde{X}_k^{m_r}(t) - \tilde{X}_k(t)| = 0$. By virtue of Theorem 5.1 of Chap. 4, it follows that $\lim_{r\to\infty} |\tilde{\tau}_k^{m_r} - \tilde{\tau}_k| = 0$ $\tilde{P}_k$-a.s., where $\tilde{\tau}_k^{m_r}$ and $\tilde{\tau}_k$ denote the first exit times of $\tilde{X}_k^{m_r}$ and $\tilde{X}_k$ from the domain D, respectively.

Similarly as above, we can verify that $\tilde{X}_k^{m_r}$ is a weak solution of $SDE(f_k, g_k^{m_r})$, because $P_k^{m_r}(\{X_k^{m_r}(s) = x\}) = \tilde{P}_k(\{\tilde{X}_k^{m_r}(s) = x\})$ for $k, m = 1, 2, \ldots$, where $\mathcal{P}_k^{m_r} = (\Omega_k^{m_r}, \mathcal{F}_k^{m_r}, \mathbb{F}_k^{m_r}, P_k^{m_r})$ is a filtered probability space such that $(\mathcal{P}_k^{m_r}, X_k^{m_r}, B_k^{m_r})$ is a weak solution of $SDE(f_k, g_k^{m_r})$ with initial condition $x_s = x$ a.s. Then $\tilde{X}_k^{m_r} \in \mathcal{X}_{s,x}(F, G)$, which by the weak compactness of $\mathcal{X}_{s,x}(F, G)$ implies that also $\tilde{X}_k \in \mathcal{X}_{s,x}(F, G)$ for every $k \geq 1$. Similarly as above, by the weak compactness of the set $\mathcal{X}_{s,x}(F, G)$, there are an increasing subsequence $(k_r)_{r=1}^{\infty}$ of the sequence $(k)_{k=1}^{\infty}$, a probability space $\hat{\mathcal{P}} = (\hat{\Omega}, \hat{\mathcal{F}}, \hat{P})$, and continuous stochastic processes $\hat{X}_{k_r}$ and $\hat{X}$ on $\hat{\mathcal{P}}$ such that $P(\tilde{X}_{k_r})^{-1} = P(\hat{X}_{k_r})^{-1}$ for $r = 1, 2, \ldots$, $\lim_{r\to\infty} \sup_{s\leq t\leq T} |\hat{X}_{k_r}(t) - \hat{X}(t)| = 0$ $\hat{P}$-a.s., and $\lim_{r\to\infty} \hat{E}|\hat{\tau}_{k_r} - \hat{\tau}| = 0$, where $\hat{\tau}_{k_r}$ and $\hat{\tau}$ denote the first exit times of $\hat{X}_{k_r}$ and $\hat{X}$ from the domain D, respectively.

Denote by $Y_k^{m_r}, \tilde{Y}_k^{m_r}, \tilde{Y}_k, \hat{Y}_k$, and $\hat{Y}$ the Itô diffusions processes defined, similarly as in the proof of Theorem 6.4, by $X_k^{m_r}, \tilde{X}_k^{m_r}, \tilde{X}_k$, and $\hat{X}$, respectively, and let $\sigma_k^{m_r}$, $\tilde{\sigma}_k^{m_r}$, $\tilde{\sigma}_k$, and $\hat{\sigma}$ be their first exit times, respectively, from the domain $(0, T) \times D$. We have $\sigma_k^{m_r} = \tau_k^{m_r} - s$ $\tilde{\sigma}_k^{m_r} = \tilde{\tau}_k^{m_r} - s$ $\tilde{\sigma}_k = \tilde{\tau}_k - s$ and $\hat{\sigma} = \hat{\tau} - s$ for $k \geq 1$, where $\tau_k^{m_r}, \tilde{\tau}_k^{m_r}, \tilde{\tau}_k$, and $\hat{\tau}$ denote the first exit times of $X_k^{m_r}, \tilde{X}_k^{m_r}, \tilde{X}_k$, and $\hat{X}$, respectively from the domain D. By Dynkin's formula, for every $k = 1, 2, \ldots$, we obtain

$$E_k^{m_r}\left[v(Y_k^{m_r}(\sigma_k^{m_r}))\right] = v(s,x) + E_k^{m_r}\left[\int_0^{\sigma^{m_r}} (\mathcal{A}_{f_k g_k^{m_r}} v)(Y_k^{m_r}(t))\mathrm{d}t\right].$$

By the definition of $Y_k^{m_r}$ and the equality $\sigma_k^{m_r} = \tau_k^{m_r} - s$, the last relation can be written in the form

$$v(s,x) = E_k^{m_r}\left[v(\tau_k^{m_r}, X_k^{m_r}(\tau_k^{m_r}))\right] - E_k^{m_r}\left[\int_s^{\tau_k^{m_r}} (\mathcal{A}_{f_k g_k^{m_r}} v)(t, X_k^{m_r}(t))\mathrm{d}t\right].$$

Hence, by virtue of Theorem 5.2 of Chap. 4 and Corollary 5.1 of Chap. 4, it follows that

$$v(s,x) = \tilde{E}_k\left[v(\tilde{\tau}_k^{m_r}, \tilde{X}_k^{m_r}(\tilde{\tau}_k^{m_r}))\right] - \tilde{E}\left[\int_s^{\tilde{\tau}_k^{m_r}} (\mathcal{A}_{f_k g_k^{m_r}} v)(t, \tilde{X}_k^{m_r}(t))\mathrm{d}t\right].$$

Let $u_k^{m_r}(t,x) = (\mathcal{A}_{f_k g_k^{m_r}} v)(t,x)$ for $(t,x) \in (0,T) \times D$ and $k, m = 1, 2, \ldots$. By virtue of (6.8), we get $\lim_{r\to\infty} \left|u(t, \tilde{X}_k^{m_r}(t)) - u_k^{m_r}(t, \tilde{X}_k^{m_r}(t))\right| \leq 1/n_k$ a.s.

uniformly with respect to $0 \le t \le T$. Therefore,

$$\lim_{r\to\infty}\left|v(s,x) - \tilde{E}_k\left[v(\tilde{\tau}_k^{m_r}, \tilde{X}_k^{m_r}(\tilde{\tau}_k^{m_r}))\right] + \tilde{E}_k\left[\int_s^{\tilde{\tau}_k^{m_r}} u(t, \tilde{X}_k^{m_r}(t))\mathrm{d}t\right]\right|$$

$$\le \lim_{r\to\infty} \tilde{E}_k\left[\int_s^{\tilde{\tau}_k^{m_r}} |u_k^{m_r}(t, \tilde{X}_k^{m_r}(t)) - u(t, \tilde{X}_k^{m_r}(t))|\mathrm{d}t\right] \le \frac{T}{k}$$

for $k \ge 1$ uniformly with respect to $t \in [0, T)$. Hence it follows that

$$\left|v(s,x) - \tilde{E}_k[v(\tilde{\tau}_k, \tilde{X}_k(\tilde{\tau}_k))] + \tilde{E}_k\left[\int_s^{\tilde{\tau}_k} u(t, \tilde{X}_k(t))\mathrm{d}t\right]\right| \le \frac{T}{k}.$$

By Theorem 5.2 of Chap. 4, Corollary 5.1 of Chap. 4, and the properties of sequences $(\tilde{X}_k)_{k=1}^{\infty}$ and $(k_r)_{r=1}^{\infty}$, one obtains

$$\left|v(s,x) - \hat{E}[v(\hat{\tau}_{k_r}, \hat{X}_{k_r}(\hat{\tau}_{k_r}))] + \hat{E}\left[\int_s^{\hat{\tau}_{k_r}} u(t, \tilde{X}_{k_r}(t))\mathrm{d}t\right]\right| \le \frac{T}{k_r},$$

for every $r \ge 1$ and $t \in [0, T)$. Therefore,

$$\lim_{r\to\infty}\left|v(s,x) - \hat{E}[v(\hat{\tau}_{k_r}, \hat{X}_{k_r}(\hat{\tau}_{k_r}))] + \hat{E}\left[\int_s^{\hat{\tau}_{k_r}} u(t, \tilde{X}_{k_r}(t))\mathrm{d}t\right]\right| = 0,$$

which implies that

$$v(s,x) = \hat{E}[\Phi(\hat{\tau}_D, \hat{X}_{s,x}(\hat{\tau}_D))] - \hat{E}\left[\int_s^{\hat{\tau}_D} u(t, \hat{X}_{s,x}(t))\mathrm{d}t\right]$$

for $(s,x) \in [0,T) \times D$, where $\hat{\tau}_D = \inf\{r \in (s,T] : \hat{X}_{s,x}(r) \notin D\}$. □

Quite similarly, we obtain the following result.

Theorem 6.6. *Assume that conditions (i), (iii), (iv′), and (v) of* $(\mathcal{A})$ *are satisfied. Let* $T > 0$ *and let* D *be a bounded domain in* $\mathbb{R}^d$. *Assume that* $\Phi \in C((0,T) \times \partial D, \mathbb{R})$, $c \in C([0,T] \times D, \mathbb{R})$, *and* $u \in C((0,T) \times D, \mathbb{R})$ *are bounded. A bounded function* $v \in C_0^{1,2}(\mathbb{R}^{d+1})$ *is a solution of the boundary problem*

$$\begin{cases} u(t,x) - v_t'(t,x) \in (\mathbb{L}_{FG}v)(t,x) - c(t,x)v(t,x) \ \text{ for } \ (t,x) \in (0,T) \times D, \\ \lim_{D \ni x \to y} v(t,x) = \Phi(t,y) \ \text{ for } \ (s,y) \in (0,T] \times \partial D, \end{cases}$$

if and only if for every $(s,x) \in [0,T) \times D$, *there exists* $\tilde{X} \in \mathcal{X}_{s,x}(F,G)$ *defined on the probability space* $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ *such that*

$$v(s,x) = \tilde{E}\left[\Phi(\tilde{\tau}_D, \tilde{X}(\tilde{\tau}_D)) \exp\left(-\int_s^{\tilde{\tau}_D} c(t, \tilde{X}(t))\mathrm{d}t\right)\right]$$
$$- \tilde{E}\left\{\int_s^{\tilde{\tau}_D}\left[u(t, \tilde{X}(t)) \exp\left(-\int_s^{s+t} c(z, \tilde{X}(z))\mathrm{d}z\right)\right]\mathrm{d}t\right\}$$

for $(s,x) \in [0,T) \times D$, *where* $\tilde{\tau}_D = \inf\{r \in (s,T] : \tilde{X}(r) \notin D\}$. □

The following results follow immediately from Theorems 5.1 and 5.3.

Theorem 6.7. *Assume that conditions (i), (iii), (iv′), and (v) of* $(\mathcal{A})$ *are satisfied. Let* $T > 0$, *let* D *be a bounded domain in* $\mathbb{R}^d$, *and let* $\Phi : (0,T) \times \partial D \to \mathbb{R}$ *be measurable and bounded. For every* $(s,x) \in (0,T) \times D$, *there exists a weak solution* $(\mathcal{P}_{\mathbb{F}}, X_{s,x}, B)$ *of* $SFI(F,G)$ *satisfying the initial condition* $X_{s,x}(s) = x$ *a.s. such that the function* $u(s,x) = E^{s,x}[\Phi(\tau_D, X_{s,x}(\tau_D))]$ *is a solution of the set-valued stochastic Dirichlet problem*

$$\begin{cases} 0 \in (\mathcal{L}_{FG}u)(t,x) \text{ for } (t,x) \in [0,T) \times D, \\ \lim_{t\to\tau_D} u(t, X_{s,x}(t) = \Phi(\tau_D, X_{s,x}(\tau_D)) \text{ for } (s,x) \in (0,T) \times \partial D, \end{cases}$$

where $\tau_D = \inf\{r \in (s,T) : X_{s,x}(r) \notin D\}$.

Proof. By the properties of F and G, we can select $f \in \mathcal{C}(F)$ and $g \in \mathcal{C}(G)$ such that for every $(s,x) \in (0,T) \times D$, there is a unique in law weak solution of $SDE(f,g)$ satisfying initial condition $X_{s,x}(s) = x$ a.s. By virtue of Theorem 5.1, the function $u(s,x) = E^{s,x}[\Phi(\tau_D, X_{s,x}(\tau_D))]$ is a solution of the stochastic Dirichlet problem (5.3), where $\mathcal{L}_{fg}$ is the characteristic operator of the Itô diffusion $Y_{s,x} = (Y_{s,x}(t))_{t\geq 0}$ defined by $Y_{s,x}(t) = (s+t, X_{s,x}(s+t))$ for fixed $0 < s < T$ and $t \in [0, T-s]$. By the definition of $\mathcal{L}_{FG}$, it follows that $(\mathcal{L}_{fg}u)(s,x) \in (\mathcal{L}_{FG}u)(s,x)$ for every $(s,x) \in (0,T) \times D$, which proves that u is a solution of the above set-valued stochastic Dirichlet problem. □

Theorem 6.8. *Assume that conditions (i), (iii), (iv′), and (v) of* $(\mathcal{A})$ *are satisfied. Let* $T > 0$, *let* D *be a bounded domain in* $\mathbb{R}^d$, *and let* $\varphi : (0,T) \times \partial D \to \mathbb{R}$ *be continuous and bounded. For every* $(s,x) \in (0,T) \times D$, *there exist a weak solution* $(\mathcal{P}_{\mathbb{F}}, X_{s,x}, B)$ *of* $SFI(F,G)$ *satisfying the initial condition* $X_{s,x}(s) = x$ *a.s. such that the function* $v(s,x) = E^{s,x}\left[\int_0^{\tau_D} \varphi(\tau_D, X_{s,x}(\tau_D))\right]$ *is a solution of the set-valued stochastic Poisson problem*

$$\begin{cases} -\varphi(s,x) \in (\mathcal{L}_{FG}v)(t,x) \text{ for } (t,x) \in [0,T) \times D, \\ \lim_{t\to\tau_D} v(t, X_{s,x}(t)) = 0 \text{ for } (s,x) \in (0,T) \times \partial D, \end{cases}$$

where $\tau_D = \inf\{r \in (s,T) : X_{s,x}(r) \notin D\}$.

Proof. The proof follows immediately from Theorem 5.3, similarly as in the proof of Theorem 6.7, because by the properties of the function φ and the boundedness of the domain D, we have $E^{s,x}\left[\int_0^{\tau_D}\varphi(t,X_{s,x}(t))\mathrm{d}t\right] \le M \cdot E^{s,x}[\tau_D] < \infty$, where $M > 0$ is such that $|\varphi(t,x)| \le M$ for $(t,x) \in (0,T) \times D$. □

Theorem 6.9. *Assume that conditions (i), (iii), (iv′), and (v) of ($\mathcal{A}$) are satisfied. Let $T > 0$ and let D be a bounded domain in $\mathbb{R}^d$. Assume that $\Phi : (0,T) \times \partial D \to \mathbb{R}$ is measurable and bounded and that $\varphi : (0,T) \times \partial D \to \mathbb{R}$ is continuous and bounded. For every $(s,x) \in (0,T) \times D$, there is a weak solution $(\mathcal{P}_{\mathbb{F}}, X_{s,x}, B)$ of $SFI(F,G)$ satisfying the initial condition $X_{s,x}(s) = x$ a.s. such that the function*

$$w(s,x) = E^{s,x}[\Phi(\tau_D, X_{s,x}(\tau_D))] + E^{s,x}\left[\int_0^{\tau_D}\varphi(t,X_{s,x}(t))\mathrm{d}t\right]$$

is a solution of the set-valued stochastic Dirichlet–Poisson problem

$$\begin{cases} -\varphi(s,x) \in (\mathcal{L}_{FG}w)(t,x) \ \text{ for } \ (t,x) \in [0,T) \times D, \\ \lim_{t\to\tau_D} w(t,X_{s,x}(t)) = \Phi(\tau_D, X_{s,x}(\tau_D)) \ \text{ for } \ (s,x) \in (0,T) \times \partial D\ , \quad a.s., \end{cases}$$

where $\tau_D = \inf\{r \in (s,T) : X_{s,x}(r) \notin D\}$.

Proof. Similarly as in the proof of Theorem 6.7, we can select $f \in \mathcal{C}(F)$ and $g \in \mathcal{C}(G)$ such that for every $(s,x) \in (0,T) \times D$, there is a unique in law weak solution of $SDE(f,g)$ satisfying initial condition $X_{s,x}(s) = x$ a.s. Let u and v be defined as above, i.e., let $u(s,x) = E^{s,x}[\Phi(\tau_D, X_{s,x}(\tau_D)]$ and $v(s,x) = E^{s,x}\left[\int_0^{\tau_D}\varphi(t,X_{s,x}(t))\mathrm{d}t\right]$ for $(s,x) \in (0,T) \times D$. By virtue of Theorems 5.1 and 5.3, one has

$$\begin{cases} \left(\mathcal{L}_{fg}u\right)(t,x) = 0 \ \text{ for } \ (t,x) \in [0,T) \times D, \\ \lim_{t\to\tau_D} u(t,x) = \Phi(\tau_D, X_{s,x}^{t,g}(\tau_D)) \ \text{ for } \ (s,x) \in (0,T) \times D \quad a.s., \end{cases}$$

and

$$\begin{cases} \left(\mathcal{L}_{fg}v\right)(t,x) = -\varphi(s,x) \ \ for \ \ (t,x) \in [0,T) \times D, \\ \lim_{t\to\tau_D} v(t, X_{s,x}^{fg}(t)) = 0 \ \ a.s. \ \ for \ \ (s,x) \in (0,T) \times D, \end{cases}$$

where $\tau_D = \inf\{r \in (s,T) : X_{s,x}(r) \notin D\}$. By the definition of $\mathcal{L}_{fg}$, it follows that $\left(\mathcal{L}_{fg}(u+v)\right)(s,x) = (\mathcal{L}_{fg}u)(s,x) + (\mathcal{L}_{fg}v)(s,x)$ for every $(s,x) \in (0,T) \times D$. Then the function $w = u + v$ satisfies $(\mathcal{L}_{fg}w)(s,x) = -\varphi(s,x)$ for $(s,x) \in (0,T) \times D$. Quite similarly, we obtain that $\lim_{t\to\tau_D} w(t,x) = \Phi(\tau_D, X_{s,x}^{t,g}(\tau_D))$ for $(s,x) \in (0,T) \times D$ a.s. Hence, similarly as in the proof of Theorem 5.4, it follows that the function w is a solution of the Dirichlet–Poisson problem. □

7 Notes and Remarks

The results of this chapter dealing with partial differential equations are based on A. Friedman [31] and B. Øksendal [86]. In particular, Theorem 3.1, Remark 3.2, and Theorem 3.2 are taken from A. Friedman [31]. Their proofs can be found in A. Friedman's book dealing with partial differential equations. All extensions of the above results to the case of partial differential inclusions are due to Kisielewicz [60,61]. The extensions of Feynman–Kac formulas to the set-valued case have been published in Kisielewicz [60]. Some results dealing with the Cauchy and Dirichlet set-valued problems are contained in Kisielewicz [61]. The stochastic characteristics of solutions of partial differential inclusions given above are good enough for solving some optimal control problems for systems described by some partial differential equations. This is a consequence of the weak compactness of the set $\mathcal{X}_{s,x}(F,G)$ of (equivalence classes of) all weak solutions of the stochastic functional inclusion $SFI(F,G)$ satisfying initial condition $x_s = x$ a.s. This property of the set $\mathcal{X}_{s,x}(F,G)$ is the basic one for solving some optimal control problems for systems described by stochastic functional and partial differential inclusions. Such optimal control problems are considered in the next chapter of the book. Example 5.1 and Fig. 6.1 are taken from B. Øksendal [86].

Chapter 7
Stochastic Optimal Control Problems

This chapter contains some optimal control problems for systems described by stochastic functional and partial differential inclusions. The existence of optimal controls and optimal solutions for such systems is a consequence of the weak compactness of the set $\mathcal{X}_{sx}(F, G)$ of all weak solutions of (equivalence classes of) $SFI(F, G)$ satisfying an initial condition $x_s = x$, measurable selection theorems, and stochastic representation theorems for solutions of partial differential inclusions presented in Chap. 6. We begin with introductory remarks dealing with optimal control problems of systems described by stochastic differential equations.

1 Optimal Control Problems for Systems Described by Stochastic Differential Equations

Assume that the state of a dynamical system starting from a point $(s, x) \in \mathbb{R}^+ \times \mathbb{R}^d$ is described at time $t \geq s$ by a weak solution of the stochastic differential equation

$$\begin{cases} \mathrm{d}x_t = f(t, x_t, u_t)\mathrm{d}t + g(t, x_t, u_t)\mathrm{d}B_t \quad \text{a.s. for } t \geq s\,, \\ x_s = x \text{ a.s.}, \end{cases} \tag{1.1}$$

depending on a control process $u = (u_t)_{t\geq 0}$, where $f : \mathbb{R}^+ \times \mathbb{R}^d \times U \to \mathbb{R}^d$ and $g : \mathbb{R}^+ \times \mathbb{R}^d \times U \to \mathbb{R}^{d\times m}$ are given functions with $U \subset \mathbb{R}^k$. Given a domain $D \subset \mathbb{R}^d$ and an initial point $(s, x) \in \mathbb{R}^+ \times D$, a system $(\mathcal{P}_{\mathbb{F}}, u, X_{s,x}, B)$ consisting of a filtered probability space $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$, $\mathbb{F}$-nonanticipative processes u and $X_{s,x}$, and an m-dimensional $\mathbb{F}$-Brownian motion $B = (B_t)_{t\geq 0}$ defined on $\mathcal{P}_{\mathbb{F}}$ satisfying (1.1) and such that $\tau_D^X < \infty$ a.s. is called an admissible system for the stochastic control system described by (1.1). As usual, τ_D^X denotes the first exit time of $X_{s,x}$ from the set D. For every $(s, x) \in \mathbb{R}^+ \times D$, we are also given a performance

M. Kisielewicz, *Stochastic Differential Inclusions and Applications*,
Springer Optimization and Its Applications 80, DOI 10.1007/978-1-4614-6756-4_7,

functional $J_D^{u,X}(s,x)$ defined for given functions $\Phi : \mathbb{R}^+ \times \mathbb{R}^d \times U \to \mathbb{R}$ and $K : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}$ and an admissible system $(\mathcal{P}_{\mathbb{F}}, u, X_{s,x}, B)$ by the formula

$$J_D^{u,X}(s,x) = E^{s,x}\left[\int_s^{\tau_D^X} \Phi(t, X_{s,x}(t), u(t))\mathrm{d}t + K\left(\tau_D^X, X_{s,x}\left(\tau_D^X\right)\right)\right], \qquad (1.2)$$

where $E^{s,x}$ denotes the mean value operator with respect to the law $Q^{s,x}$ of $X_{s,x}$. For every admissible system $(\mathcal{P}_{\mathbb{F}}, u, X_{s,x}, B)$, a pair $(u, X_{s,x})$ is said to be an admissible pair for (1.1). The set of all admissible pairs for the control system (1.1) is denoted by $\Lambda_{fg}(s,x)$. For every $(u, X_{s,x}) \in \Lambda_{fg}(s,x)$, a process $X_{s,x}$ is called an admissible trajectory corresponding to an admissible control u. The performance functional $J_D^{u,X}(s,x)$ can be regarded as a functional defined on the set $\Lambda_{fg}(s,x)$.

An admissible pair $(\bar{u}, \bar{X}_{s,x}) \in \Lambda_{fg}(s,x)$ is said to be optimal for an optimal control problem (1.1) and (1.2) if $J_D^{\bar{u},\bar{X}}(s,x) = \sup\{J_D^{u,X}(s,x) : (u, X_{s,x}) \in \Lambda_{fg}(s,x)\}$ for every $(s,x) \in \mathbb{R}^+ \times D$. If $(\bar{u}, \bar{X}_{s,x})$ is the optimal pair for (1.1) and (1.2), then $\bar{u}$ is called the optimal control, and $\bar{X}_{s,x}$ the optimal trajectory for the optimal control problem described by (1.1) and (1.2). The function $v : \mathbb{R}^+ \times D \to \mathbb{R}$ defined by $v(s,x) = \sup\{J_D^{u,X}(s,x) : (u, X_{s,x}) \in \Lambda_{fg}(s,x)\}$ for every $(s,x) \in \mathbb{R}^+ \times D$ is said to be the value function associated to the optimal control problem (1.1) and (1.2). An admissible pair $(\bar{u}, \bar{X}_{s,x})$ is optimal if $v(s,x) = J_D^{\bar{u},\bar{X}}(s,x)$ for every initial condition $(s,x) \in \mathbb{R}^+ \times D$. The problem consisting in finding for each $(s,x) \in \mathbb{R}^+ \times D$ the number $v(s,x)$ for the optimal control problem (1.1) and (1.2) will be denoted by

$$\begin{cases} \mathrm{d}x_t = f(t, x_t, u_t)\mathrm{d}t + g(t, x_t, u_t)\mathrm{d}B_t \quad \text{a.s. for } t \geq s\,, \\ x_s = x \quad \text{a.s.}\,, \\ J_D^{u,X}(s,x) \xrightarrow{\Lambda_{fg}} \max\,. \end{cases} \qquad (1.3)$$

Let us observe that if the optimal pair $(\bar{u}, \bar{X}_{s,x}) \in \Lambda_{fg}(s,x)$ exists and $(f(\cdot,\cdot,z), g(\cdot,\cdot,z))$ is such that $SDF(f(\cdot,\cdot,z), g(\cdot,\cdot,z))$ possesses for every fixed $z \in U$ a unique in law weak solution $X_{s,x}^z$ satisfying initial condition $X_{s,x}^z(s) = x$ a.s. for $(s,x) \in \mathbb{R}^+ \times \mathbb{R}^d$, then the standard approach to determine an optimal pair is to solve the Hamilton–Jacobi–Bellman (HJB) equation

$$\begin{cases} \sup_{z \in U}\{\Phi(s,x,z) + (\mathcal{A}_{fg}^z v)(s,x)\} = 0 \quad \text{for} \quad (s,x) \in \mathbb{R}^+ \times D\,, \\ v(s,x) = K(s,x) \quad \text{for} \quad (s,x) \in \mathbb{R}^+ \times \partial D, \end{cases}$$

where $\mathcal{A}_{fg}^z$ is the infinitesimal generator of a $(d+1)$-dimensional Itô diffusion defined, similarly as in Sect. 11 of Chap. 1, by $X_{s,x}^z$ for every fixed $z \in U$. If the above supremum is attained, i.e., if there exists an optimal control $\bar{u}(s,x)$, then

$$\begin{cases} \Phi(s,x,\bar{u}(s,x)) + (\mathcal{A}_{\bar{f}\bar{g}}v)(s,x) = 0 \quad \text{for} \quad (s,x) \in \mathbb{R}^+ \times D\,, \\ v(\bar{\tau}_D, \bar{X}_{s,x}) = K(\bar{\tau}_D, \bar{X}_{s,x}) \quad \text{for} \quad (s,x) \in \mathbb{R}^+ \times \partial D, \end{cases}$$

where $\bar{f}(s,x) = f(s,x,\bar{u}(s,x))$, $\bar{g}(s,x) = g(s,x,\bar{u}(s,x))$ for $(s,x) \in \mathbb{R}^+ \times D$, $\mathcal{A}_{\bar{f}\bar{g}}$ is an infinitesimal generator defined by a unique in law weak solution $\bar{X}_{s,x}$ of $SDE(\bar{f},\bar{g})$ satisfying an initial condition $\bar{X}_{s,x}(s) = x$ a.s. for $(s,x) \in \mathbb{R}^+ \times \mathbb{R}^d$, and $\bar{\tau}_D$ denotes the first exit time of $\bar{X}_{s,x}$ from the set D. Immediately from Theorem 5.5 of Chap. 6, it follows that if $v \in C_0^{1,2}([0,T] \times D, \mathbb{R})$ and $\bar{X}_{s,x}$ is such that $\bar{E}^{s,x}[\int_0^{\bar{\tau}_D} \Phi(t,\bar{X}_{s,x}(t))\mathrm{d}t] < \infty$ and there exists a number $C > 0$ such that $|v(t,x)| \le C(1 + \bar{E}^{s,x}[\int_0^{\bar{\tau}_D} \Phi(t,\bar{X}_{s,x}(t))\mathrm{d}t])$ for every $(s,x) \in (0,T) \times \mathbb{R}^d$, then

$$v(s,x) = \bar{E}^{s,x}[K(\bar{\tau}_D, \bar{X}_{s,x})] + \bar{E}^{s,x}\left[\int_0^{\bar{\tau}_D} \Phi(t,\bar{X}_{s,x}(t))\mathrm{d}t\right],$$

where $\bar{E}^{s,x}$ is a mean value operator taken with respect to a distribution of $\bar{X}_{s,x}$.

We shall consider now the optimal control problem (1.3) with continuous deterministic control parameters with values in a closed set $U \subset \mathbb{R}^k$ and a strong solution $X_{s,x}$ of (1.1) defined for a given m-dimensional $\mathbb{F}$-Brownian motion $B = (B_t)_{t\ge0}$ on a given complete filtered probability space $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\ge0}$ satisfying the usual conditions. We consider a control system (1.1) with measurable functions $f : \mathbb{R}^+ \times \mathbb{R}^d \times U \to \mathbb{R}^d$ and $g : \mathbb{R}^+ \times \mathbb{R}^d \times U \to \mathbb{R}^{d\times m}$ satisfying the following conditions (H).

(H): There exist $k, m \in \mathbb{L}(\mathbb{R}^+, \mathbb{R}^+)$ such that

(i) $\max(|f(t,x,z)|, \|g(t,x,z)\|) \le m(t)$ for every $(t,x,z) \in \mathbb{R}^+ \times \mathbb{R}^d \times U$.
(ii) $\max(|f(t,x,z) - f(t,\bar{x},\bar{z})|^2, \|g(t,x,z) - g(t,\bar{x},\bar{z})\|^2) \le k(t)(|x-\bar{x}|^2 + |z-\bar{z}|^2)$ for every $t \ge 0$, $x,\bar{x} \in \mathbb{R}^d$, and $z,\bar{z} \in U$.
(iii) $g(t,x,z) \cdot g(t,x,z)^*$ is positive definite on $\mathbb{R}^+ \times \mathbb{R}^d$ for every fixed $z \in U$.

In what follows, by $\mathcal{U}_T$ we denote a nonempty compact subset of the Banach space $(C([0,T],\mathbb{R}^k), \|\cdot\|_T)$ with the supremum norm $\|\cdot\|_T$ such that $u_t \in U$ for every $u \in \mathcal{U}_T$ and $t \in [0,T]$.

Remark 1.1. Similarly as in the proof of Theorem 1.1 of Chap. 4, by an appropriate changing of the norm of the space $\mathcal{X}$ defined in the proof of Theorem 1.1 of Chap. 4, we can verify that if conditions (i) and (ii) of (H) are satisfied, then for every $(s,x) \in \mathbb{R}^+ \times \mathbb{R}^d$, $T > s$, a filtered probability space $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$, an m-dimensional $\mathbb{F}$-Brownian motion $B = (B_t)_{t\ge0}$, and $u \in \mathcal{U}_T$, there exists a unique strong solution $X^u_{s,x}$ of (1.1) defined on $[s,T] \times \Omega$.

Proof. Let $(s,x) \in \mathbb{R}^+ \times \mathbb{R}^d$, $T > s$, a filtered probability space $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$, and an m-dimensional $\mathbb{F}$-Brownian motion $B = (B_t)_{t\ge0}$ be given. Define, for fixed $u \in \mathcal{U}_T$, set-valued mappings F and G by taking $F(t,x) = \{f(t,x,u_t)\}$ and $G(t,x) = \{g(t,x,u_t)\}$ for $(t,x) \in [0,T] \times \mathbb{R}^d$. Let $X^{\alpha\beta}_{s,x}(t)$ be defined by $X^{\alpha\beta}_{s,x}(t) = x + \int_s^t \alpha_\tau \mathrm{d}\tau + \int_s^t \beta_\tau \mathrm{d}B_\tau$ for every $t \in [s, T]$ and $(\alpha, \beta) \in \mathcal{X}$. Similarly as in the proof of Theorem 1.1 of Chap. 4, we define on $\mathcal{X}$ an operator Q, which in the case of the above-defined multifunctions F and G, has the form $Q(\alpha, \beta) = \{(f(\cdot, X^{\alpha\beta}_{s,x}, u), g(\cdot, X^{\alpha\beta}_{s,x}, u))\}$ for every $(\alpha, \beta) \in \mathcal{X}$.

Let us define on $\mathbb{L}^2([0,T] \times \Omega, \Sigma_{\mathbb{F}}, \mathbb{R}^d)$ a family $\{\|\cdot\|_\lambda\}_{\lambda>0}$ of norms $\|\cdot\|_\lambda$ equivalent to the norm $|\cdot|$ of this space by setting $\|w\|^2_\lambda = \int_0^T \exp[-lK(t)]E|w_t|^2\mathrm{d}t$ for $w \in \mathbb{L}^2([0,T] \times \Omega, \Sigma_{\mathbb{F}}, \mathbb{R}^d)$, where $l = 1/\lambda^2$ and $K(t) = \int_0^t k(\tau)\mathrm{d}\tau$ with $k \in \mathbb{L}(\mathbb{R}^+, \mathbb{R}^+)$ satisfying conditions (H). For every $(\alpha, \beta), (\tilde{\alpha}, \tilde{\beta}) \in \mathcal{X}$, one gets

$$\begin{aligned}
&\|f(\cdot, X^{\alpha\beta}_{s,x}, u) - f(\cdot, X^{\tilde{\alpha}\tilde{\beta}}_{s,x}, u)\|^2_\lambda \\
&\quad = \int_0^T \exp[-lK(t)]E|f(t, X^{\alpha\beta}_{s,x}(t), u_t) - f(t, X^{\tilde{\alpha}\tilde{\beta}}_{s,x}(t), u_t)|^2\mathrm{d}t \\
&\quad \le \int_0^T k(t)\exp[-lK(t)]E|X^{\alpha\beta}_{s,x}(t) - X^{\tilde{\alpha}\tilde{\beta}}_{s,x}(t)|^2\mathrm{d}t.
\end{aligned}$$

Similarly as in the proof of Theorem 1.1 of Chap. 4, we get

$$\begin{aligned}
E[|X^{\alpha\beta}_{s,x}(t) - X^{\tilde{\alpha}\tilde{\beta}}_{s,x}(t)|^2] &= E\left|\int_s^t (\alpha_\tau - \tilde{\alpha}_\tau)\mathrm{d}\tau + \int_s^t (\beta_\tau - \tilde{\beta}_\tau)\mathrm{d}B\tau\right|^2 \\
&\le 2T\int_0^t E|\alpha_\tau - \tilde{\alpha}_\tau|^2\mathrm{d}\tau + 2\int_0^t E|\beta_\tau - \tilde{\beta}_\tau|^2\mathrm{d}\tau.
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\|f(\cdot, X^{\alpha\beta}_{s,x}, u) - f(\cdot, X^{\tilde{\alpha}\tilde{\beta}}_{s,x}, u)\|^2_\lambda \\
&\quad \le 2T\int_0^T\int_0^t k(t)\exp[-lK(t)]E|\alpha_\tau - \tilde{\alpha}_\tau|^2\mathrm{d}\tau\mathrm{d}t \\
&\qquad + 2\int_0^T\int_0^t k(t)\exp[-lK(t)]E|\beta_\tau - \tilde{\beta}_\tau|^2\mathrm{d}\tau\mathrm{d}t.
\end{aligned}$$

By interchanging the order of integration, we obtain

$$\begin{aligned}
\int_0^T\int_0^t k(t)\exp[-lK(t)]E|\alpha_\tau-\tilde{\alpha}_\tau|^2\mathrm{d}\tau\mathrm{d}t &= \int_0^T\int_\tau^T E|\alpha_\tau-\tilde{\alpha}_\tau|^2 k(t)\exp[-lK(t)]\mathrm{d}t\mathrm{d}\tau \\
= -\lambda^2 e^{-lK(T)}\int_0^T E|\alpha_\tau - \tilde{\alpha}_\tau|^2\mathrm{d}\tau &+ \lambda^2\int_0^T k(\tau)\exp[-lK(\tau)]E|\alpha_\tau - \tilde{\alpha}_\tau|^2\mathrm{d}\tau \\
\le \lambda^2\,\|\alpha - \tilde{\alpha}\|^2_\lambda.
\end{aligned}$$

In a similar way, we obtain

$$\int_0^T\int_0^t k(t)\exp[-lK(t)]E|\beta_\tau - \tilde{\beta}_\tau|^2\mathrm{d}\tau\mathrm{d}t \le \lambda^2\,\|\beta - \tilde{\beta}\|_\lambda.$$

Therefore,

$$\|f(\cdot, X_{s,x}^{\alpha\beta}, u) - f(\cdot, X_{s,x}^{\tilde{\alpha}\tilde{\beta}}, u)\|_\lambda^2 \le 2\,\lambda^2(1+T)\,\|(\alpha,\beta) - (\tilde{\alpha} - \tilde{\beta}\|_\lambda^2\,,$$

where $\|(\alpha,\beta) - (\tilde{\alpha},\tilde{\beta})\|_\lambda = \max(\|\alpha - \tilde{\alpha}\|_\lambda, \|\beta - \tilde{\beta}\|_\lambda)$. In a similar way, for every $\lambda > 0$, we can define on the space $\mathbb{L}^2([0,T] \times \Omega, \Sigma_{\mathbb{F}}, \mathbb{R}^{d\times m})$ an equivalent norm, denoted again by $\|\cdot\|_\lambda$, and get

$$\|g(\cdot, X_{s,x}^{\alpha\beta}, u) - g(\cdot, X_{s,x}^{\tilde{\alpha}\tilde{\beta}}, u)\|_\lambda \le 2\,\lambda^2(1+T)\ \|(\alpha,\beta) - (\tilde{\alpha},\tilde{\beta})\|_\lambda.$$

Therefore, for every $\lambda > 0$ and $(\alpha,\beta), (\tilde{\alpha},\tilde{\beta}) \in \mathcal{X}$, one has

$$d_\lambda(Q(\alpha,\beta), Q(\tilde{\alpha},\tilde{\beta})) \le \lambda\,\sqrt{2(1+T)}\,\|(\alpha,\beta) - (\tilde{\alpha},\tilde{\beta})\|_\lambda,$$

where

$$\begin{aligned} & d_\lambda(Q(\alpha,\beta), Q(\tilde{\alpha},\tilde{\beta})) \\ &\quad = \max\{\|f(\cdot, X_{s,x}^{\alpha\beta}, u) - f(\cdot, X_{s,x}^{\tilde{\alpha}\tilde{\beta}}, u)\|_\lambda\,,\ \|g(\cdot, X_{s,x}^{\alpha\beta}, u) - g(\cdot, X_{s,x}^{\tilde{\alpha}\tilde{\beta}}, u)\|_\lambda\}. \end{aligned}$$

Taking in particular $\lambda \in (0, 1/\sqrt{2(1+T)})$, we obtain a contraction mapping Q defined on the complete metric space $(\mathcal{X}, d_\lambda)$. Then there exists a unique fixed point $(\alpha,\beta) \in \mathcal{X}$ of Q, which generates exactly one strong solution $X_{s,x}^{\alpha\beta}$ of (1.1) defined on $[s,T] \times \Omega$. □

Let $X_{s,x}^u$ be the unique strong solution of (1.1) defined for given $(s,x) \in \mathbb{R}^+ \times \mathbb{R}^d$, $T > s$, and $u \in \mathcal{U}_T$ on the interval $[s,T]$. We can extend such a solution to the whole interval $[0,T]$ by taking $X_{s,x}^u(t) = x$ a.s. for $0 \le t < s$ and define on $\mathcal{U}_T$ an operator $\lambda_{s,x}$ with values in $C_{\mathbb{F}}^T$ by setting $\lambda_{s,x}(u) = \tilde{X}_{s,x}^u$, where $\tilde{X}_{s,x}^u = \mathbb{I}_{[0,s)}x + \mathbb{I}_{[s,T]}X_{s,x}^u$ and $(C_{\mathbb{F}}^T, \|\cdot\|)$ denotes the space of all $\mathbb{F}$-adapted d-dimensional continuous square integrable stochastic processes $X = (X_t)_{0\le t\le T}$ with norm $\|X\| = \{E[\sup_{0\le t\le T}|X_t|^2]\}^{1/2}$.

Lemma 1.1. *Let $B = (B_t)_{t\ge 0}$ be an m-dimensional $\mathbb{F}$-Brownian motion on a filtered probability space $\mathcal{P}_{\mathbb{F}}$, $(s,x) \in \mathbb{R}^+ \times \mathbb{R}^d$, and $T > s$. If f and g are measurable and satisfy (i) and (ii) of conditions (H), then $\lambda_{s,x}$ is a continuous mapping on $\mathcal{U}_T$ depending continuously on $(s,x) \in \mathbb{R}^+ \times \mathbb{R}^d$.*

Proof. By virtue of Remark 1.1, for every $u \in \mathcal{U}_T$, there exists a unique strong solution of (1.1) defined on $[s,T] \times \Omega$. Let $u \in \mathcal{U}_T$, and let $(u_n)_{n=1}^\infty$ be a sequence of $\mathcal{U}_T$ such that $\|u_n - u\|_T \to 0$ as $n \to \infty$. By the definition of the mapping $\lambda_{s,x}$, we have $\lambda_{s,x}(u) = \tilde{X}_{s,x}$ and $\lambda_{s,x}(u_n) = \tilde{X}_{s,x}^n$ for $n = 1, 2, \ldots$. By Corollary 4.4 of Chap. 1, for every $n \ge 1$ and $s \le t \le T$, we get

$$
\begin{aligned}
E\left[\sup_{0\le z\le t}|\tilde{X}^n_{s,x}(z)-\tilde{X}_{s,x}(z)|^2\right] &= E\left[\sup_{s\le z\le t}|X^n_{s,x}(z)-X_{s,x}(z)|^2\right]\\
&\le 2E\left(\sup_{s\le z\le t}\left|\int_s^z [f(\tau,X^n_{s,x}(\tau),u^n_\tau)-f(\tau,X_{s,x}(\tau),u_\tau)]\mathrm{d}\tau\right|^2\right)\\
&\quad + 2E\left(\sup_{s\le z\le t}\left|\int_s^z [g(\tau,X^n_{s,x}(\tau),u^n_\tau)-g(\tau,X_{s,x}(\tau),u_\tau)]\mathrm{d}B_\tau\right|^2\right)\\
&\le 2TE\int_s^t |f(\tau,X^n_{s,x}(\tau),u^n_\tau)-f(\tau,X_{s,x}(\tau),u_\tau)|^2\mathrm{d}\tau\\
&\quad + 8E\int_s^t |g(\tau,X^n_{s,x}(\tau),u^n_\tau)-g(\tau,X_{s,x}(\tau),u_\tau)|^2\mathrm{d}\tau\\
&\le 2(T+4)\,\|u^n-u\|_T^2\int_0^T k(t)\mathrm{d}t\\
&\quad + 2(T+4)\int_0^t k(\tau)\mathrm{d}\tau\, E\left[\sup_{s\le z\le \tau}|X^n_{s,x}(z)-X_{s,x}(z)|^2\right]\mathrm{d}\tau,
\end{aligned}
$$

which by Gronwall's inequality (see [49], p. 22) implies that

$$
\begin{aligned}
\|\tilde{X}^n_{s,x}-\tilde{X}_{s,x}\|^2 &= E\left[\sup_{0\le t\le T}|\tilde{X}^n_{s,x}(t)-\tilde{X}_{s,x}(t)|^2\right]\\
&\le 2(T+4)\left(\int_0^T k(t)\mathrm{d}t\right)\exp\left[2(T+4)\int_0^T k(t)\mathrm{d}t\right]\|u^n-u\|_T^2.
\end{aligned}
$$

Therefore, $\lim_{n\to\infty}\|\lambda_{s,x}(u_n)-\lambda_{s,x}(u)\|_T=0$ for every $u\in\mathcal{U}_T$ and every sequence $(u_n)_{n=1}^\infty$ of $\mathcal{U}_T$ converging to $u\in\mathcal{U}_T$. Finally, immediately from the definition of $\lambda_{s,x}$, for every $(s,x),(\bar{s},\bar{x})\in\mathbb{R}^+\times\mathbb{R}^d$ with $s<\bar{s}$, one gets

$$
\sup\{|\lambda_{s,x}(u)-\lambda_{\bar{s},\bar{x}}(u)|:u\in\mathcal{U}_T\}\le 2\left[|x-\bar{x}|+(\sqrt{T}+1)\sqrt{\int_s^{\bar{s}} m^2(t)\mathrm{d}t}\right],
$$

which implies that the mapping $\mathbb{R}^+\times\mathbb{R}^d\ni(s,x)\to\lambda_{s,x}(u)\in\mathbb{R}^d$ is uniformly continuous with respect to $u\in\mathcal{U}_T$. Similarly, this is true for the case $\bar{s}<s$. □

Now we can prove the following existence theorem.

Theorem 1.1. *Let f and g be measurable and satisfy conditions (H). If $K:\mathbb{R}^+\times\mathbb{R}^d\to\mathbb{R}$ and $\Phi:\mathbb{R}^+\times\mathbb{R}^d\times U\to\mathbb{R}$ are continuous and bounded, then for every bounded domain D, filtered probability space $\mathcal{P}_{\mathbb{F}}$, m-dimensional $\mathbb{F}$-Brownian motion $B=(B_t)_{t\ge 0}$ defined on $\mathcal{P}_{\mathbb{F}}$, and $(s,x)\in\mathbb{R}^+\times\mathbb{R}^d$, there* *exists* $\bar{u} \in \mathcal{U}_T$ *such that* $I^D_{s,x}(\bar{u}, \bar{X}^{\bar{u}}_{s,x}) = \sup\{I^D_{s,x}(u, X^u_{s,x}) : u \in \mathcal{U}_T\}$, *where* $I^D_{s,x}(u, X^u_{s,x}) = J^{u,X}_D(s,x)$ *and* $X^u_{s,x}$ *is the unique strong solution of (1.1) on the filtered probability space* $\mathcal{P}_{\mathbb{F}}$ *corresponding to the Brownian motion* B *and* $u \in \mathcal{U}_T$.

Proof. Similarly as above, by virtue of Remark 1.1, for every $u \in \mathcal{U}_T$, there exists a unique strong solution of (1.1) defined on $[s,T]\times\Omega$. Observe that $\sup\{I^D_{s,x}(u, X^u_{s,x}) : u \in \mathcal{U}_T\} = \sup\{I^D_{s,x}(u, \lambda_{s,x}(u)) : u \in \mathcal{U}_T\}$. Let $\alpha = \sup\{I^D_{s,x}(u, \lambda_{s,x}(u)) : u \in \mathcal{U}_T\}$, and let $(u_n)_{n=1}^\infty$ be a sequence of $\mathcal{U}_T$ such that $\alpha = \lim_{n\to\infty} I^D_{s,x}(u_n, \lambda_{s,x}(u_n))$. By the compactness of $\mathcal{U}_T$, there exist an increasing subsequence $(n_k)_{k=1}^\infty$ of $(n)_{n=1}^\infty$ and $\bar{u} \in \mathcal{U}_T$ such that $\|u_{n_k} - \bar{u}\|_T \to 0$ as $k \to \infty$. By virtue of Lemma 1.1, it follows that $\|\lambda_{s,x}(u_{n_k}) - \lambda_{s,x}(u)\|_T \to 0$ as $k \to \infty$. By the definitions of the operator $\lambda_{s,x}$ and the norm $\|\cdot\|$, it follows that there exists a subsequence, still denoted by $(X^{n_k}_{s,x})_{k=1}^\infty$, of the sequence $(X^{n_k}_{s,x})_{k=1}^\infty$ such that $\sup_{0\le t\le T} |\tilde{X}^{n_k}_{s,x} - \bar{X}_{s,x}| \to 0$ a.s. as $k \to \infty$, where $\bar{X}_{s,x} = \lambda_{s,x}(\bar{u})$. By virtue of Lemma 10.1 of Chap. 1 and Theorem 5.1 of Chap. 4, we have $\tilde{\tau}^{n_k}_D \to \bar{\tau}_D$ a.s. as $k \to \infty$, where $\tilde{\tau}^{n_k}_D$ and $\bar{\tau}_D$ denote the first exit times of $\tilde{X}^{n_k}_{s,x}$ and $\bar{X}_{s,x}$, respectively, from the domain D. Hence, by the continuity of Φ and K, it follows that $\alpha = \lim_{k\to\infty} I^D_{s,x}(u_{n_k}, \lambda_{s,x}(u_{n_k})) = I^D_{s,x}(\bar{u}, \lambda_{s,x}(\bar{u})) = I^D_{s,x}(\bar{u}, \bar{X}_{s,x})$. Thus $(\bar{u}, \bar{X}_{s,x}|_{[s,T]})$ is an optimal pair for (1.3).□

We can consider the above optimal control problem with a special type of controls $u = (u_t)_{t\ge0}$ of the form $u_t = \varphi(t, X_t)$ a.s. for $t \ge 0$ and a measurable function $\varphi : \mathbb{R}^+ \times \mathbb{R}^d \to U \subset \mathbb{R}^k$. Such controls are called Markov controls, because with such u, the corresponding process $X = (X_t)_{t\ge0}$ becomes an Itô diffusion, in particular a Markov process. In what follows, the above Markov control will be identified with a measurable function φ, and this function will be simply called a Markov control. The set of all such Markov controls will be denoted by $\mathcal{M}(U)$. The set of all restrictions of all $\varphi \in \mathcal{M}(U)$ to the set $[0,T]\times\mathbb{R}^d$ is denoted by $\mathcal{M}_T(U)$. Immediately from Theorem 1.1, it follows that for all measurable functions f and g satisfying conditions (H), there exists an optimal control for (1.3) in the set $\mathcal{S}_T$ consisting of all bounded and uniformly Lipschitz continuous Markov controls $\varphi \in \mathcal{M}_T(U)$, i.e., with the property that there exists a number $L > 0$ such that $|\varphi(t,z) - \varphi(s,v)| \le L(|t-s| + |z-v|)$ for every $\varphi \in \mathcal{S}_T$, $t,s \in [0,T]$, and $z, v \in \mathbb{R}^d$. Indeed, for functions f, g, and $\varphi, \psi \in \mathcal{S}_T \subset \mathcal{M}_T(U)$ as given above, we have

$$|f(t,x,\varphi(t,x)) - f(t,z,\psi(t,z))|^2 \le 2|f(t,x,\varphi(t,x)) - f(t,z,\varphi(t,z))|^2 + 2|f(t,z,\varphi(t,z)) - f(t,z,\psi(t,z))|^2 \le 2k(t)\left[(1+L^2)|x-z| + 2L^2\|\varphi-\psi\|^2_T\right]$$

and

$$\|g(t,x,\varphi(t,x)) - g(t,z,\psi(t,z))\|^2 \le 2k(t)\left[(1+L^2)|x-z| + 2L^2\|\varphi-\psi\|^2_T\right]$$

for every $t \in [0,T]$ and $x, z \in \mathbb{R}^d$, where $\|\cdot\|_T$ denotes the supremum norm of the space $\mathbf{C}([0,T]\times\mathbb{R}^d, \mathbb{R}^k)$ of all continuous bounded functions $v : [0,T]\times\mathbb{R}^d \to \mathbb{R}^k$.

Hence, similarly as in the proof of Lemma 1.1, it follows that the mapping $\lambda_{s,x}$: $\mathcal{S}_T \ni u \to \tilde{X}^{\varphi}_{s,x} \in C^T_{\mathbb{F}}$ with $\tilde{X}^{\varphi}_{s,x}$ and $C^T_{\mathbb{F}}$ as above is continuous. Here $X^{\varphi}_{s,x}$ is a strong solution of (1.1) corresponding to the Markov control $\varphi \in \mathcal{S}_T$. Therefore, immediately from Theorem 1.1, we obtain the existence in $\mathcal{S}_T$ of the optimal control for (1.3).

2 Optimal Control Problems for Systems Described by Stochastic Functional Inclusions

We shall now extend the above optimal control problem (1.3) on the case in which the dynamics of a control system is described by stochastic functional inclusions $SFI(F,G)$ of the form

$$\begin{cases} X_t - X_s \in \int_s^t F(\tau, X_\tau)\mathrm{d}\tau + \int_s^t G(\tau, X_\tau)\mathrm{d}B_\tau \; for \quad t \geq s\,, \\ X_s = x \quad a.s. \end{cases} \tag{2.1}$$

with the performance functional depending only on the weak solution $(\mathcal{P}_{\mathbb{F}}, X, B)$ of $SFI(F,G)$, i.e., with the performance functional $J^X_D(s,x)$ of the form

$$J^X_D(s,x) = E^{s,x}\left[\int_s^{\tau_D} \Psi(t, X_t)\mathrm{d}t + K(\tau_D, X_{\tau_D})\right], \tag{2.2}$$

where D is a bounded subset of $\mathbb{R}^d$, and $\Psi : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}$ and $K : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}$ are given continuous functions. By a solution of such a stochastic optimal control problem we mean a weak solution $(\tilde{\mathcal{P}}_{\tilde{\mathbb{F}}}, \tilde{X}, \tilde{B})$ of (2.1) such that $J^{\tilde{X}}_D(s,x) = \sup\{J^X_D(s,x) : X \in \mathcal{X}^D_{s,x}\}$, where $\mathcal{X}^D_{s,x}$ denotes the set of all weak solutions of (equivalence classes of) the stochastic functional inclusion $SFI(F,G)$ satisfying an initial condition $X(s) = x$ and such that $\tau_D = \inf\{t > s : X_{s,x}(t) \notin D\} < \infty$. Such an optimal control problem will be denoted by

$$\begin{cases} X_t - X_s \in \int_s^t F(\tau, X_\tau)\mathrm{d}\tau + \int_s^t G(\tau, X_\tau)\mathrm{d}B_t\tau \text{ for } t \geq s\,, \\ X_s = x \;\; a.s.\,, \\ J^X_D(s,x) \xrightarrow{\mathcal{X}^D_{s,x}} \max, \end{cases} \tag{2.3}$$

and called an optimal control problem for the control system described by the stochastic functional inclusion $SFI(F,G)$. In this case, the set $\mathcal{X}^D_{s,x}$ is said to be an admissible set for the optimal control problem (2.3). If there is $(\tilde{\mathcal{P}}_{\tilde{\mathbb{F}}}, \tilde{X}, \tilde{P}) \in \mathcal{X}^D_{s,x}$ such that $J^{\tilde{X}}_D(s,x) = \sup\{J^X_D(s,x) : X \in \mathcal{X}^D_{s,x}\}$, then $(\tilde{\mathcal{P}}_{\tilde{\mathbb{F}}}, \tilde{X}, \tilde{P})$ is called the optimal solution of the optimal control problem (2.3). Similarly as above, it will be simply denoted by $\tilde{X}$. We shall consider the optimal control problems of the form (2.3) with set-valued mappings $F : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ such that the set $\mathcal{X}_{s,x}(F,G)$ of all weak solutions of $SFD(F,G)$ satisfying the initial condition $X(s) = x$ is weakly compact in distribution and such that $\mathcal{X}^D_{s,x}(F,G) \neq \emptyset$. Hence, immediately from Theorem 5.1 of Chap. 4, it will follow that $\mathcal{X}^D_{s,x}$ is also weakly compact. We apply the result obtained to the case of F and G defined by $F(t,x) = \{f(t,x,z) : z \in U\}$ and $G(t,x) = \{g(t,x,z) : z \in U\}$. Hence in particular, the existence of optimal pairs for the optimal control problems of the system described by (1.1) and performance functionals of the form

$$J^X_D(s,x) = E^{s,x}\left[\int_s^{\tau_D} \sup_{u\in U} \Phi(t,X_t,u))\mathrm{d}t + K(\tau_D, X_{\tau_D})\right] \tag{2.4}$$

and

$$J^X_D(s,x) = E^{s,x}\left[\int_s^{\tau_D} \sup_{n\geq 1} \Phi(t,X_t,\varphi^n(t,X_t))\mathrm{d}t + K(\tau_D, X_{\tau_D})\right] \tag{2.5}$$

will follow, where $(\varphi^n)_{n=1}^\infty$ is a dense sequence of a bounded set $\mathcal{U} \subset C(\mathbb{R}^+\times\mathbb{R}^d, U)$. In what follows, we shall still denote by $(\mathcal{P})$ and $(\mathcal{A})$ the assumptions defined in Sect. 1 of Chap. 6.

Theorem 2.1. *Let $F : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ be convex-valued, continuous, and bounded, and let $\Psi : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}$ be a uniformly integrally bounded Carathéodory function. Assume that G is diagonally convex and satisfies item (iv′) of conditions $(\mathcal{A})$. Let D be a bounded domain in $\mathbb{R}^d$. If $K : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}$ is continuous and bounded, then for every $(s,x) \in \mathbb{R}^+ \times D$, the optimal control problem (2.3) possesses an optimal solution.*

Proof. Let us observe that $\mathcal{X}^D_{s,x}$ is nonempty and weakly compact in distribution. Indeed, similarly as in the proof of Theorem 4.1 of Chap. 4, we can verify that $\mathcal{X}_{s,x}(F,G)$ is weakly compact in distribution for every $(s,x) \in \mathbb{R}^+ \times \mathbb{R}^d$. By property $(\mathcal{P})$ of G, for every $(f,g) \in \mathcal{C}(F) \times \mathcal{C}(G)$, there exists a unique in law solution $(\tilde{\mathcal{P}}_{\tilde{\mathbb{F}}}, \tilde{x}, \tilde{B})$ of $SDE(f,g)$ with initial condition $\tilde{x}_s = x$ a.s., which by the properties of functions f and g, implies that $(\tilde{\mathcal{P}}_{\tilde{\mathbb{F}}}, \tilde{x}, \tilde{B}) \in \mathcal{X}_{s,x}(F,G)$. By Remark 10.4 of Chap. 1, we have $\tilde{\tau}_D < \infty$ a.s., where $\tilde{\tau}_D$ is the first exit time of $\tilde{x}$ from the set D. Then $(\tilde{\mathcal{P}}_{\tilde{\mathbb{F}}}, \tilde{x}, \tilde{B}) \in \mathcal{X}^D_{s,x}$. To verify that $\mathcal{X}^D_{s,x}$ is weakly compact, let us observe that by the weak compactness of $\mathcal{X}_{s,x}(F,G)$ and the relation $\mathcal{X}^D_{s,x} \subset \mathcal{X}_{s,x}(F,G)$, it is enough to verify that $\mathcal{X}^D_{s,x}$ is weakly closed.

Let $(x^r)_{r=1}^\infty$ be a sequence of $\mathcal{X}^D_{s,x}$ convergent in distributions. Then there exists a probability measure $\mathcal{P}$ on $\beta(C(\mathbb{R}^+,\mathbb{R}^d))$ such that $P(x^r)^{-1} \Rightarrow \mathcal{P}$ as $r \to \infty$. By virtue of Theorem 2.3 of Chap. 1, there exist a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ and random variables $\tilde{x}^r : \tilde{\Omega} \to C(\mathbb{R}^+,\mathbb{R}^d)$ and $\tilde{x} : \tilde{\Omega} \to C(\mathbb{R}^+,\mathbb{R}^d)$ for $r = 1,2,\ldots$ such that $P(x^r)^{-1} = P(\tilde{x}^r)^{-1}$ for $r = 1,2,\ldots$, $\tilde{P}(\tilde{x})^{-1} = \mathcal{P}$ and $\lim_{r\to\infty} \rho(\tilde{x}^r, \tilde{x}) = 0$ with $(\tilde{P}.1)$, where ρ is the metric defined in $C(\mathbb{R}^+,\mathbb{R}^d)$ as in Theorem 2.4 of Chap. 1. For every $r \geq 1$, one has $\tau_D^r < \infty$ a.s., where τ_D^r is the first exit time of x^r from the set D, which by Theorem 5.2 of Chap. 4, implies that $\tilde{\tau}_D^r < \infty$ a.s., where $\tilde{\tau}_D^r$ denotes the first exit time of $\tilde{x}^r$ from the set D. Hence, by the properties of the sequence $(\tilde{x}^r)_{r=1}^{\infty}$, it follows that $\tilde{\tau}_D < \infty$ a.s., where $\tilde{\tau}_D$, is the first exit time of $\tilde{x}$ from D. Similarly as in the proof of Theorem 4.1 of Chap. 4, we can verify now that by virtue of Theorem 1.3 of Chap. 4, there exist a standard extension $\hat{\mathcal{P}}_{\hat{\mathbb{F}}} = (\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{F}}, \hat{P})$ of $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{F}}, \tilde{P})$ and an m-dimensional Brownian motion $\hat{B}$ such that $(\hat{\mathcal{P}}_{\hat{\mathbb{F}}}, \hat{x}, \hat{B})$ is a weak solution of $SFI(F, G, \mu)$, with $\mu = P\tilde{x}_s^{-1}$ and such that $x^r \Rightarrow \hat{x}$. Furthermore, we have $P\hat{x}^{-1} = P\tilde{x}^{-1}$, which by Theorem 5.2 of Chap. 4, implies that $P\hat{\tau}_D^{-1} = P\tilde{\tau}_D^{-1}$. Hence in particular, it follows that $\hat{\tau}_D < \infty$. Thus $\mathcal{X}_\mu^D$ is weakly closed with respect to weak convergence in the sense of distributions.

By (2.2) and the properties of the functions Ψ and K, one has $\alpha := \sup\{J_D^X(s,x) : X \in \mathcal{X}_{s,x}^D\} < \infty$, because

$$\int_s^{\tau_D^X} \Psi(t, X(t))\mathrm{d}t \leq \int_s^{\tau_D^X} |\Psi(t, X(t))|\mathrm{d}t \leq \int_0^{\tau_D^X} m(t)\mathrm{d}t \leq \int_0^{\infty} m(t)\mathrm{d}t < \infty,$$

where $m \in \mathbb{L}(\mathbb{R}^+, \mathbb{R}^+)$ is such that $|\Psi(t,x)| \leq m(t)$ and there is $M > 0$ such that $|K(t,x)| \leq M$ for $x \in \mathbb{R}^d$ and $t \geq 0$. Let $(\mathcal{P}_{\mathbb{F}_n}^n, X^n, B^n) \in \mathcal{X}_{s,x}^D$ be for $n = 1, 2, \ldots$ such that $\alpha = \lim_{n\to\infty} J_D^n(s,x)$, with

$$J_D^n(s,x) = E^{s,x}\left[\int_s^{\tau_D^n} \Psi(t, X^n(t))\mathrm{d}t + K(\tau_D^n, X^n(\tau_D^n)),\right]$$

where $E_n^{s,x}$ denotes the mean value operator with respect to the probability law $Q_n^{s,x}$ of X^n and $\tau_D^n = \inf\{r > s : X^n(r) \notin D\}$ for $n = 1, 2, \ldots$. By the weak compactness of $\mathcal{X}_{s,x}^D$ and Theorem 2.3 of Chap. 1, there are an increasing subsequence $(n_k)_{k=1}^{\infty}$ of the sequence $(n)_{n=1}^{\infty}$, a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$, and continuous processes $\tilde{X}^{n_k}$ and $\tilde{X}$ on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ such that $P(X^{n_k})^{-1} = P(\tilde{X}^{n_k})^{-1}$ for $k = 1, 2, \ldots$ and $\rho(\tilde{X}^{n_k}, \tilde{X}) \to 0$, $\tilde{P}$-a.s. as $k \to \infty$, which by Corollary 3.3 of Chap. 1, implies that $P(X_s^{n_k})^{-1} \Rightarrow P\tilde{X}_s^{-1}$ as $k \to \infty$. Let $\tilde{\mathcal{F}}_t = \bigcap_{\varepsilon>0} \sigma(\{\tilde{X}(u) : s \leq u \leq t + \varepsilon\})$ for $t \geq s$ and put $\tilde{\mathbb{F}} = (\tilde{\mathcal{F}}_t)_{t\geq s}$. It is clear that $\tilde{X}$ is $\tilde{\mathbb{F}}$-adapted. By virtue of Lemma 1.3 of Chap. 4, we have $\mathcal{M}_{FG}^{\tilde{X}} \neq \emptyset$, and therefore, there exist $\tilde{f} \in S_{\tilde{\mathbb{F}}}(F \circ \tilde{X})$ and $\tilde{g} \in S_{\tilde{\mathbb{F}}}(G \circ \tilde{X})$ such that for every $h \in C_0^2(\mathbb{R}^d)$, a stochastic process $\varphi_h^{\tilde{X}} = ((\varphi_h^{\tilde{X}})_t)_{t\geq s}$ with $(\varphi_h^{\tilde{X}})_t = h(\tilde{X}_t) - h(\tilde{X}_s) - \int_s^t (\mathbb{L}_{\tilde{f}\tilde{g}}^{\tilde{X}} h)_\tau d\tau$ for $t \geq s$ is a continuous local $\tilde{\mathbb{F}}$-martingale on the filtered probability space $\tilde{\mathcal{P}}_{\tilde{\mathbb{F}}} = (\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{F}}, \tilde{P})$. Hence, by Theorem 1.3 of Chap. 4, it follows that there exists a standard extension of $\tilde{\mathcal{P}}_{\tilde{\mathbb{F}}}$, still denoted by $\tilde{\mathcal{P}}_{\tilde{\mathbb{F}}}$, and an m-dimensional $\tilde{\mathbb{F}}$-Brownian motion $\tilde{B} = (\tilde{B}_t)_{t\geq 0}$ such that $(\tilde{\mathcal{P}}_{\tilde{\mathbb{F}}}, \tilde{X}, \tilde{B})$ is a weak solution of $SFI(F, G, \mu)$ with an initial distribution $\mu = P\tilde{X}_s^{-1}$. Immediately from the properties of the stochastic processes $\tilde{X}^{n_k}$ and $\tilde{X}$, it follows that $\widetilde{X}_s^{n_k} = x$, P-a.s., and $P(\widetilde{X}_s^{n_k})^{-1} \Rightarrow P(\widetilde{X}_s)^{-1}$ as $k \to \infty$, which implies that $\widetilde{X}_s = x$, P̃-a.s. Therefore, $(\tilde{\mathcal{P}}_{\tilde{\mathbb{F}}}, \tilde{X}, \tilde{B}) \in \mathcal{X}_{s,x}(F, G)$. Similarly as above, we can verify that $(\tilde{\mathcal{P}}_{\tilde{\mathbb{F}}}, \tilde{X}, \tilde{B}) \in \mathcal{X}^D_{s,x}$. On the other hand, By (2.2) and the properties of processes X^{n_k} and $\tilde{X}^{n_k}$ and Theorem 5.2 of Chap. 4, it follows that $P(\tau^k_D)^{-1} = P(\tilde{\tau}^k_D)^{-1}$ for $k = 1, 2, \ldots$. Then $J^{n_k}_D(s,x) = \tilde{J}^{n_k}_D(s,x)$ for every $k = 1, 2, \ldots$, where

$$\tilde{J}^{n_k}_D(s,x) = \tilde{E}^{s,x}\left[\int_s^{\tilde{\tau}^k_D} \Psi(t, \tilde{X}^{n_k}(t))\mathrm{d}t + K(\tilde{\tau}^k_D, \tilde{X}^{n_k}(\tilde{\tau}^k_D))\right]$$

for $k = 1, 2, \ldots$ with $\tilde{\tau}^k_D$ and τ^k_D defined as above with $\widetilde{X}^{n_k}_t = \widetilde{X}^{n_k}(t)$. Hence, by Theorem 5.1 of Chap. 4, it follows that

$$\lim_{k\to\infty} \tilde{J}^{n_k}_D(s,x) = \tilde{E}^{s,x}\left[\int_s^{\tilde{\tau}_D} \Psi(t, \tilde{X}(t))\mathrm{d}t + K(\tilde{\tau}_D, \tilde{X}(\tilde{\tau}_D))\right],$$

where $\tilde{\tau}_D = \inf\{r > s : \tilde{X} \notin D\}$. But $\alpha = \lim_{k\to\infty} J^{n_k}_D(s,x) = \lim_{k\to\infty} \tilde{J}^{n_k}_D(s,x)$. Therefore,

$$\alpha = \tilde{E}^{s,x}\left[\int_s^{\tilde{\tau}_D} \Psi(t, \tilde{X}(t))\mathrm{d}t + K(\tilde{\tau}_D, \tilde{X}(\tilde{\tau}_D))\right].$$

□

Remark 2.1. Similarly as above, we can consider the following viable optimal control problem:

$$\begin{cases} X_t - X_s \in \int_s^t F(\tau, X_\tau)\mathrm{d}\tau + \int_s^t G(\tau, X_\tau)\mathrm{d}B_t\tau, \text{ for } t \geq s\,, \\ X_t \in \Gamma(t) \text{ a.s. for } t \geq s, \\ J(X) \xrightarrow{\mathcal{X}^\Gamma_D} \max, \end{cases}$$

where Γ is a given target set mapping and $\mathcal{X}^\Gamma_D$ denotes the set of all weak Γ-viable solutions $(\mathcal{P}_{\mathbb{F}}, X, B)$ of the stochastic functional inclusion $SFI(F,G)$ such that $\tau^X_D = \inf\{t > s : X(t) \notin D\} < \infty$. □

We shall consider now the existence of the optimal control problem (1.3) with a performance functional $J^X_D(s,x)$ defined by (2.4) and (2.5) above. Let us recall that for a given nonempty set $U \subset \mathbb{R}^k$, a bounded domain D, an initial point $(s,x) \in \mathbb{R}^+ \times D$, and functions $f : \mathbb{R} + \times\mathbb{R}^d \times U \to \mathbb{R}^d$, $g : \mathbb{R} + \times\mathbb{R}^d \times U \to \mathbb{R}^{d\times m}$, $\Psi : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}$, and $K : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}$, we are interested in the existence of an admissible pair $(\bar{u}, \bar{X}^{\bar{u}}) \in \Lambda_{fg}(s,x)$ such that $J^{\bar{X}}_D(s,x) = \sup\{J^X_D(s,x) : (u, X^u) \in \Lambda_{fg}(s,x)\}$. We shall show that such an optimal pair $(\bar{u}, \bar{X}) \in \Lambda_{fg}(s,x)$ exists if $f : \mathbb{R}^+ \times \mathbb{R}^d \times U \to \mathbb{R}^d$ and $g : \mathbb{R}^+ \times \mathbb{R}^d \times U \to \mathbb{R}^{d\times m}$ satisfy the following conditions (C):

(i) f and g are continuous and bounded such that $f(t,x,\cdot)$, $g(t,x,\cdot)$, and $(g \cdot g^*)(t,x,\cdot)$ are affine for every fixed $(t,x) \in \mathbb{R}^+ \times \mathbb{R}^d$ on the compact convex set $U \subset \mathbb{R}^k$.

(ii) g is such that $g \cdot g^*$ is uniformly positive definite.
(iii) $\Phi : \mathbb{R}^+ \times \mathbb{R}^d \times U \to \mathbb{R}$ is a uniformly integrally bounded Carathéodory function and $K : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}$ is continuous and bounded.

Lemma 2.1. *If f and g satisfy conditions (C), then for every nonempty compact convex set $U \subset \mathbb{R}^k$, the set-valued mappings F and G defined by $F(t,x) = \{f(t,x,z) : z \in U\}$ and $G(t,x) = \{g(t,x,z) : z \in U\}$ satisfy $(\mathcal{P})$ and conditions (i), (iii), (iv′), and (v) of $(\mathcal{A})$.*

Proof. Immediately from (ii) of conditions (C), it follows that G satisfies the condition $(\mathcal{P})$. Let $\mathcal{T}_U$ be the induced topology in U. Then $(U, \mathcal{T}_U)$ is a compact topological space. Let $(\bar{t}, \bar{x}) \in \mathbb{R}^+ \times \mathbb{R}^d$ and $\bar{z} \in U$ be fixed and V an open set in $\mathbb{R}^d$. Suppose $(\bar{t}, \bar{x}, \bar{z})$ is such that $f(\bar{t}, \bar{x}, \bar{z}) \in V$. By the continuity of $f(\cdot,\cdot,\bar{z})$ at $(\bar{t}, \bar{x})$, there is a neighborhood $\mathcal{N}$ of $(\bar{t}, \bar{x})$ such that $f(t,x,\bar{z}) \in V$ for every $(t,x) \in \mathcal{N}$. Therefore, for every $(t,x) \in \mathcal{N}$, one has $F(t,x) \cap V \neq \emptyset$. Then F is l.s.c. In a similar way, we can also verify that G is l.s.c. By the compactness of the set U and continuity of $f(t,x,\cdot)$ and $g(t,x,\cdot)$, it follows that $F(t,x)$ and $G(t,x)$ are compact subsets of $\mathbb{R}^d$ and $\mathbb{R}^{d\times m}$, respectively, for every $(t,x) \in \mathbb{R}^+ \times \mathbb{R}^d$. Similarly, by the convexity of U and affineness of $f(t,x,\cdot)$, $g(t,x,\cdot)$, and $(g \cdot g^*)(t,x,\cdot)$, it follows that F and G are convex-valued and G is diagonally convex. We shall verify that F and G are also u.s.c. Indeed, similarly as above, let $(\bar{t}, \bar{x}) \in \mathbb{R}^+ \times \mathbb{R}^d$ be arbitrarily fixed and suppose V is an open neighborhood of $F(\bar{t}, \bar{x})$. By the continuity of f, for every fixed $z \in U$ there exist neighborhoods W^z and $\mathcal{O}^z$ of $(\bar{t}, \bar{x}) \in \mathbb{R}^+ \times \mathbb{R}^d$ and $\bar{z} \in U$, respectively, such that $f(W^z \times \mathcal{O}^z) \subset V$. By the compactness of the topological space $(U, \mathcal{T}_U)$, there are $z_1, \ldots, z_n \in U$ such that $\bigcup_{i=1}^n \mathcal{O}^{z_i} = U$. For every $i = 1, 2, \ldots, n$, we have $f(W^{z_i} \times \mathcal{O}^{z_i}) \subset V$. Therefore, $\bigcup_{i=1}^n f(W^{z_i} \times \mathcal{O}^{z_i}) \subset V$. But

$$\bigcup_{i=1}^n f\left(\left[\bigcap_{i=1}^n W^{z_i}\right] \times \mathcal{O}^{z_i}\right) = f\left(\left[\bigcap_{i=1}^n W^{z_i}\right] \times \left[\bigcup_{i=1}^n \mathcal{O}^{z_i}\right]\right)$$
$$= f\left(\left[\bigcap_{i=1}^n W^{z_i}\right] \times U\right) \subset \bigcup_{i=1}^n f\left(W^{z_i} \times \mathcal{O}^{z_i}\right) \subset V.$$

Therefore, $F(\bigcap_{i=1}^n W^{z_i}) = f([\bigcap_{i=1}^n W^{z_i}] \times U) \subset V$. Then F is u.s.c. at $(\bar{t}, \bar{x}) \in \mathbb{R}^+ \times \mathbb{R}^d$. In a similar way, we can verify that also G is u.s.c.

Let $\sigma \in \mathcal{C}(l(G))$ be a continuous selector of $D(G) = l(G)$, where $l(u) = u \cdot u^*$ for every $u \in \mathbb{R}^{d\times m}$, and let $\lambda(t,x,z) = l(g(t,x,z)) = g(t,x,z) \cdot g^*(t,x,z)$ for $(t,x,z) \in \mathbb{R}^+ \times \mathbb{R}^d \times U$. We have $\sigma(t,x) \in \lambda(t,x,U)$ for $(t,x) \in \mathbb{R}^+ \times \mathbb{R}^d$. Therefore, by virtue of Theorem 2.2 of Chap. 2, there exists a sequence $(z_n)_{n=1}^\infty$ of continuous functions $z_n : \mathbb{R}^+ \times \mathbb{R}^d \to U$ such that $\sup_{(t,x)} |\sigma(t,x) - l(g_n(t,x))| \to 0$ as $n \to \infty$, where $g_n(t,x) = g(t,x,z_n(t,x)) \in G(t,x)$ for $n = 1, 2, \ldots$ and $(t,x) \in \mathbb{R}^+ \times \mathbb{R}^d$. Then there exists a sequence $(g_n)_{n=1}^\infty$ of continuous selectors of G such that $l(g_n) \to \sigma$ uniformly in $(t,x) \in \mathbb{R}^+ \times \mathbb{R}^d$ as $n \to \infty$. Thus (iv′) of conditions $(\mathcal{A})$ is also satisfied. □

We can now prove the existence of an optimal pair for the optimal control problem (1.3) with the performance functionals defined by (2.4) and (2.5).

Theorem 2.2. *Let D be a bounded domain in $\mathbb{R}^d$ and assume that conditions (C) are satisfied. There exists an optimal pair of the optimal control problem (1.3) with a performance functional defined by* (2.4)*.*

Proof. Let F and G be defined as above. By virtue of Lemma 2.1, the multi-functions F and G satisfy the conditions of Theorem 2.1. Therefore, for every $(s,x) \in \mathbb{R}^+ \times \mathbb{R}^d$, there exists a weak solution $(\tilde{\mathcal{P}}_{\tilde{\mathbb{F}}}, \tilde{X}, \tilde{B})$ of $SFI(F,G)$ satisfying the initial condition $\tilde{X}(s) = x$, $\tilde{P}$-a.s., with $\tilde{\mathcal{P}}_{\tilde{\mathbb{F}}} = (\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{F}}, \tilde{P})$ such that $J_D^{\tilde{X}}(s,x) = \sup\{J_D^X(s,x) : X \in \mathcal{X}_{s,x}\}$, where

$$J_D^{\tilde{X}}(s,x) = \tilde{E}^{s,x}\left[\int_s^{\tilde{\tau}_D} \Psi(t,\tilde{X}(t))\mathrm{d}t + K(\tilde{\tau}_D, \tilde{X}(\tilde{\tau}_D))\right]$$

with $\Psi(t,x) = \sup\{\Phi(t,x,u) : u \in U\}$ and $\tilde{\tau}_D = \inf\{r > s : \tilde{X}(r) \notin D\}$. By virtue of Theorem 1.5 of Chap. 3, there are $\tilde{f} \in S_{\tilde{\mathbb{F}}}(F \circ \tilde{X})$ and $\tilde{g} \in S_{\tilde{\mathbb{F}}}(G \circ \tilde{X})$ such that $\tilde{X}(t) = x + \int_s^t \tilde{f}_\tau \mathrm{d}\tau + \int_s^t \tilde{g}_\tau \mathrm{d}\tilde{B}_\tau$, $\tilde{P}$-a.s. for $t \geq s$. Let $\Gamma(t,x) = \{(f(t,x,z), g(t,x,z)) : z \in U\}$ for $(t,x) \in \mathbb{R}^+ \times \mathbb{R}^d$. Similarly as in the proof of Lemma 2.1, we can verify that Γ is a continuous and bounded set-valued mapping with compact values in $\mathbb{R}^d \times \mathbb{R}^{d\times m}$. Therefore, the set-valued process $\tilde{\Gamma} = (\tilde{\Gamma}_t)_{t\geq s}$ defined by $\tilde{\Gamma}_t = \Gamma(t,\tilde{X}(t))$ is $\tilde{\mathbb{F}}$-nonanticipative and such that $(\tilde{f}_t, \tilde{g}_t) \in \tilde{\Gamma}_t$, $\tilde{P}$-a.s. for $t \geq s$. By virtue of Theorem 2.5 of Chap. 2, there exists an $\tilde{\mathbb{F}}$-nonanticipative process $\tilde{u} = (\tilde{u}_t)_{t\geq s}$ with values in the set U such that $(\tilde{f}_t, \tilde{g}_t) = (f(t,\tilde{X}(t),\tilde{u}_t), g(t,\tilde{X}(t),\tilde{u}_t))$, $\tilde{P}$-a.s. for $t \geq s$. Then an optimal solution $\tilde{X}$ of the optimal control problem (1.3) with the performance functional (2.4) can be expressed by the formula

$$\tilde{X}(t) = x + \int_s^t f(\tau,\tilde{X}(\tau),\tilde{u}_\tau)\mathrm{d}\tau + \int_s^t g(\tau,\tilde{X}(\tau),\tilde{u}_\tau)\mathrm{d}\tilde{B}_\tau$$

$\tilde{P}$-a.s. for $t \geq s$. Therefore, $(\tilde{u},\tilde{X}) \in \Lambda_{fg}(s,x)$. In a similar way, we deduce that for every weak solution $(\mathcal{P}_{\mathbb{F}}, X, B)$ of $SFI(F,G)$ satisfying the initial condition $X(s) = x$ a.s. with the above-defined set-valued mappings F and G, there exists an $\mathbb{F}$-nonanticipative stochastic process $u = (u_t)_{t\geq s}$ with values in U such that $(u,X) \in \Lambda_{fg}(s,x)$. By the properties of the performance functional $J_D^X(s,x)$ defined by (2.4), one has

$$J_D^{\tilde{X}}(s,x) = \sup\{J_D^X(s,x) : X \in \mathcal{C}_{s,x}^D\} = \sup\{J_D^X(s,x) : (u,X) \in \Lambda_{f,g}(s,x)\}$$

with $\mathcal{C}_{s,x}^D = \pi(\Lambda_{f,g}(s,x))$, where $\pi(u,X) = X$ for $(u,X) \in \Lambda_{fg}(s,x)$. Then $(\tilde{u},\tilde{X})$ is the optimal pair for the optimal control problem (1.3) with the performance functional defined by (2.4). □

In a similar way, we can prove the following existence theorem.

Theorem 2.3. *Let D be a bounded domain in $\mathbb{R}^d$, $\mathcal{U}$ a bounded subset of $C(\mathbb{R}^+ \times \mathbb{R}^d, U)$, and $(\varphi^n)_{n=1}^{\infty}$ a dense sequence of $\mathcal{U}$. Assume that conditions (C) are satisfied and that f and g are such that $f(t,x,\cdot)$ and $g(t,x,\cdot)$ are linear. There exists an optimal pair $(\tilde{u}, \tilde{X})$ for the optimal control problem (1.3) with the performance functional $J_D^X(s,x)$ defined by (2.5) and $\tilde{u} = \lim^w_{j\to\infty} \sum_{k=1}^{m_j} \mathbb{I}_{C_k^j} \varphi^{n_j^k}(\cdot, \tilde{X}.)$, where $\lim^w$ denotes the weak limit of sequences in the space $\mathbb{L}(\mathbb{R}^+ \times \Omega, \Sigma_{\mathbb{F}}, \mathbb{R}^k)$, $\{C_1^j, \dots, C_{m_j}^j\}$ is a finite $\Sigma_{\mathbb{F}}$-partition of $\mathbb{R}^+ \times \Omega$, and $\{\varphi^{n_j^1}, \dots, \varphi^{n_j^{m_j}}\} \subset \{\varphi^n : n \ge 1\}$ for every $j \ge 1$.*

Proof. Let F and G be defined by $F(t,x) = \{f(t,x,\varphi(t,x)) : \varphi \in \mathcal{U}\}$ and $G(t,x) = \{g(t,x,\varphi(t,x)) : \varphi \in \mathcal{U}\}$ for $(t,x) \in \mathbb{R}^+ \times \mathbb{R}^d$. By virtue of Lemma 2.1, F and G satisfy the conditions of Theorem 2.1. Therefore, for every $(s,x) \in \mathbb{R}^+ \times \mathbb{R}^d$, there exists a weak solution $(\tilde{\mathcal{P}}_{\tilde{\mathbb{F}}}, \tilde{X}, \tilde{B})$ of $SFI(F,G)$ satisfying the initial condition $\tilde{X}(s) = x$, $\tilde{P}$-a.s., with $\tilde{\mathcal{P}}_{\tilde{\mathbb{F}}} = (\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{F}}, \tilde{P})$ such that $J_D^{\tilde{X}}(s,x) = \sup\{J_D^X(s,x) : X \in \mathcal{X}_{s,x}\}$. By virtue of Theorem 1.5 of Chap. 3, there are $\tilde{f} \in S_{\tilde{\mathbb{F}}}(F \circ \tilde{X})$ and $\tilde{g} \in S_{\tilde{\mathbb{F}}}(G \circ \tilde{X})$ such that $\tilde{X}_t = x + \int_s^t \tilde{f}_\tau \mathrm{d}\tau + \int_s^t \tilde{g}_\tau \mathrm{d}\tilde{B}_\tau$, $\tilde{P}$-a.s. for $t \ge s$. By the properties of the sequence $(\varphi^n)_{n=1}^{\infty}$, it follows that $F(t,x) = \mathrm{cl}\{f(t,x,\varphi^n(t,x)) : n \ge 1\}$ and $G(t,x) = \mathrm{cl}\{g(t,x,\varphi^n(t,x)) : n \ge 1\}$ for $(t,x) \in \mathbb{R}^+ \times \mathbb{R}^d$. Therefore, by virtue of Lemma 4.1 of Chap. 2, it follows that $S_{\tilde{\mathbb{F}}}(F \circ \tilde{X}) = \overline{\mathrm{dec}}\{f(\cdot, \tilde{X}., \varphi^n(\cdot, \tilde{X}.)) : n \ge 1\}$ and $S_{\tilde{\mathbb{F}}}(G \circ \tilde{X}) = \overline{\mathrm{dec}}\{g(\cdot, \tilde{X}., \varphi^n(\cdot, \tilde{X}.)) : n \ge 1\}$. Hence it follows that $(\tilde{f}, \tilde{g}) \in \overline{\mathrm{dec}}\{(f,g)(\cdot, \tilde{X}., \varphi^n(\cdot, \tilde{X}.)) : n \ge 1\}$. Thus there exists a sequence $(\alpha_j)_{j=1}^{\infty}$ of $\mathrm{dec}\{f,g)(\cdot, \tilde{X}., \varphi^n(\cdot, \tilde{X}.)) : n \ge 1\}$ converging to $(\tilde{f}, \tilde{g})$ in the metric topology of $\mathbb{L}^2(\mathbb{R}^+ \times \Omega, \Sigma_{\mathbb{F}}, \mathbb{R}^d \times \mathbb{R}^{d\times m})$. But $\mathrm{dec}\{(f,g)(\cdot, \tilde{X}., \varphi^n(\cdot, \tilde{X}.)) : n \ge 1\} = (f,g)(\cdot, \tilde{X}., \mathrm{dec}\{\varphi^n(\cdot, \tilde{X}.) : n \ge 1\})$. Therefore, for every $j \ge 1$, there exist a finite $\Sigma_{\mathbb{F}}$-partition $\{C_1^j, \dots, C_{m_j}^j\}$ of $\mathbb{R}^+ \times \Omega$ and a family of $\{\varphi^{n_j^1}, \dots, \varphi^{n_j^{m_j}}\} \subset \{\varphi^n : n \ge 1\}$ such that $\alpha_j = (f,g)(\cdot, \tilde{X}., \sum_{k=1}^{m_j} \mathbb{I}_{C_k^j} \varphi^{n_j^k}(\cdot, \tilde{X}.))$ for $j \ge 1$. By the boundedness of the set $\mathcal{U}$, it follows that the sequence $(\sum_{k=1}^{m_j} \mathbb{I}_{C_k^j} \varphi^{n_j^k}(\cdot, \tilde{X}.))_{j=1}^{\infty}$ is relatively sequentially weakly compact. Then there exist $\tilde{u} \in \mathbb{L}(\mathbb{R}^+ \times \Omega, \Sigma_{\mathbb{F}}, \mathbb{R}^k)$ and a subsequence, still denoted by $(\sum_{k=1}^{m_j} \mathbb{I}_{C_k^j} \varphi^{n_j^k}(\cdot, \tilde{X}.))_{j=1}^{\infty}$, weakly converging to $\tilde{u}$. Hence, by the properties of the functions f and g, it follows that $(\tilde{f}, \tilde{g}) = (f(\cdot, \tilde{X}., \tilde{u}), g(\cdot, \tilde{X}., \tilde{u}))$. Similarly as in the proof of Theorem 2.3, it follows that $(\tilde{u}, \tilde{X})$, with the optimal control $\tilde{u}$ described above, is the optimal pair for the optimal control problem (1.3) with the performance functional $J_D^X(s,x)$ defined by (2.5). □

3 Optimal Problems for Systems Described by Partial Differential Inclusions

Let $F : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ be such that the following conditions $(\mathcal{D})$ are satisfied:

(i) F and G are bounded, continuous, and convex-valued, and for every $g \in \mathcal{C}(G)$, the matrix-valued mapping $l(g) = g \cdot g^*$ is uniformly positive definite.
(ii) G is diagonally convex, i.e., for every $(t,x) \in \mathbb{R}^+ \times \mathbb{R}^d$, the set $D(G)(t,x) = \{v \cdot v^* : v \in G(t,x)\}$ is convex.
(iii) For every $\sigma \in \mathcal{C}(D(G))$, there exists a sequence $(g^n)_{n=1}^{\infty}$ of $\mathcal{C}(G)$ such that $\sup_{(t,x)\in\mathbb{R}^+\times\mathbb{R}^d} |\sigma_n(t,x) - \sigma(t,x)| \to 0$ as $n \to \infty$, where $\sigma_n = l(g_n)$ for $n \geq 1$.

For a bounded domain $D \subset \mathbb{R}^d$, $T > 0$, $(s,x) \in \mathbb{R}^+ \times \mathbb{R}^d$, $\tilde{h} \in C_0^2(\mathbb{R}^{d+1})$, $u \in C([0,T] \times D, \mathbb{R})$, and a continuous function $\Phi : (0,T) \times \partial D \to \mathbb{R}$, we shall consider the initial and boundary values problems (6.3) and (6.4) of Chap. 6 of the form:

$$\begin{cases} v_t'(t,s,x) - v_s'(t,s,x) \in (\mathbb{L}_{FG} v(t,\cdot))\,(s,x) - c(s,x)v(t,s,x) \\ \text{for } (s,x) \in [0,T) \times \mathbb{R}^d \text{ and } t \in [0, T-s], \\ v(0,s,x) = \tilde{h}(s,x) \text{ for } (s,x) \in [0,T) \times \mathbb{R}^d\ , \end{cases}$$

and

$$\begin{cases} u(t,x) - v_t'(t,x) \in (\mathbb{L}_{FG} v)\,(t,x) - c(t,x)v(t,x) \text{ for } (t,x) \in (0,T) \times D, \\ \lim_{D\ni x\to y} v(t,x) = \Phi(t,y) \text{ for } (t,y) \in (0,T] \times \partial D. \end{cases}$$

Let $H : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ be measurable and uniformly integrably bounded and let $\Lambda_{FG}(c,\tilde{h})$ and $\Gamma_{FG}(c,u,\Phi)$ denote the sets of all solutions of the above initial and boundary value problems, respectively. For every $(s,x) \in [0,T) \times \mathbb{R}^d$, let $\mathcal{H}_{s,x}$ and $\mathcal{Z}_x$ denote the mappings defined on $\Lambda_{FG}(c,\tilde{h})$ and $\Gamma_{FG}(c,u,\Phi)$, respectively, by setting

$$\mathcal{H}_{s,x}(v) = \int_0^T H(t, v(t,s,x))\mathrm{d}t \quad \text{for} \quad v \in \Lambda_{FG}(c,\tilde{h})$$

and

$$\mathcal{Z}_x(w) = \int_0^T H(t, w(t,x))\mathrm{d}t \quad \text{for} \quad w \in \Gamma_{FG}(c,u,\Phi).$$

For every fixed $(s,x) \in [0,T) \times \mathbb{R}^d$, we shall look for $\tilde{v} \in \Lambda_{FG}^{\mathcal{C}}(c,\tilde{h})$ and $\tilde{v} \in \Gamma_{FG}^{\mathcal{C}}(c,u,\Phi)$ such that $\mathcal{H}_{sx}(\tilde{v}) = \inf\{\mathcal{H}_{s,x}(v) : v \in \Lambda_{FG}^{\mathcal{C}}(c,\tilde{h})\}$ and $\mathcal{Z}_x(\tilde{v}) = \inf\{\mathcal{Z}_x(u) : u \in \Gamma_{FG}^{\mathcal{C}}(c,u,\Phi)\}$, where $\Lambda_{FG}^{\mathcal{C}}(c,\tilde{h}) = \Lambda_{FG}(c,\tilde{h}) \cap C_b^{1,1,2}(\mathbb{R}^{d+2})$ and $\Gamma_{FG}^{\mathcal{C}}(c,u,\Phi) = \Gamma_{FG}(c,u,\Phi) \cap C^{1,2}(\mathbb{R}^{d+1})$.

Theorem 3.1. *Assume that conditions* $(\mathcal{D})$ *are satisfied. Let* $c \in C([0,T] \times \mathbb{R}^d, \mathbb{R})$ *be bounded,* $\tilde{h} \in C^{1,2}(\mathbb{R}^{d+1})$*, and let* $H : [0,T] \times \mathbb{R} \to \mathbb{R}$ *be measurable and uniformly integrally bounded such that* $H(t,\cdot)$ *is continuous. If* F *and* G *are furthermore such that for* $\tilde{h}$ *and* c *as given above, the set* $\Lambda^{\mathcal{C}}_{FG}(c,\tilde{h})$ *is nonempty, then there is* $\tilde{X} \in \mathcal{X}_{s,x}(F,G)$ *such that the function* $\tilde{v} \in \Lambda^{\mathcal{C}}_{FG}(c,\tilde{h})$ *defined by*

$$\tilde{v}(t,s,x) = \tilde{E}\left[\exp\left(-\int_s^{s+t} c(\tau,\tilde{X}(\tau))\mathrm{d}\tau\right)\tilde{h}(s+t,\tilde{X}(s+t))\right]$$

for every $(s,x) \in [0,T) \times \mathbb{R}^d$ *and* $t \in [0,T-s]$ *satisfies* $\mathcal{H}_{s,x}(\tilde{v}) = \inf\{\mathcal{H}_{s,x}(v) : v \in \Lambda^{\mathcal{C}}_{FG}(c,\tilde{h})\}$.

Proof. Let $(s,x) \in [0,T) \times \mathbb{R}^d$ be fixed. The set $\{\mathcal{H}_{s,x}(v) : v \in \Lambda^{\mathcal{C}}_{FG}(c,\tilde{h})\}$ is nonempty and bounded, because there is $k \in \mathbb{L}([0,T],\mathbb{R}_+)$ such that $|\mathcal{H}_{s,x}(v)| \le \int_0^T k(t)\mathrm{d}t$ for every $v \in \Lambda^{\mathcal{C}}_{FG}(c,\tilde{h})$. Therefore, there exists a sequence $(v^n)_{n=1}^{\infty}$ of $\Lambda^{\mathcal{C}}_{FG}(c,\tilde{h})$ such that $\alpha =: \inf\{\mathcal{H}_{s,x}(v) : v \in \Lambda^{\mathcal{C}}_{FG}(c,\tilde{h})\} = \lim_{n\to\infty}\mathcal{H}_{s,x}(v^n)$. By virtue of Theorem 6.4 of Chap. 6, for every $n = 1,2,\ldots$ and $(s,x) \in [0,T) \times \mathbb{R}^d$, there is $X^n_{s,x} \in \mathcal{X}_{s,x}(F,G)$ such that

$$v^n(t,s,x) = E^{s,x}\left[\exp\left(-\int_s^{s+t} c(\tau,X^n_{s,x}(\tau))\mathrm{d}\tau\right)\tilde{h}(s+t,X^n_{s,x}(s+t))\right]$$

for $(t,x) \in [0,T-s]\times\mathbb{R}^d$. By the weak compactness of $\mathcal{X}_{s,x}(F,G)$ and Theorem 2.3 of Chap. 1, there are an increasing subsequence $(n_k)_{k=1}^{\infty}$ of $(n)_{n=1}^{\infty}$, a probability space $(\tilde{\Omega},\tilde{\mathcal{F}},\tilde{P})$, and stochastic processes $\tilde{X}^{n_k}$ and $\tilde{X}$ on $(\tilde{\Omega},\tilde{\mathcal{F}},\tilde{P})$ such that $P(X^{n_k}_{s,x})^{-1} = P(\tilde{X}^{n_k})^{-1}$ for $k = 1,2,\ldots$ and $\sup_{0\le t\le T}|\tilde{X}^{n_k}(t) - \tilde{X}(t)| \to 0$ a.s. Hence in particular, it follows that

$$\begin{aligned}v^{n_k}(t,s,x) &= E^{s,x}\left[\exp\left(-\int_s^{s+t} c(\tau,X^{n_k}_{s,x}(\tau))\mathrm{d}\tau\right)\tilde{h}(s+t,X^{n_k}_{s,x}(s+t))\right]\\ &= \tilde{E}\left[\exp\left(-\int_s^{s+t} c(\tau,\tilde{X}^{n_k}(\tau))\mathrm{d}\tau\right)\tilde{h}(s+t,\tilde{X}^{n_k}(s+t))\right],\end{aligned}$$

where $\tilde{E}$ is the mean value operator taken with respect to the probability measure $\tilde{P}$. By the properties of processes $\tilde{X}^{n_k}$, $\tilde{X}$ and functions c and $\tilde{h}$, it follows that

$$\lim_{k\to\infty} v^{n_k}(t,s,x) = \tilde{E}\left[\exp\left(-\int_s^{s+t} c(\tau,\tilde{X}(\tau))\mathrm{d}\tau\right)\tilde{h}(s+t,\tilde{X}(s+t))\right].$$

By virtue of Theorem 6.3 of Chap. 6, it follows that the function $\tilde{v}(t,s,x) =: \lim_{k\to\infty} v^{n_k}(t,s,x)$ belongs to $\Lambda^{\mathcal{C}}(F,G,\tilde{h},c)$, because $(\mathcal{A}_{FG}v(t\cdot))(s,x) \subset v'_s(t,s,x) + (\mathbb{L}_{FG}v(t\cdot))(s,x)$ for $(s,x) \in [0,T] \times \mathbb{R}^d$ and $t \in [0,T-s]$. Hence, by the properties of the function H, we get $\alpha = \lim_{k\to\infty}\mathcal{H}_{s,x}(v^{n_k}) = \mathcal{H}_{s,x}(\tilde{v})$. □

Theorem 3.2. *Assume that conditions* $(\mathcal{D})$ *are satisfied,* $T > 0$*, and* D *is a bounded domain in* $\mathbb{R}^d$*. Let* $c \in C([0,T]\times\mathbb{R}^d, \mathbb{R})$*,* $u \in C((0,T)\times D, \mathbb{R})$*, and* $\Phi \in C([0,T]\times \partial D, \mathbb{R})$ *be bounded. Assume that* $H : [0,T]\times\mathbb{R} \to \mathbb{R}$ *is measurable and uniformly integrally bounded such that* $H(t,\cdot)$ *is continuous. If* F *and* G *are furthermore such that for* Φ*,* u*, and* c *given above, the set* $\Gamma^{\mathcal{C}}_{FG}(c,u,\Phi)$ *belongs to* $[0,T)\times\mathbb{R}^d$ *for every* (s,x)*, then there is* $\tilde{X}_{s,x} \in \mathcal{X}_{s,x}(F,G)$ *such that the function* $\tilde{v} \in \Gamma^{\mathcal{C}}_{FG}(c,u,,\Phi)$ *defined by*

$$\tilde{v}(s,x) = E^{s,x}\left[\Phi(\tau_D, \tilde{X}_{s,x}(\tau_D)) \exp\left(-\int_s^{\tau_D} c(t,\tilde{X}_{s,x}(t))\mathrm{d}t\right)\right] - E^{s,x}\left\{\int_s^{\tau_D}\left[u(t,\tilde{X}_x(t)) \exp\left(-\int_s^{s+t} c(z,\tilde{X}_{s,x}(z))\mathrm{d}z\right)\right]\mathrm{d}t\right\}$$

with $\tau_D = \inf\{r \in (s,T] : \tilde{X}_{s,x}(r) \notin D\}$ *satisfies* $\mathcal{Z}_x(\tilde{v}) = \inf\{\mathcal{Z}_x(v) : v \in \Gamma^{\mathcal{C}}_{FG}(c,u,\Phi)\}$.

Proof. Similarly as above, we can select a sequence $(v_n)_{n=1}^{\infty}$ of $\Gamma^{\mathcal{C}}_{FG}(c,u,,\Phi)$ such that $\alpha = \sup\{\mathcal{Z}_x(v) : v \in \Gamma^{\mathcal{C}}_{FG}(c,u,,\Phi)\} = \lim_{n\to\infty}\mathcal{Z}_x(v_n)$ for fixed $x \in \mathbb{R}^d$. By virtue of Theorem 6.6 of Chap. 6, for every $(s,x) \in [0,T)\times\mathbb{R}^d$, there exists a sequence $(X^n_{s,x})_{n=1}^{\infty}$ of $\mathcal{X}_{s,x}(F,G)$ such that

$$v_n(s,x) = E_n^{s,x}\left[\Phi(\tau_D^n, X^n_{s,x}(\tau_D)) \exp\left(-\int_s^{\tau_D^n} c(t,X^n_{s,x}(t))\mathrm{d}t\right)\right] - E_n\left\{\int_s^{\tau_D^n}\left[u(t,X^n_{s,x}(t)) \exp\left(-\int_s^{s+t} c(z,X^n_{s,x}(z))\mathrm{d}z\right)\right]\mathrm{d}t\right\}$$

for $n \geq 1$, where $\tau_D^n = \inf\{r \in (s,T] : X^n_{s,x}(r) \notin D\}$. By virtue of Theorem 4.1 of Chap. 4 and Theorem 2.3 of Chap. 1, there are an increasing subsequence $(n_k)_{k=1}^{\infty}$ of $(n)_{n=1}^{\infty}$, a probability space $(\tilde{\Omega},\tilde{\mathcal{F}},\tilde{P})$, and stochastic processes $\tilde{X}^{n_k}_{s,x}$ and $\tilde{X}_{s,x}$ on $(\tilde{\Omega},\tilde{\mathcal{F}},\tilde{P})$ such that $P(X^{n_k}_{s,x})^{-1} = P(\tilde{X}^{n_k}_{s,x})^{-1}$ for $k = 1,2,\ldots$ and $\sup_{s\le t\le T}|\tilde{X}^{n_k}_{s,x}(t) - \tilde{X}_{s,x}(t)| \to 0$ a.s. Hence by Theorem 5.2 of Chap. 4, it follows that

$$\begin{aligned} v_{n_k}(s,x) &= E^{s,x}_{n_k}\left[\Phi(\tau_D^{n_k}, X^{n_k}_{s,x}(\tau_D^{n_k})) \exp\left(-\int_s^{\tau_D^{n_k}} c(t,X^{n_k}_x(t))\mathrm{d}t\right)\right] \\ &\quad - E^{s,x}_{n_k}\left\{\int_s^{\tau_D^{n_k}}\left[u(t,X^{n_k}_{s,x}x(t)) \exp\left(-\int_s^{s+t} c(z,X^{n_k}_{s,x}(z))\mathrm{d}z\right)\right]\mathrm{d}t\right\} \\ &= \tilde{E}\left[\Phi(\tilde{\tau}_D^{n_k}, \tilde{X}^{n_k}_{s,x}(\tilde{\tau}_D^{n_k})) \exp\left(-\int_s^{\tilde{\tau}_D^{n_k}} c(t,\tilde{X}^{n_k}_{s,x}(t))\mathrm{d}t\right)\right] \\ &\quad - \tilde{E}\left\{\int_s^{\tilde{\tau}_D^{n_k}}\left[u(t,\tilde{X}^{n_k}_{s,x}(t)) \exp\left(-\int_s^{s+t} c(z,\tilde{X}^{n_k}_{s,x}(z))\mathrm{d}z\right)\right]\mathrm{d}t\right\} = \tilde{v}_{n_k}(s,x) \end{aligned}$$

for $(s,x) \in [0,T) \times D$ and $k \geq 1$, where $\tilde{\tau}_D^{n_k} = \inf\{r \in (s,T] : \tilde{X}_{s,x}^{n_k}(r) \notin D\}$. Therefore, by Lemma 10.1 of Chap. 1, Theorem 5.1 of Chap. 4, and the properties of the sequence $(\tilde{X}_{s,x}^{n_k})_{k=1}^{\infty}$, one obtains

$$\lim_{k\to\infty} \tilde{v}_{n_k}(s,x) = \tilde{E}\left[\Phi(\tilde{\tau}_D, \tilde{X}_{s,x}(\tilde{\tau}_D)) \exp\left(-\int_s^{\tilde{\tau}_D} c(t, \tilde{X}_x(t))\mathrm{d}t\right)\right] - \tilde{E}\left\{\int_s^{\tilde{\tau}_D}\left[u(t, \tilde{X}_{s,x}(t)) \exp\left(-\int_s^{s+t} c(z, \tilde{X}_{s,x}(z))\mathrm{d}z\right)\right]\mathrm{d}t\right\},$$

where $\tilde{\tau}_D = \inf\{r \in (s,T] : \tilde{X}_{s,x}(r) \notin D\}$. Immediately from Theorem 6.6 of Chap. 6, it follows that the function $\tilde{v}$ defined by

$$\tilde{v}(s,x) = \tilde{E}\left[\Phi(\tilde{\tau}_D, \tilde{X}_{s,x}(\tilde{\tau}_D)) \exp\left(-\int_s^{\tilde{\tau}_D} c(t, \tilde{X}_x(t))\mathrm{d}t\right)\right] - \tilde{E}\left\{\int_s^{\tilde{\tau}_D}\left[u(t, \tilde{X}_{s,x}(t)) \exp\left(-\int_s^{s+t} c(z, \tilde{X}_{s,x}(z))\mathrm{d}z\right)\right]\mathrm{d}t\right\}$$

belongs to $\Gamma_{FG}^{\mathcal{C}}(c,u,\Phi)$. Finally, similarly as above, we get $\alpha = \lim_{k\to\infty} \mathcal{Z}_x(\tilde{v}_{n_k}) = \mathcal{Z}_x(\tilde{v})$. □

In a similar way, we can also prove similar theorems for control systems described by set-valued stochastic Dirichlet, Poisson, and Dirichlet–Poisson problems. To formulate them, let us recall the basic notation dealing with such problems. Let $T > 0$ and let $D \subset \mathbb{R}^d$ be a nonempty bounded domain. Assume that $F : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ are measurable and bounded, and let $\Phi : (0,T) \times \partial D \to \mathbb{R}$, $\varphi : (0,T) \times D \to \mathbb{R}$ and $\psi : (0,T) \times D \to \mathbb{R}$ be continuous and bounded. Let $\mathcal{D}_{FG}(\Phi)$, $\mathcal{P}_{FG}(\varphi)$ and $\mathcal{R}_{FG}(\Phi,\psi)$ be defined by

$$\mathcal{D}_{FG}(\Phi) = \{u(s,x) = E^{s,x}[\Phi(\tau_D, X_{s,x}(\tau_D))] : X_{s,x} \in \mathcal{X}_{s,x}(F,G)\},$$

$$\mathcal{P}_{FG}(\varphi) = \left\{v(s,x) = E^{s,x}\left[\int_0^{\tau_D} \varphi(\tau_D, X_{s,x}(\tau_D))\right] : X_{s,x} \in \mathcal{X}_{s,x}(F,G)\right\},$$

and

$$\mathcal{R}_{FG}(\Phi,\psi) = \left\{w : w(s,x) = E^{s,x}[\Phi(\tau_D, X_{s,x}(\tau_D))] + E^{s,x}\left[\int_0^{\tau_D} \varphi(\tau_D, X_{s,x}(\tau_D))\right] : X_{s,x} \in \mathcal{X}_{s,x}(F,G)\right\}.$$

Immediately from Theorem 6.7, Theorem 6.8, and Theorem 6.9 of Chap. 6, it follows that $\mathcal{D}_{FG}(\Phi)$, $\mathcal{P}_{FG}(\varphi)$, and $\mathcal{R}_{FG}(\Phi,\psi)$ are subsets of the sets of all solutions of the following stochastic set-valued boundary value problems:

$$\begin{cases} 0 \in (\mathcal{L}_{FG}u)\,(t,x) \text{ for } (t,x) \in [0,T) \times D, \\ \lim_{t\to\tau_D} u(t, X_{s,x}(t) = \Phi(\tau_D, X_{s,x}(\tau_D)) \text{ for } (s,x) \in (0,T) \times D \quad a.s., \end{cases}$$

$$\begin{cases} -\varphi(s,x) \in (\mathcal{L}_{FG}v)\,(t,x) \text{ for } (t,x) \in [0,T) \times D, \\ \lim_{t\to\tau_D} v(t, X_{s,x}(t)) = 0)) \text{ for } (s,x) \in (0,T) \times D \quad a.s., \end{cases}$$

and

$$\begin{cases} -\varphi(s,x) \in (\mathcal{L}_{FG}w)\,(t,x) \text{ for } (t,x) \in [0,T) \times D, \\ \lim_{t\to\tau_D} w(t, X_{s,x}(t)) = \Phi(\tau_D, X_{s,x}(\tau_D)) \text{ for } (s,x) \in (0,T) \times D \quad a.s., \end{cases}$$

respectively. Similarly as above, we obtain the following results.

Theorem 3.3. *Assume that conditions* $(\mathcal{D})$ *are satisfied,* $T > 0$*, and* D *is a bounded domain in* $\mathbb{R}^d$*. Let* $\Phi \in C([0,T] \times \partial D, \mathbb{R})$ *be continuous and bounded. Assume that* $H : [0,T] \times \mathbb{R} \to \mathbb{R}$ *is measurable and uniformly integrally bounded such that* $H(t,\cdot)$ *is continuous. For every* $(s,x) \in (0,T) \times D$*, there is* $\tilde{X}_{s,x} \in \mathcal{X}_{s,x}(F,G)$ *such that the function* $\tilde{u}(s,x) = E^{s,x}[\Phi(\tilde{\tau}_D, \tilde{X}_{s,x}(\tilde{\tau}_D))]$ *satisfies* $\mathcal{Z}_x(\tilde{u}) = \sup\{\mathcal{Z}_x(u) : u \in \mathcal{D}_{FG}(\Phi)\}$*, where* $\tilde{\tau}_D = \inf\{r \in (0,T] : \tilde{X}_{s,x}(r) \notin D\}$*.*

Proof. Similarly as above, we can select a sequence $(u_n)_{n=1}^{\infty}$ of $\mathcal{D}_{FG}(\Phi)$ such that $\alpha = \sup\{\mathcal{Z}_x(u) : u \in \mathcal{D}_{FG}(\Phi)\} = \lim_{n\to\infty} \mathcal{Z}_x(u_n)$. By the definition of $\mathcal{D}_{FG}(\Phi)$, there exists a sequence $(X_x^n)_{n=1}^{\infty}$ of $\mathcal{X}_{s,x}(F,G)$ such that $u_n(s,x) = E^{s,x}[\Phi(\tau_D^n, X_{s,x}^n(\tau_D^n))]$, where $\tau_D^n = \inf\{r \in (0,T] : X_x^n(r) \notin D\}$. By virtue of Theorem 4.1 of Chap. 4 and Theorem 2.3 of Chap. 1, there are an increasing subsequence $(n_k)_{k=1}^{\infty}$ of $(n)_{n=1}^{\infty}$, a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$, and stochastic processes $\tilde{X}^{n_k}$ and $\tilde{X}$ on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ such that $P(X_{s,x}^{n_k})^{-1} = P(\tilde{X}^{n_k})^{-1}$ for $k = 1, 2, \ldots$ and $\sup_{0\le t\le T} |\tilde{X}^{n_k}(t) - \tilde{X}(t)| \to 0$ a.s. Hence, similarly as in the proof of Theorem 3.2, it follows that $\alpha = \lim_{n\to\infty} \mathcal{Z}_x(u_{n_k}) = \mathcal{Z}_x(\tilde{u})$, where $\tilde{u}(s,x) = E^{s,x}[\Phi(\tilde{\tau}_D, \tilde{X}_{s,x}(\tilde{\tau}_D))]$.□

Theorem 3.4. *Assume that conditions* $(\mathcal{D})$ *are satisfied,* $T > 0$*, and* D *is a bounded domain in* $\mathbb{R}^d$*. Let* $\varphi : (0,T) \times D \to \mathbb{R}$ *be continuous and bounded, and let* $H : [0,T] \times \mathbb{R} \to \mathbb{R}$ *be measurable and uniformly integrally bounded such that* $H(t,\cdot)$ *is continuous. For every* $(s,x) \in (0,T) \times D$*, there is* $\tilde{X}_{s,x} \in \mathcal{X}_{s,x}(F,G)$ *such that the function* $\tilde{u}(s,x) = E^{s,x}[\Phi(\tilde{\tau}_D, \tilde{X}_{s,x}(\tilde{\tau}_D))]$ *satisfies* $\mathcal{Z}_x(\tilde{u}) = \sup\{\mathcal{Z}_x(u) : u \in \mathcal{P}_{FG}(\varphi)\}$*, where* $\tilde{\tau}_D = \inf\{r \in (0,T] : \tilde{X}_{s,x}(r) \notin D\}$*.* □

Theorem 3.5. *Assume that conditions* $(\mathcal{D})$ *are satisfied,* $T > 0$*, and* D *is a bounded domain in* $\mathbb{R}^d$*. Let* $\Phi \in C((0,T) \times \partial D, \mathbb{R})$ *and* $\psi : (0,T) \times D \to \mathbb{R}$ *be continuous and bounded, and let* $H : [0,T] \times \mathbb{R} \to \mathbb{R}$ *be measurable and uniformly integrally* *bounded such that* $H(t,\cdot)$ *is continuous. For every* $(s,x)\in(0,T)\times D$, *there exists* $\tilde{X}_{s,x}\in\mathcal{X}_{s,x}(F,G)$ *such that the function* $\tilde{u}(s,x)=E^{s,x}[\Phi(\tilde{\tau}_D,\tilde{X}_{s,x}(\tilde{\tau}_D))]$ *satisfies* $\mathcal{Z}_x(\tilde{u})=\sup\{\mathcal{Z}_x(u):u\in\mathcal{R}_{FG}(\Phi,\psi)\}$, *where* $\tilde{\tau}_D=\inf\{r\in(0,T]:\tilde{X}_{s,x}(r)\notin D\}$. □

4 Notes and Remarks

The results of this chapter are consequences of the properties of the set $\mathcal{X}_{s,x}(F,G)$ of all (equivalence classes of) weak solutions for $SFI(F,G)$ and the representation theorems presented in Chap. 6. It is possible to consider problems with weaker assumptions. It is important to observe that such an approach reduces the optimal control problems described by stochastic functional and partial differential inclusions to the existence of optimal problems of functionals defined on weakly compact subsets of the space $\mathcal{M}(\mathcal{X})$ of probability measures defined on a Borel σ-algebra $\beta(\mathcal{X})$ of a complete metric space $\mathcal{X}$. Furthermore, this approach, together with representation theorems, leads to the representation of optimal solutions of the above type of optimal control problems by weak solutions of stochastic functional inclusions. This allows us in some special cases to determine explicit solutions of such optimal control problems. Some applications of weak solutions of multivalued stochastic equations to optimal control problems are given by A. Zălinescu in [97]. Some optimal control problems described by stochastic differential equations depending on control parameters can be solved explicitly by solving appropriate HJB equations. As pointed out (see B. Øksendal [86]) at the beginning of this chapter, some solutions of these equations can also be represented by weak solutions of stochastic differential equations. More information dealing with such problems can be found in B. Øksendal [86] and J. Yong and X.Y. Zhou [96].

Let us observe (see [45]) that there are three major approaches to stochastic optimal control: dynamic programming, duality, and the maximum principle. Dynamic programming obtains, by means of the optimality principle of Bellman, the Hamilton–Jacobi–Bellman equation, which characterizes the value function (see [28, 29, 37, 64, 98]). Under some smoothness and regularity assumptions on the solution, it is possible to obtain, at least implicitly, the optimal control. This is the content of the so-called verification theorem, which appears in W.H. Fleming and R.M. Rishel [28] or W.H. Fleming and H.M. Soner [29]. However, the problem of recovering the optimal control from the gradient of the value function by means of solving a static optimization remains, and this can be difficult to do. Duality methods, also known in stochastic control theory as the martingale approach, have become very popular in recent years, because they provide powerful tools for studying some classes of stochastic control problems, usually connected with some approximative procedures (see [73]). Martingale methods are particularly useful for problems appearing in finance (see [26]), such as the model of R.C. Merton [74]. Duality reduces the original problem to one of finite dimension. The approach is based on the martingale representation theorem and the Girsanov transformation.

The stochastic maximum principle has been developed completely in recent years in S. Peng [87]. It is a counterpart of the maximum principle for deterministic problems. The distinctive feature is the use of the concept of forward–backward stochastic differential equations, which arise naturally, governing the evolution of the state variables. See H.J. Kushner [67], J.M. Bismut [19,20], or U.G. Haussmann [36].

Control problems and optimal control problems for systems described by stochastic and partial differential equations have been considered by many authors. The classical optimal control problems for systems described by stochastic differential equations and inclusions were considered by, among others, N.A. Ahmed [1], A. Friedman [30], W.H. Fleming and M. Nisio [27], and M. Michta [75]. Optimal control problems for partial differential equations were considered by, for example, W. Huckbusch in [34]

References

1. Ahmed, N.U.: Optimal control of stochastic dynamical systems. Inf. Contr. **22**(1), 13–30 (1973)
2. Ahmed, N.U.: Nonlinear stochastic differential inclusions in Banach space. Stoch. Anal. Appl. **12**(1), 1–10 (1994)
3. Aitalioubrahim, M., Sajid, S.: Second order viability result in Banach spaces. Discuss. Math. Disc. Math. DICO 30, 5–21 (2010)
4. Alexiewicz, A.: Functional Analysis. P.W.N., Warszawa (1969) (in Polish)
5. Aubin, J.P., Cellina, A.: Differential Inclusions. Springer, New York (1984)
6. Aubin, J.P.: Viability Theory. Birkhäuser, Boston (1991)
7. Aubin, J.P., Da Prato, G.: Stochastic viability and invariance. Ann. Scouli Norm. Pisa **27**, 595–614 (1990)
8. Aubin, J.P., Da Prato, G.: Stochastic Nagumo's viability theorem. Stoch. Anal. Appl. **13**(1), 1–11 (1995)
9. Aubin, J.P., Da Prato, G.: The viability theorem for stochastic differential inclusions. Stoch. Anal. Appl. **16**, 1–15 (1998)
10. Aubin, J.P., Da Prato, G., Frankowska, H.: Stochastic invariance for differential inclusions. Set-valued Anal. **8**, 181–201 (2000)
11. Aubin, J.P., Frankowska, H.: Hyperbolic Systems of Partial Differential Inclusions, Ann. Scuoli Norm. Pisa **18**(4), 541–562 (1991)
12. Aubin, J.P., Frankowska, H.: Set-Valued Analysis. Birkhäuser, Boston (1990)
13. Aubin, J.P., Pujol, D., Saint-Pierre, P.: Dynamical Management of Portfolios Under Contingent Uncertainty. Cener Rech. Viabilité, Jeux, Contrôle (2000)
14. Aumann, R.: Integrals of set-valued functions. J. Math. Anal. Appl. **12**, 1–12 (1965)
15. Bartuzel, G., Fryszkowski, A.: Abstract differential inclusions with some applications to partial differential ones. Ann. Polon. Math. **LIII**, 67–78 (1991)
16. Bartuzel, G., Fryszkowski, A.: Stability of principal eigenvalue of schrödinger type problem for differential inclusions. Topol. Meth. Nonlinear Anal. **16**, 181–194 (2000)
17. Bartuzel, G., Fryszkowski, A.: A class of retracts in L^p with some applications to differential inclusions. Discuss. Math. DIO **22**, 213–224 (2002)
18. Benoit, T.V., Ha, T.X.D.: Existence of viable solutions for a nonconvex stochastic differential inclusion. Disc. Math. DI **17**, 107–131 (1997)
19. Bismut, J.M.: Conjugate convex functions in optimal stochastic control. J. Math. Anal. Appl. **44**, 484–504 (1973)
20. Bismut, J.M.: An introductory approach to duality in optimal stochastic control. SIAM Rev. **20**, 62–78 (1978)
21. Billingsley, P.: Convergence of Probability Measures. Wiley, New York (1976)

M. Kisielewicz, *Stochastic Differential Inclusions and Applications*, 275
Springer Optimization and Its Applications 80, DOI 10.1007/978-1-4614-6756-4,

22. Bocşan, G.: On Wiener stochastic integrals of multifunctions. Univ. Tim. FSN **87**, 1–7 (1987)
23. Da Prato, G., Frankowska, H.: A stochastic Fillipov theorem. Stoch. Anal.Appl. **12**, 408–426 (1994)
24. Dawidowski, M., Kubiaczyk, I.: An existence theorem for generalized hyperbolic equation $z''_{xy} \in F(x, y, z)$ in Banach space. Comm. Math. **30**, 10–25 (1990)
25. Dawidowski, M., Kisielewicz, M., Kubiaczyk, I.: Existence theorem for hyperbolic differential inclusions with Carathéodory right-hand side. Discuss. Math. **10**, 69–75 (1990)
26. Duffie, D.: Dynamic Asset Pricing Theory. Princeton University Press, Princeton (1996)
27. Fleming, W.H., Nisio, M.: On stochastic relaxed control for partially observed diffusion. Nagoya Math. J. **93**, 71–108 (1984)
28. Fleming, W.H., Rishel, R.M.: Deterministic and Stochastic Optimal Control. Springer, New York (1975)
29. Fleming, W.H., Soner, H.M.: Controlled Markov Processes and Viscosity Solutions. Springer, New York (1993)
30. Friedman, A.: Nonlinear optimal control problems to parabolic equations. SIAM J. Control. Optim. **22**(5), 805–816 (1984)
31. Friedman, A.: Stochastic Differential Equations and Applications. Academic, New York (1975)
32. Fryszkowski, A.: Carathéodory type selectors of set-valued maps of two variables. Bull. Acad. Polon. Sci. Sér. Sci Math. **26**, 41–46 (1977)
33. Gihman, I.I., Skorohod, A.V.: The Theory of Stochastic Processes I–III. Springer, Berlin (1974, 1975, 1979)
34. Hackbusch, W.: On the fast solving of parabolic boundary control problems. SIAM J. Control. Optim. **17**(2), 231–244 (1979)
35. Halmos, P.R.: Measure Theory. D. Van Nostrand Company Inc., New York (1956)
36. Haussmann, U.G.: A Stochastic Maximum Principle for Optimal Control Diffusion. Pitman Research Notes in Mathematics, vol. 151. Longman, Harlow (1986)
37. Hernández-Lerma, O.: Lectures on continuous-time Markov control processes. In: Aportaciones Matemáticas. Sociedad Matemática Mexicana, Mexico (1994)
38. Hiai, F.: Multivalued stochastic integrals and stochastic inclusions. Devision of Applied Mathematics, Research Institute of Applied Electricity, Sapporo 060. Japan (unpublished)
39. Hiai, F., Umegaki, H.: Integrals, conditional expectations and martingales of multivalued functions. J. Math. Anal. **7**, 149–182 (1977)
40. Hildenbrand, W.: Core and Equilibria of a Large Economy. Princ. Univ. Press, Princeton (1974)
41. Hu, S., Papageorgiou, N.S.: Handbook of Multivalued Analysis I. Kluwer Acad. Publ., Dordrecht (1997)
42. Hukuhara, M.: Intégration des applications mesurables dont la valeur est un compact convexe. Funk. Ekv. **10**, 205–223 (1967)
43. Ikeda, N., Watanabe, S.: Stochastic Differential Equations and Diffusion Processes. North Holland Publ., Amsterdam (1981)
44. Jacod, J., Shiryaev, A.N.: Limit Theorems for Stochastic Processes. Springer, New York (1987)
45. Josa–Fombellida, R., Rincón-Zapatero, J.P.: New approach to stochastic optimal control. J. Optim. Theory Appl. **135**, 163–177 (2007)
46. Jung, E.J., Kim, J.H.: On Set-valued Stochastic Integrals. Stoch. Anal. Appl. **21**, 401–418 (2003)
47. Karatzas, I., Shreve, S.E.: Brownian Motion and Stochastic Calculus. Springer, New York (1991)
48. Kim, B.K., Kim, J.H.: Stochastic integrals of set-valued processes and fuzzy processes. J. Math. Anal. Appl. **2336**, 480–502 (1999)
49. Kisielewicz, M.: Differential Inclusions and Optimal Control. Kluwer Acad. Publ., New York (1991)
50. Kisielewicz, M.: Properties of solution set of stochastic inclusions. J. Appl. Math. Stoch. Anal. **6**, 217–236 (1993)
51. Kisielewicz, M.: Set-valued stochastic integrals and stochastic inclusions. Discuss. Math. **13**, 119–126 (1993)

52. Kisielewicz, M.: Existence theorem for nonconvex stochastic inclusions. J. Appl. Math. Stoch. Anal. **7**, 151–15 (1994)
53. Kisielewicz, M.: Strong and weak solutions to stochastic inclusions. Banach Center Publ. **32**, 227–286 (1995)
54. Kisielewicz, M.: Viability theorem for stochastic differential inclusions. Discuss. Math. **16**, 61–74 (1995)
55. Kisielewicz, M.: Set-valued stochastic integrals and stochastic inclusions. Stoch. Anal. Appl. **15**(5), 783–800 (1997)
56. Kisielewicz, M.: Weak compactness of solution sets to stochastic differential inclusions with convex right hand sides. Topol. Math Nonlinear Anal. **18**, 149–169 (2001)
57. Kisielewicz, M.: Continuous selection theorems. Discuss. Math. Control. Optim. **25**, 159–163 (2005)
58. Kisielewicz, M.: Weak compactness of solution sets to stochastic differential inclusions with non-convex right hand sides. Stoch. Anal. Appl. **23**, 871–901 (2005)
59. Kisielewicz, M.: Backward stochastic differential inclusions. Dynam. Syst. Appl. **16**, 121–140 (2007)
60. Kisielewicz, M.: Stochastic differential inclusions and diffusion processes. J. Math. Anal. Appl. **334**, 1039–1054 (2007)
61. Kisielewicz, M.: Stochastic representation of partial differential inclusions. J. Math. Anal. Appl. **353**, 592–606 (2009)
62. Kisielewicz, M.: Some properties of set-valued stochastic integrals. J. Math. Anal. Appl. **388**, 984–995 (2012)
63. Klein, E., Thompson, A.C.: Theory of Correspondences. Wiley Interscience Publication, New York (1984)
64. Krylov, N.V.: Controlled Diffusion Processes. Springer, New York (1980)
65. Kubiaczyk, I.: An existence theorem for the generalized hyperbolic equation $z''_{xy} \in F(x, y, z)$ in Banach space. Comm. Math. **30**, 35–43 (1990)
66. Kucia, A., Nowak, A.: On Carathéodory type selectors of set-valued maps in Hilbert space. Ann. Math. Sil. **2**(14), 47–52 (1986)
67. Kuhner, H.J.: Necessary conditions for continuous parameter stochastic optimization problems. SIAM J. Control **10**, 550–565 (1972)
68. Kunita, H.: Stochastic Flows and Stochastic Differential Equations. Cambridge University Press, New York (1990)
69. Kuratowski, C.: Topologie, vol. I. PWN, Warszawa (1958)
70. Kuratowski, K.: Ryll-Nardzewski, C.: A general theorem on selectors. Bull. Polon. Acad. Sci. **13**, 397–403 (1965)
71. Levakov, A.A.: Stochastic differential inclusions. Diff. Uravn. **33**(2), 212–220 (1997)
72. Lipcer, R.S., Shiryaev, A.N.: Statistics of Stochastic Processes. PWN, Warszawa (1981) (in Polish)
73. Marti, K.: Approximative solutions of stochastic control problems by means of quasilinearization. In: Hamza, M.H. (ed.) Measurement and Control, pp. 183–188. ACTA Press, Anaheim (1977)
74. Merton, R.C.: Optimum consumption and portfolio rules in a continuous time model. J. Econ. Theory **3**, 373–413 (1971)
75. Michta, M.: Optimal solutions to stochastic differential inclusions. Appl. Math. (Warsaw) **29**(4), 387–398 (2002)
76. Michta, M.: On solutions to stochastic differential inclusions. Discr. Cont. Dynam. Syst. **140**, 618–622 (2003)
77. Michta, M.: On weak solutions to stochastic differential inclusions driven by semimartingales. Stoch. Anal. Appl. **22**(5), 1341–1361 (2004)
78. Michta, M., Motyl, J.: Compactness of solutions of second order dynamical systems. Dynam. Cont. Discr. Imp. Syst. Ser. A: Math. Anal. **14**(4), 525–545 (2007)
79. Millian, A.: A note on the stochastic invariance for Itô equation. Bull. Acad. Sci. Math. **41**(2), 139–150 (1993)

80. Mishina, A.P., Proskurjakov, I.W.: Higher Algebra. Nauka, Moskov (1965) (in Russian)
81. Motyl, J.: Note on strong solutions of stochastic inclusions. J. Math. Anal. Appl. **8**(3), 291–297 (1995)
82. Motyl, J.: Existence of solutions of set-valued Itô equation. Bull. PAN **46**(4), 419–430 (1998)
83. Motyl, J.: Stability problem for stochastic inclusions. Stoch. Anal. Appl. **16**, 933–944 (1998)
84. Motyl, J.: Stochastic functional inclusion driven by semimartingale. Stoch. Anal. Appl. **16**(3), 517–532 (1998)
85. Motyl, J.: Viable solutions to set-valued stochastic equations. Optimization **48**, 157–176 (2000)
86. Øksendal, B.: Stochastic Differential Equations. Springer, Berlin (1998)
87. Peng, S.: A general stochastic maximum principle for optimal control problems. SIAM J. Control. Optim. **28**, 966–979 (1990)
88. Plaskacz, S.: Value functions in control systems and differential games: A viability method. Lecture Notes in Nonlinear Analysis. Journal of Schauder Center for Nonlinear Analysis, N.C.University (2003)
89. Protter, P.: Stochastic Integration and Differential Equations. Springer, Berlin (1990)
90. Przesławski, K.: Lipschitz continuous selectors. J. Convex Anal. **5**(2), 249–267 (1998)
91. Rybiński, L.: On Carathéodory type selections. Fund. Math. **125**, 187–193 (1985)
92. Sosulski, W.: Properties of Solutions Set to Generalized Partial Differential Equations of Hyperbolic Type. Technical University of Zielona Góra Press, Zielona Gora (1982) (in Polish)
93. Sosulski, W.: Existence of optimal solutions for some problems described by neutral partial functional-differential inclusions of hyperbolic type. Discuss. Math. **10**, 123–130 (1990)
94. Strook, D.W., Varadhan, S.R.S.: Diffusion processes with continuous coefficients. I, II. Comm. Pure Appl. Math. **22**(345–400), 479–530 (1969)
95. Strook, D.W., Varadhan, S.R.S.: Multidimensional Diffusion Processes. Springer, Berlin (1997)
96. Yong, J., Zhou, X.Y.: Stochastic Controls: Hamiltonian Systems and HJB Equations. Springer, New York (1999)
97. Zălinescu, A.: Weak solutions and optimal control for multivalued stochastic differential equations. Nonlinear Diff. Equat. Appl. **15**, 511–533 (2008)
98. Zhang, J., Li, S., Mitoma, I., Okazaki, Y.: On set-valued stochastic integrals in an M-type 2 Banach space. J. Math. Anal. Appl. **350**, 216–233 (2009)

Index

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