Stochastic differential inclusions and diffusion processes

Michał Kisielewicz

Faculty of Mathematics, Computer Science and Econometrics, University of Zielona Góra, Podgórna 50, 65-246 Zielona Góra, Poland
Received 1 February 2006
Available online 23 January 2007
Submitted by H. Frankowska

Abstract
Connections between weak solutions of stochastic differential inclusions and solutions of partial differential inclusions, generated by given set-valued mappings are considered. The main results are based on some continuous approximation selection theorem and weak compactness of the set of all weak solutions to a given stochastic differential inclusion.
© 2007 Elsevier Inc. All rights reserved.

Keywords: Stochastic differential inclusions; Diffusion processes

1. Introduction

The main results of the paper deal with some relations between weak solutions of stochastic differential inclusions and solutions of partial differential inclusions generated by given set-valued mappings. The first papers concerning general stochastic differential inclusions are due to F. Hiai [5] and M. Kisielewicz [7,8], where independently stochastic differential inclusions of the form

\[ x_t - x_s \in \text{cl}_{L^2} \left( \int_s^t F(\tau, x_\tau) \, d\tau + \int_s^t G(\tau, x_\tau) \, dB_\tau \right) \] (1)
have been investigated. They have to be satisfied by an $L^2$-continuous $\mathcal{F}_t$-nonanticipative stochastic process $(x_t)_{0 \leq t \leq T}$ for every $0 \leq s < t \leq T$, i.e. by an $\mathcal{F}_t$-nonanticipative square integrable process $(x_t)_{0 \leq t \leq T}$ that is continuous with respect to the norm topology of the space $L^2(\Omega, \mathcal{F}, \mathbb{P})$. Such inclusions were considered on a given complete filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbb{P})$ satisfying the usual hypotheses, i.e. with a filtration $(\mathcal{F}_t)_{0 \leq t \leq T}$ such that $\mathcal{F}_0$ contains all $\mathbb{P}$-null sets of $\mathcal{F}$ and $\mathcal{F}_t = \bigcap_{t' \geq t} \mathcal{F}_{t'+\varepsilon}$. Apart from the set-valued mappings $F : [0, T] \times \mathbb{R}^n \to \text{Cl}(\mathbb{R}^n)$ and $G : [0, T] \times \mathbb{R}^n \to \text{Cl}(\mathbb{R}^{n \times m})$ or with values in $\text{Cl}(H)$, where $H$ is a Hilbert space, some $\mathcal{F}_t$-Brownian motions $(B_t)_{0 \leq t \leq T}$ and $(W_t)_{0 \leq t \leq T}$ with values at $\mathbb{R}^m$ or $H$, respectively, also have been given. As usual $\text{Cl}(X)$ denotes the space of all nonempty closed subsets of a metric space $(X, \rho)$. Similarly as in the theory of stochastic differential equations, the process $(x_t)_{0 \leq t \leq T}$ mentioned above, is said to be a strong solution to (1). Such solutions have been considered by N.U. Ahmed [1], J.P. Aubin and G. Da Prato [2], G. Da Prato and H. Frankowska [3], J. Motyl [13–15] and others. For the existence of strong solutions some Lipschitz type continuity for $F(t, \cdot)$ and $G(t, \cdot)$ is needed. Such type assumptions are often too strong for practical applications. For instance in optimal control problems we have to deal with multifunctions defined by $F(t, x) = \{ f(t, x, u) : u \in U \}$ and $G(t, x) = \{ g(t, x, u) : u \in U \}$, where $U$ is a metric space whereas $f$ and $g$ are given functions on $[0, T] \times \mathbb{R}^n \times U$ with values in $\mathbb{R}^n$ and $\mathbb{R}^{n \times m}$, respectively. Usually such multifunctions are measurable and continuous if $f$ and $g$ have such properties. In general they are not Lipschitz continuous, even if $f$ and $g$ are Lipschitz continuous. Therefore, we are interested in a weaker notion of solutions that are not restrictive in the existence theory and are extremely useful and fruitful in both theory and applications. Such type of solutions is known as weak ones (see [7,10]). In the present paper we shall consider inclusions of solutions is known as weak ones (see [7,10]). In the present paper we shall consider inclusions (1) with $F$ and $G$ taking their values in the spaces $\text{Conv}(\mathbb{R}^n)$ and $\text{Conv}(\mathbb{R}^{n \times m})$ of all nonempty compact convex subsets of $\mathbb{R}^n$ and $\mathbb{R}^{n \times m}$, respectively. In such a case (see [8]) the stochastic differential inclusion (1) has the form

$$
\int_s^t F(\tau, x_\tau) d\tau + \int_s^t G(\tau, x_\tau) dB_\tau
$$

and its solution $(x_t)_{0 \leq t \leq T}$ is a continuous process on $(\Omega, \mathcal{F}, \mathbb{P})$. A weak solution to (2) is understood as a system including a complete filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbb{P})$ satisfying the usual hypotheses, an $\mathcal{F}_t$-Brownian motion $(B_t)_{0 \leq t \leq T}$ and an $\mathcal{F}_t$-nonanticipative continuous processes $(x_t)_{0 \leq t \leq T}$ satisfying the relation (2) when $F$ and $G$ are given. In what follows we shall identify such system with the pair $(x, B)$ of process $x = (x_t)_{0 \leq t \leq T}$ and $B = (B_t)_{0 \leq t \leq T}$ or simply by a process $(x_t)_{0 \leq t \leq T}$ defined on the above filtered probability space. We will say that $x = (x_t)_{0 \leq t \leq T}$ is a weak solution to (2) on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbb{P})$. If $(x_t)_{0 \leq t \leq T}$ is such that $P_{x_0^{-1}} = \mu$, where $\mu$ is a given probability measure on the Borel $\sigma$-algebra $\beta(\mathbb{R}^n)$ and $P_{x_0^{-1}}$ denotes the distribution of $x_0$ then we say that $x = (x_t)_{0 \leq t \leq T}$ is a weak solution to (2) on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbb{P})$ with an initial distribution $\mu$ on $\beta(\mathbb{R}^n)$.

It was proved in [7] that for the existence of weak solutions to (2) with a given initial distribution it is enough to assume that $F$ and $G$ are Borel measurable, bounded, convex-valued and such that $F(t, \cdot)$ and $G(t, \cdot)$ are lower semicontinuous for fixed $t \in [0, T]$. Stochastic differential inclusions considered in [7] and [8] are defined for one-dimensional Brownian motions. In the present paper we shall consider the general case with $m$-dimensional Brownian motions. Therefore $G$ has to take its values from the space $\text{Cl}(\mathbb{R}^{n \times m})$, where $\mathbb{R}^{n \times m}$ denotes the space of all $n \times m$-type matrices. We shall consider $\mathbb{R}^{n \times m}$ as a normed space with the norm defined by $\|g\| = (\sum_{i=1}^n \sum_{j=1}^m s_{ij}^2)^{1/2}$ for $g = (g_{ij})_{n \times m}$. In particular $\mathbb{R}^{n \times 1}$ is simply denoted by $\mathbb{R}^n$. 
Throughout the paper we assume that $F$ and $G$ are convex-valued or that $G$ is diagonally convex-valued (see [10]), i.e. that the set $\{g: g^T \in G(t, x)\}$ is convex for every fixed $(t, x) \in [0, T] \times \mathbb{R}^n$, where $g^T$ denotes the transposition of $g$. It can be verified (see [10, Proposition 2]) that a set-valued mapping $G$ with convex values in $\mathbb{R}^{1 \times m}$ is diagonally convex.

Finally, let us recall that a continuous $n$-dimensional stochastic process $x = (x_t)_{0 \leq t \leq T}$ on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, P)$ can be equivalently defined as an $(\mathcal{F}, \beta(C^n_T))$-measurable random function $x: \Omega \rightarrow C^n_T$, where $C^n_T = C([0, T], \mathbb{R}^n)$ and $\beta(C^n_T)$ denotes the Borel $\sigma$-algebra on $C^n_T$. Such defined continuous process determines on $\beta(C^n_T)$ its distribution denoted by $P x^{-1}$ and understood as a probability measure on $\beta(C^n_T)$ of the form $(P x^{-1})(A) = P(x^{-1}(A))$ for every $A \in \beta(C^n_T)$, where $x^{-1}(A) = \{\omega \in \Omega: X(\omega) \in A\}$. It admits the definition of the convergence in distribution of sequences of continuous processes called also a weak convergence. It is well known (see [6]) that a sequence $(x^r)_{r=1}^{\infty}$ of continuous processes $x^r$ defined on probability spaces $(\Omega^r, \mathcal{F}^r, P^r)$ with $r = 1, 2, \ldots$, weakly converges to a continuous process $x$ defined on $(\Omega, \mathcal{F}, P)$ if \( \lim_{r \to \infty} E^r f(x^r) = Ef(x) \) for every continuous bounded function $f: C^n_T \rightarrow \mathbb{R}$, where $E^r$ and $E$ denote the expectations with respect to $P^r$ and $P$, respectively. In particular, for continuous processes $x$ and $\tilde{x}$ on $(\Omega, \mathcal{F}, P)$ and $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$, respectively such that $P x^{-1} = \tilde{P} \tilde{x}^{-1}$ we have $Ef(x) = \tilde{E}f(\tilde{x})$ for every continuous bounded function $f: C^n_T \rightarrow \mathbb{R}$.

2. Weak compactness of weak solutions set of stochastic differential inclusions and continuous selection theorem

For given $F: [0, T] \times \mathbb{R}^n \rightarrow Cl(\mathbb{R}^n)$ and $G: [0, T] \times \mathbb{R}^n \rightarrow Cl(\mathbb{R}^{n \times m})$ and a probability measure $\mu$ on $\beta(\mathbb{R}^n)$ we denote by $\mathcal{X}_\mu(F, G)$ the set of all weak solutions of (2) with an initial distribution $\mu$. A sequence $(x^r)_{r=1}^{\infty}$ of $\mathcal{X}_\mu(F, G)$ is said to be convergent in distribution if there is a probability measure $\mathcal{P}$ on $\beta(C^n_T)$ such that $P(x^r)^{-1} \rightarrow \mathcal{P}$ weakly in the space $\mathcal{M}(C^n_T)$ of all probability measures on $\beta(C^n_T)$ as $r \to \infty$. It was proved in [10] that for every bounded compact convex-valued mappings $F$ and $G$ satisfying Carathéodory conditions and such that $G$ is diagonally convex-valued the set $\mathcal{X}_\mu(F, G)$ is nonempty and sequentially weakly closed with respect to the convergence in distributions. We shall verify that by the above assumptions $\mathcal{X}_\mu(F, G)$ is also sequentially weakly compact with respect to the convergence in distribution. Such result for the martingal problem has been proved in [11].

**Theorem 1.** Let $F: [0, T] \times \mathbb{R}^n \rightarrow Cl(\mathbb{R}^n)$ and $G: [0, T] \times \mathbb{R}^n \rightarrow Cl(\mathbb{R}^{n \times m})$ be bounded measurable and convex-valued mappings such that $F(t, \cdot)$ and $G(t, \cdot)$ are continuous for fixed $t \in [0, T]$. Assume $G$ is diagonally convex-valued. Then for every probability measure $\mu$ on $\beta(\mathbb{R}^n)$ the set $\mathcal{X}_\mu(F, G)$ is nonempty and sequentially weakly compact with respect to the convergence in distribution.

**Proof.** By [10, Theorem 12] the set $\mathcal{X}_\mu(F, G)$ is nonempty and sequentially weakly closed with respect to the convergence in distribution. Let $(x^r, B^r)_{r=1}^{\infty}$ be a sequence of $\mathcal{X}_\mu(F, G)$ on filtered probability spaces $(\Omega^r, \mathcal{F}^r, (\mathcal{F}_t^r)_{0 \leq t \leq T}, P^r)$ such that $(x^r, B^r)$ satisfies (2) for every $r = 1, 2, \ldots$. By [8, Theorem 4] for every $r = 1, 2, \ldots$ there are $f^r \in S(F \circ x^r)$ and $g^r \in S(G \circ x^r)$ such that $dx^r_t = f^r_t \, dt + g^r_t \, dB^r_t$ for $t \in [0, T]$. Similarly as in [6, Theorems IV.2.2 and 1.4.2] we can prove that there are a subsequence $(r_k)_{k=1}^{\infty}$ of $(r)_{k=1}^{\infty}$, a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mu})$ and a continuous process $\tilde{x}$ on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mu})$ such that $P(x^r_k)^{-1} \rightarrow \tilde{P} \tilde{x}^{-1}$ weakly in $\mathcal{M}(C^n_T)$ as $k \to \infty$. Hence, by the sequential weak closedness of $\mathcal{X}_\mu(F, G)$ it follows that $\tilde{x} \in \mathcal{X}_\mu(F, G)$.  \(\square\)
In what follows we shall need some continuous approximation selection theorem. Let \((X, \rho), (Y, | \cdot |)\) and \((Z, \| \cdot \|)\) be Polish and Banach spaces, respectively and denote by \(\text{Cl}(Y)\) a family of all nonempty closed subsets of \(Y\). Recall that a set-valued mapping \(F : X \to \mathcal{P}(Y)\), where \(\mathcal{P}(Y)\) denotes a family of all nonempty subsets of \(Y\) is said to be lower semicontinuous (l.s.c.) at \(\bar{x} \in X\) if for every open set \(U\) in \(Y\) with \(F(\bar{x}) \cap U \neq \emptyset\) there is a neighbourhood \(V_{\bar{x}}\) of \(\bar{x}\) such that \(F(x) \cap U \neq \emptyset\) for every \(x \in V_{\bar{x}}\).

**Lemma 2.** Let \(\lambda : X \times Y \to Z\) and \(u : X \to Z\) be continuous and let \(F : X \to \mathcal{P}(Y)\) be lower semicontinuous such that \(u(x) \in \lambda(x, F(x))\) for \(x \in X\). Then for every lower semicontinuous function \(\varepsilon : X \to (0, \infty)\) a set-valued mapping \(\Phi : X \to \mathcal{P}(Y)\) defined by

\[
\Phi(x) = F(x) \cap \left\{ u \in Y : \|\lambda(x, u) - u(x)\| < \varepsilon(x) \right\}
\]

is lower semicontinuous on \(X\).

**Proof.** Let \(\eta > 0\) be given and let \((\bar{x}, \bar{u}) \in \text{Graph}(\Phi)\). There is \(\sigma > 0\) such that \(\|\lambda(\bar{x}, \bar{u}) - u(\bar{x})\| = \varepsilon(\bar{x}) - \sigma\). There is \(\delta > 0\) 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 is \(y_x \in F(x)\) such that \(\|y_x - \bar{u}\| < \min(\eta, 1/3\sigma, \delta)\). There is \(\sigma_2 > 0\) such that \(\|u(x) - u(\bar{x})\| < 1/3\sigma\) for every \(x \in X\) satisfying \(\rho(x, \bar{x}) < \sigma_2\). Furthermore by the lower semicontinuity of \(\varepsilon\) there is \(\sigma_3 > 0\) such that \(\varepsilon(x) > \varepsilon(\bar{x}) - 1/3\sigma\) for every \(x \in X\) such that \(\rho(x, \bar{x}) < \sigma_3\).

Now for every \(x \in X\) satisfying \(\rho(x, \bar{x}) < \min(\delta, \sigma_1, \sigma_2, \sigma_3)\) one gets

\[
\|\lambda(x, y_x) - u(x)\| \leq \|\lambda(x, y_x) - \lambda(\bar{x}, \bar{u})\| + \|\lambda(\bar{x}, \bar{u}) - u(\bar{x})\| + \|u(\bar{x}) - u(x)\| < 1/3\sigma + \varepsilon(\bar{x}) - \sigma + 1/3\sigma < \varepsilon(x).
\]

Then \(y_x \in \Phi(x)\) and \(\|y_x - \bar{u}\| < \eta\). Let us observe now that for every open set \(U \subset Y\) such that \(U \cap \Phi(\bar{x}) \neq \emptyset\) there are \(\bar{u} \in \Phi(\bar{x})\) and \(\eta > 0\) such that \((\bar{u} + \eta B^0) \subset U\). To such \(\bar{u} \in \Phi(\bar{x})\) and \(\eta > 0\) we can select \(\bar{\varepsilon} = \min(\delta, \sigma_1, \sigma_2, \sigma_3)\) such that \((\bar{u} + \eta B^0) \cap \Phi(x) \neq \emptyset\) for \(x \in (\bar{x} + \bar{\varepsilon} B^0)\). Therefore for every open set \(U \subset Y\) such that \(U \cap \Phi(\bar{x}) \neq \emptyset\) there is \(\bar{\varepsilon} > 0\) such that \(U \cap \Phi(x) \supset (\bar{u} + \eta B^0) \cap \Phi(x) \neq \emptyset\) for every \(x \in (\bar{x} + \bar{\varepsilon} B^0)\).

Now we can prove the following selection theorem.

**Theorem 3.** Let \(\lambda : X \times Y \to Z\) and \(u : X \to Z\) be continuous and \(F : X \to \text{Cl}(Y)\) be l.s.c. such that \(u(x) \in \lambda(x, F(x))\) for \(x \in X\). Assume \(\lambda(x, \cdot)\) is affine and \(F(x)\) are convex subsets of \(Y\) for fixed \(x \in X\). Then for every \(\varepsilon > 0\) there is a continuous function \(f_\varepsilon : X \to Y\) such that \(f_\varepsilon(x) \in F(x)\) and \(\|\lambda(x, f_\varepsilon(x)) - u(x)\| \leq \varepsilon\) for \(x \in X\).

**Proof.** By virtue of Lemma 2 for every \(\varepsilon > 0\) a set-valued mapping \(\Phi_\varepsilon : X \to \mathcal{P}(Y)\) defined by (3) with \(\varepsilon(x) \equiv \varepsilon\) is l.s.c. Therefore the set-valued mapping \(\text{cl}(\Phi_\varepsilon)\) is also l.s.c. on \(X\). Furthermore by the convexity of \(F(x)\) and the properties of \(\lambda(x, \cdot)\) it follows that \(\Phi_\varepsilon(x)\) and \(\text{cl}(\Phi_\varepsilon)(x)\) are convex subsets of \(Y\) for every \(x \in X\). Hence by Michael’s continuous selection theorem (see [9, Theorem II.4.1]) there is a continuous selector \(f_\varepsilon\) for \(\text{cl}(\Phi_\varepsilon)\). It is clear that \(f_\varepsilon\) satisfies conditions our theorem.
3. Diffusion processes generated by differential operators

Denote by $C^2_0(\mathbb{R}^d)$ the space of all continuous functions $h: \mathbb{R}^d \to \mathbb{R}$ with continuous derivatives up to the order two with compact supports in $\mathbb{R}^d$. For given functions $f: \mathbb{R}^d \to \mathbb{R}^d$ and $g: \mathbb{R}^d \to \mathbb{R}^{d \times m}$ let us define on $C^2_0(\mathbb{R}^d)$ the differential operator $A_{fg}$ of the form:

$$(A_{fg}h)(x) = \{\partial_t h(x), f(x)\} + \frac{1}{2} \text{tr}(\partial_{xx} h(x) \cdot (g \cdot g^T)(x))$$

for $x \in \mathbb{R}^d$ and $h \in C^2_0(\mathbb{R}^d)$, where $\partial_x = (\partial_{x_1}, \ldots, \partial_{x_d})$ and $\partial_{xx} = (\partial_{x_ix_j})_{1 \leq i \leq d, 1 \leq j \leq d}$.

Let us observe that if $m = 1$ then we can define $(A_{fg}h)(x)$ only for $f, g, h: \mathbb{R}^1 \to \mathbb{R}^1$ and $(A_{fg}h)(x) = f(x)h'(x) + \frac{1}{2}g^2h''(x)$. Therefore the theory of diffusion processes defined by stochastic differential equations with multidimensional noises is technically much more complicated than with one-dimensional Brownian motions. Let $C^d_T = C([0, T], \mathbb{R}^d)$ and let $\beta(C^d_T)$ be the Borel $\sigma$-algebra on $C^d_T$.

The diffusion measures determined by the operator $A_{fg}$ or simply the $A_{fg}$-diffusion is defined as a system $\{P_x: x \in \mathbb{R}^d\}$ of the strong Markov’s probability distributions on $(C^d_T, \beta(C^d_T))$ satisfying the following conditions:

(I) $P_x(\{w \in C^d_T: w(0) = x\}) = 1$ for $x \in \mathbb{R}^d$,

(II) $h(w(t)) - h(w(0)) - f^i_t(A_{fg}h)(w(s)) ds$ is a $(P_x, \beta_1(C^d_T))$-martingale for every $h \in C^2_0(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$, where $\beta_1(C^d_T) = \sigma(\bigcup_{w \in C^d_T} \{w(s): 0 \leq s \leq t\})$.

We say that an $A_{fg}$-diffusion satisfies the uniqueness condition if for every system $\{P_x': x \in \mathbb{R}^d\}$ of the strong Markov’s probability distributions on $(C^d_T, \beta(C^d_T))$ satisfying conditions (I) and (II) we have $P_x' = P_x$ for $x \in \mathbb{R}^d$.

A stochastic process $X: \Omega \to C^d_T$ on $(\Omega, \mathcal{F}, P)$ is said to be $A_{fg}$-diffusion process if its distribution $PX^{-1}$ satisfies

$$(PX^{-1})(\cdot) = \int_{\mathbb{R}^d} P_x(\cdot) \mu(dx),$$

where $\{P_x: x \in \mathbb{R}^d\}$ is $A_{fg}$-diffusion and $\mu$ is the probability distribution of $X(0)$.

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 consider a stochastic differential equation

$$dx_i = f(x_i) dt + g(x_i) dB_t.$$  \hspace{1cm} (4)

Similarly as it was mentioned above by a weak solution to (4) we mean a system containing a continuous $\mathcal{F}_t$-adapted stochastic process $X$, an $m$-dimensional $\mathcal{F}_t$-Brownian motion $B = (B_t)_{0 \leq t \leq T}$ on a complete filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, P)$ satisfying the usual hypothesis and such that the pair $(X, B)$ satisfies (4). It will be still denoted by $X$. We will say that (4) has a weak uniqueness property if for every its weak solutions $X$ and $X'$ defined on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, P)$ and $(\Omega', \mathcal{F}', (\mathcal{F}_t')_{0 \leq t \leq T}, P')$, respectively with the same initial distributions we have $PX^{-1} = P(X')^{-1}$.

Given $d \times m$-matrix $g(x)$ for $x \in \mathbb{R}^d$ by $\lambda_i(g \cdot g^T)(x)$ we denote the eigenvalues of $(g \cdot g^T)(x)$ for $i = 1, 2, \ldots, d$. We say that $g$ generates a uniformly positive symmetric matrix if $\sup_{x \in \mathbb{R}^d} \inf_{1 \leq i \leq d} \lambda_i(g \cdot g^T)(x) > 0$. In such a case we will also say that $g$ is the uniformly
positive generator. It can be proved [6, Theorem IV.6.1] that if \( f \) and \( g \) are continuous and bounded and \( g \) is the uniformly positive generator then a family \( \{ P_x : x \in \mathbb{R}^d \} \) of distributions of a family \( \{ X_t : t \in [0, T] \} \) of weak solutions to (4) satisfying \( X_t(0) = x \) with (P.1) is an \( \mathcal{A}_{fg} \)-diffusion. Hence in particular it follows that for every \( x \in \mathbb{R}^d \) the weak solution \( X_t \) to (4) satisfying \( X_t(0) = x \) with (P.1) is a time-homogeneous Itô diffusion process with the generator \( \mathcal{A}_{fg} \) satisfying

\[
(\mathcal{A}_{fg} h)(x) = \lim_{t \searrow 0} \frac{E h(X_t(t)) - h(x)}{t}
\]

for \( t \in (0, T] \) and every \( h : \mathbb{R}^d \to \mathbb{R} \) such that the above limit exists for all \( x \in \mathbb{R}^d \). Therefore the backward Kolmogorov’s theorem (see [16, Theorem 8.1.1]) is also true for such diffusions. The above result is also true (see [4, Theorem III.7]) for continuous mappings \( f : \mathbb{R}^d \to \mathbb{R}^d \) and \( g : \mathbb{R}^d \to \mathbb{R}^{d \times d} \) such that \( g^{-1} \) is locally bounded and there is \( C > 0 \) such that

\[
\| f(x) \| + \| g(x) \| < C (1 + \| x \|),
\]

and

\[
\limsup_{x \to x_0} \text{tr} (l(g)(x) - l(g)(x_0))^2 < \| l(g^{-1})(x_0) \|,
\]

for every \( x \in \mathbb{R}^d \) and \( x_0 \in \mathbb{R}^d \), where \( l(v) = v \cdot v^T \) for \( v \in \mathbb{R}^{d \times d} \).

Let us observe that solutions of non-autonomous stochastic differential equations are not time-homogeneous processes (see [16, p. 108]) and therefore are not strong Markov. However with an extra argument, by extending their state spaces, we can reduce such equations to the autonomous case. In what follows we apply such procedure to weak solutions of non-autonomous stochastic differential equations. Assume \( b : [0, T] \times \mathbb{R}^n \to \mathbb{R}^n \) and \( \sigma : [0, T] \times \mathbb{R}^n \to \mathbb{R}^{n \times m} \) are continuous and bounded and let \( f = (1, \beta(T)) \) and \( g = (0, \sigma_1, \ldots, \sigma_n)^T \) with \( 0, \sigma_i \in \mathbb{R}^{1 \times m} \), where \( 0 = (0, \ldots, 0) \) and \( \sigma_i \) denotes the \( i \)th row of \( \sigma \) for \( i = 1, \ldots, n \). If \( \sigma \) is a uniformly positive generator then by virtue of [6, Theorem IV.2.2] and the Strook and Varadhan uniqueness theorem (see [17]) for every \( x \in \mathbb{R}^n \) there is exactly one weak solution \( X_t \) to

\[
dx_t = b(t, x_t) dt + \sigma(t, x_t) \, dB_t \tag{5}
\]

satisfying \( X_t(0) = x \) with (P.1). Therefore for every \( x \in \mathbb{R}^n \) there is exactly one weak solution \( Y_{(0,x)}(t) = (t, X_t(t))^T \) to (4) with \( f \) and \( g \) defined above such that \( Y_{(0,x)}(0) = (0, x) \) with (P.1). Indeed, it is easy to see that \( Y_{(0,x)} \) satisfies (4) with \( f \) and \( g \) defined above. To verify the uniqueness in law of \( Y_{(0,x)} \) let us assume that \( X_t \) and \( \widetilde{X}_t \) are solutions to (5) with Brownian motions \( B \) and \( \tilde{B} \) on probability spaces \( (\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}), \mathbb{P} \) and \( (\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_t)_{0 \leq t \leq T}), \tilde{\mathbb{P}} \), respectively. By the Strook and Varadhan uniqueness theorem we have \( \mathbb{P} X_t^{-1} = \tilde{\mathbb{P}} \tilde{X}_t^{-1} \) on \( \beta(C^2_0) \) that is equivalent to \( P(X_t(t_1), \ldots, X_t(t_r))^{-1} = \tilde{P}(\tilde{X}_t(t_1), \ldots, \tilde{X}_t(t_r))^{-1} \) on \( \beta(C^{n \times n}_0) \) for every \( 0 \leq t_1 < \cdots < t_r \leq T \). Let \( Y_{(0,x)}(t) = (t, X_t(t))^T \) and \( \tilde{Y}_{(0,x)}(t) = (t, \tilde{X}_t(t))^T \) and put \( Q = \{ A \times B : A \in \beta([0, T)^r], B \in \beta(C^{n 	imes n}_0) \} \). It is clear that \( Q \) is \( \pi \)-system such that \( \beta([0, T)^r \times C^{n \times n}_0) = \sigma(Q) \). For every \( 0 \leq t_1 < \cdots < t_r \leq T \) and \( A \times B \in Q \) one has \( P(Y_{(0,x)}(t_1), \ldots, Y_{(0,x)}(t_r))^{-1} (A \times B) = 1_A(t_1) \ldots t_r) P(X_t(t_1), \ldots, X_t(t_r))^{-1} (B) = 1_A(t_1) \ldots t_r) \tilde{P}(\tilde{X}_t(t_1), \ldots, \tilde{X}_t(t_r))^{-1} (B) = \tilde{P}(\tilde{Y}_{(0,x)}(t_1), \ldots, \tilde{Y}_{(0,x)}(t_r))^{-1} (A \times B) \). Hence by Dynkin’s theorem it follows \( P(Y_{(0,x)}(t_1), \ldots, Y_{(0,x)}(t_r))^{-1} = \tilde{P}(\tilde{Y}_{(0,x)}(t_1), \ldots, \tilde{Y}_{(0,x)}(t_r))^{-1} \) on \( \sigma(Q) \) for every \( 0 \leq t_1 < \cdots < t_r \leq T \) that is equivalent to \( P Y_{(0,x)}^{-1} = \tilde{P} \tilde{Y}_{(0,x)}^{-1} \) on \( \beta(C_1^{n \times n}) \). Therefore \( Y_{(0,x)} \) is a time-homogeneous Itô diffusion with the generator \( \mathcal{A}_{fg} \). Then by virtue of [16, Theorem 8.1.1] it follows that for every \( \tilde{h} \in C_0^\infty(\mathbb{R}^{1+n}) \) the function \( U(t, (0, x)) = E \tilde{h}(Y_{(0,x)}(t)) \) belongs to \( C^{1,2}([0, T] \times \mathbb{R}^{n+1}, \mathbb{R}) \) and
satisfies $U'_t(t, (0, x)) = (A_{fg} U(t, \cdot))(t, x)$ with $U(0, (0, x)) = \hat{h}(0, x)$ for $t \in [0, T]$ and $x \in \mathbb{R}^n$. Thus the function $u(t, x) = U(t, (0, x))$ satisfies $u'_t(t, x) = (A_{fg} u(t, \cdot))(t, x)$, $u(0, x) = \hat{h}(0, x)$ and $u(t, x) = E\hat{h}((t, X^T_t(t)))$ for $(t, x) \in [0, T] \times \mathbb{R}^n$. Taking $\hat{h} = h \circ \Pi_{\mathbb{R}^n}$ for $h \in C^2_0(\mathbb{R}^n)$, where $\Pi_{\mathbb{R}^n}$ denotes the orthogonal projection of $\mathbb{R}^{1+n}$ onto $\mathbb{R}^n$ we obtain $u(t, x) = Eh(X_t(t))$, $u'_t(t, x) = (A_{fg} h(t, \cdot))(t, x)$ and $u(0, x) = h(x)$ for $(t, x) \in [0, T] \times \mathbb{R}^n$. Let us observe that for $f$ and $g$ defined above and $\hat{h} \in C^{1,2}_0([0, T] \times \mathbb{R}^n, \mathbb{R})$ one has

$$(A_{fg} \hat{h})(t, x) = \hat{h}'_t(t, x) + (L_{bo} \hat{h})(t, x)$$

for $t \in [0, T]$ and $x \in \mathbb{R}^n$, where

$$(L_{bo} \hat{h})(t, x) = \sum_{i=1}^{n} \hat{h}'_{x_i}(t, x)b_i(t, x) + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \hat{h}''_{x_i x_j}(t, x)a_{ij}(t, x),$$

where $a_{ij}(t, x)$ denote for $i, j = 1, \ldots, n$ elements of the matrix $a(t, x) = \sigma(t, x) \cdot \sigma^T(t, x)$ for $t \in [0, T]$ and $x \in \mathbb{R}^n$. In a particular case for $h \in C^2_0(\mathbb{R}^n)$ we have $(A_{fg} h)(t, x) = (L_{bo} h)(t, x)$ for $t \in [0, T]$ and $x \in \mathbb{R}^n$. Therefore for every $h \in C^2_0(\mathbb{R}^n)$ a function $u(t, x) = Eh(X_t(t))$ satisfies $u'_t(t, x) = (L_{bo} u(t, \cdot))(t, x)$ and $u(0, x) = h(x)$ for $(t, x) \in [0, T] \times \mathbb{R}^n$.

Conversely, if $b$ and $\sigma$ are such as above and $u \in C^{1,2}_0([0, T] \times \mathbb{R}^n, \mathbb{R})$ is bounded such that $\nu(0, x) = h(x)$ and $u'_t(t, x) = (L_{bo} u(t, \cdot))(t, x)$ for $h \in C^2_0(\mathbb{R}^n)$, $t \in [0, T]$ and $x \in \mathbb{R}^n$ then for every $x \in \mathbb{R}^n$ there exists exactly one weak solution $X_t(x)$ to (5) satisfying $X_t(0) = x$ with (P.1) and such that $\nu(t, x) = Eh(X_t(x))$. Indeed, by virtue of [6, Theorem IV.2.2] and the Strick and Varadhan uniqueness theorem for every $x \in \mathbb{R}^n$ there is exactly one weak solution $X_t(x)$ such that $X_t(x) = x$ with (P.1). Fix $(s, x) \in [0, T] \times \mathbb{R}^n$ and define a process $Y_{(0,x)}(t)$ by taking $Y_{(0,x)}(t) = (s - t, X^T_s(t))$ for $t \in [0, T]$. Let $\mathcal{A}$ be defined by $(\mathcal{A} u)(t, x) = -u'_t(t, x) + (L_{bo} u(t, \cdot))(t, x)$ for $u \in C^{1,2}_0([0, T] \times \mathbb{R}^n, \mathbb{R})$ and $t \in [0, T]$. We have $(\mathcal{A} u)(t, x) = 0$ for $t \in [0, T]$ and $x \in \mathbb{R}^n$. We can verify that $\mathcal{A}$ is the generator of $Y_{(0,x)}$. Therefore, by Dynkin’s formula (see [16, Theorem 7.4.1]) we obtain $\nu(s, x) = E[\nu(Y_{(0,x)}(t)))]$ for $t \in [0, T]$. In particular, choosing $t = s$ we get $\nu(t, x) = E[\nu(Y_{(0,x)}(t)))] = E[\nu(h(X^T_s(t))) = E[h(X_t(x))].$ Quite similar by virtue of [16, Theorem 8.2.1] we can obtain the Feynman–Kac type formulas for non-autonomous stochastic differential equations.

We can formulate the following theorems.

**Theorem 4.** Let $f \colon [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ and $g \colon [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times m}$ be continuous and bounded and let $g$ be a uniformly positive generator. Assume for every $x \in \mathbb{R}^n$ a process $X_t \colon \Omega \rightarrow C_T$ is a weak solution to the stochastic differential equation $dx_t = f(t, x_t) dt + g(t, x_t) dB_t$ defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, P)$ such that $X_t(0) = x$ with (P.1). Let $u(t, x) = E[\exp(\int_0^t c(\tau, X^T_\tau(\tau)) d\tau) h(X_t(\tau))]$ for $h \in C^2_0(\mathbb{R}^n)$, $c \in C_b([0, T] \times \mathbb{R}^n, \mathbb{R})$, $t \in [0, T]$ and $x \in \mathbb{R}^d$. Then $u \in C^{1,2}_0([0, T] \times \mathbb{R}^n, \mathbb{R})$ and

$$
\begin{cases}
\frac{\partial u(t, x)}{\partial t} = (L_{fg} u(t, \cdot))(t, x) + c(t, x) u(t, x) & \text{for } t \in [0, T], \ x \in \mathbb{R}^d, \\
u(0, x) = h(x) & \text{for } x \in \mathbb{R}^n.
\end{cases}
$$
Theorem 5. Let \( f : [0, T] \times \mathbb{R}^n \to \mathbb{R}^n \) and \( g : [0, T] \times \mathbb{R}^n \to \mathbb{R}^{n \times m} \) be continuous and bounded. If \( g \) is a uniformly positive generator and \( v \in C^{1,2}([0, T] \times \mathbb{R}^n, \mathbb{R}) \) is bounded and such that

\[
\begin{align*}
\frac{\partial v(t,x)}{\partial x} &= (L_{f,g}v(t,\cdot))(t,x) + c(t,x)v(t,x) \quad \text{for } t \in [0,T], \ x \in \mathbb{R}^n, \\
v(0,x) &= h(x) \quad \text{for } x \in \mathbb{R}^n,
\end{align*}
\]

for \( c \in C_b([0, T] \times \mathbb{R}^n, \mathbb{R}) \) and \( h \in C^2_0(\mathbb{R}) \) then for every \( x \in \mathbb{R}^n \) there is exactly one weak solution \( X_x \) to the stochastic differential equation \( dx_t = f(t,x_t) \, dt + g(t,x_t) \, dB_t \) defined on a filtered probability space \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, P)\) satisfying an initial condition \( X_x(0) = x \) with \((P,1)\) and such that \( v(t,x) = E[\exp(\int_0^T c(\tau, X_x(\tau)) \, d\tau)h(X_x(t))]) \) for \( t \in [0,T] \) and \( x \in \mathbb{R}^n \).

4. **Set-valued diffusion generator and continuous approximation selection theorems**

For given \( h \in C^2_b(\mathbb{R}^n), u \in \mathbb{R}^n \) and \( v \in \mathbb{R}^{n \times m} \) let us define \( (\mathcal{L}u,v,h)(x) \) by setting \( (\mathcal{L}u,v,h)(x) = (\partial_x h(x), u) + \frac{1}{2} \text{tr}(\partial_{xx} h(x) \cdot v^T) \) for \( x \in \mathbb{R}^n \), where \( \partial_x = (\partial_{x_1}, \ldots, \partial_{x_n}) \) and \( \partial_{xx} = (\partial_{x_i}x_j)_{1 \leq i \leq n, 1 \leq j \leq n} \). If \( F : [0,T] \times \mathbb{R}^n \to \mathcal{C}(\mathbb{R}^n) \) and \( G : [0,T] \times \mathbb{R}^n \to \mathcal{C}(\mathbb{R}^{n \times m}) \) we define \( \mathcal{L}_{FG} \) on \( C^2_b(\mathbb{R}^n) \) by taking \( (\mathcal{L}_{FG} h)(t,x) = (\mathcal{L}u,v,h)(x) \) for \( t \in [0,T], \ x \in \mathbb{R}^n \) and \( h \in C^2_b(\mathbb{R}) \), where \( \gamma(x,u,\sigma) = (\partial_x h(x), u) + \frac{1}{2} \text{tr}(\partial_{xx} h(x) \cdot \sigma) \) for \( u \in \mathbb{R}^n \) and \( \sigma \in \mathcal{L}(\mathbb{R},x) \), where \( v = v \cdot v^T \) for \( v \in \mathbb{R}^{n \times m} \).

We shall show that for some lower semicontinuous bounded set-valued mappings \( F : [0,T] \times \mathbb{R}^n \to \mathcal{C}(\mathbb{R}^n), \, G : [0,T] \times \mathbb{R}^n \to \mathcal{C}(\mathbb{R}^{n \times m}) \), every \( k = 1,2, \ldots \) and a bounded function \( u \in C^{1,2}([0,T] \times \mathbb{R}^n, \mathbb{R}) \) satisfying \( u'_k(t,x) \in (\mathcal{L}_{FG} u(t,\cdot))(t,x) \) for \( (t,x) \in [0,T] \times \mathbb{R}^n \) there are continuous selectors \( f_k \) and \( g_k \) for \( F \) and \( G \), respectively such that \( |u'_k(t,x) - (\mathcal{L}_{fg} u(t,\cdot))(t,x)| \leq \frac{1}{k} \) for \( (t,x) \in [0,T] \times \mathbb{R}^n \). Let us observe first that if \( F : [0,T] \times \mathbb{R}^n \to \mathcal{C}(\mathbb{R}^n) \) and \( D : [0,T] \times \mathbb{R}^n \to \mathcal{C}(\mathbb{R}^{n \times m}) \) are lower semicontinuous and bounded then a set-valued mapping \( F \times D : [0,T] \times \mathbb{R}^n \to \mathcal{C}(\mathbb{R}^n \times \mathbb{R}^{n \times m}) \) is lower semicontinuous and bounded, too. Furthermore for a given \( u \in C^{1,2}([0,T] \times \mathbb{R}^n, \mathbb{R}) \) the function \( \lambda : X \times Y \to \mathbb{R} \) defined by \( \lambda(u(x),u,\sigma) = \gamma(u(t,x),u,\sigma) \) for \( (t,x) \in X \) and \( (u,\sigma) \in Y \) with \( X = [0,T] \times \mathbb{R}^n \) and \( Y = \mathbb{R}^n \times \mathbb{R}^{n \times m} \) satisfies the assumptions of Theorem 3. Therefore, immediately from this theorem the following result follows.

**Lemma 6.** Let \( F : [0,T] \times \mathbb{R}^n \to \mathcal{C}(\mathbb{R}^n) \) and \( D : [0,T] \times \mathbb{R}^n \to \mathcal{C}(\mathbb{R}^{n \times m}) \) be lower semicontinuous and bounded with convex values and assume \( u \in C^{1,2}([0,T] \times \mathbb{R}^n, \mathbb{R}) \) is such that \( u'_k(t,x) \in \gamma(u(t,x), F(t,x) \times D(t,x)) \) for \( (t,x) \in [0,T] \times \mathbb{R}^n \). Then for every \( k = 1,2, \ldots \) there are continuous selectors \( f_k \) and \( \sigma_k \) for \( D \) and \( D \), respectively such that \( |u'_k(t,x) - \gamma(u(t,x), f_k(t,x), \sigma_k(t,x))| \leq \frac{1}{k} \) for \( (t,x) \in [0,T] \times \mathbb{R}^n \).

Given \( G : [0,T] \times \mathbb{R}^n \to \mathcal{C}(\mathbb{R}^{n \times m}) \) we shall consider the set-valued function \( D \) defined by \( D(t,x) = l(G(t,x)) \) for \( (t,x) \in [0,T] \times \mathbb{R}^n \), where \( l(v) = v \cdot v^T \) for \( v \in \mathbb{R}^{n \times m} \). It is clear that there are a lot of set-valued mappings \( \tilde{G} : [0,T] \times \mathbb{R}^n \to \mathcal{C}(\mathbb{R}^{n \times m}) \) such that \( l(G(t,x)) = l(\tilde{G}(t,x)) \). Therefore we introduce in \( \mathbb{R}^{n \times m} \) an equivalent relation \( R_l \) by setting \( x R_l y \) if and only if \( l(y) = l(x) \). Put \( X = \mathbb{R}^{n \times m} \) and let \( \tilde{X} = X / R_l \) be the \( R_l \)-quotient space. Let \( q : \tilde{X} \to \tilde{X} \) be the quotient mapping defined in the usual way by setting \( \tilde{x} \mapsto q(x) = [x] \in \tilde{X} \), where \([x] = \{ z \in X : z R_l x \} \). Denote by \( T_{\tilde{X}} \) a norm topology in \( \tilde{X} \) and let \( \overline{T} \) be a natural topology in \( \tilde{X} \) defined by \( \overline{T} = \{ V \subset \tilde{X} : q^{-1}(V) \in T_{\tilde{X}} \} \). It is clear that \( q \) is \( (T_{\tilde{X}}, \overline{T}) \)-continuous. Let us introduce in \( \tilde{X} \) a topology \( T_{\tilde{I}} = \{ q^{-1}(V) : V \in \overline{T} \} \). We have \( T_{\tilde{I}} \subset T_{\tilde{X}} \).
Lemma 7. For a given set-valued mapping $G : [0, T] \times \mathbb{R}^n \to \text{Cl}(\mathbb{R}^{n \times m})$ let $f : [0, T] \times \mathbb{R}^n \to \mathbb{R}^{n \times m}$ be continuous and such that $l(f(t, x)) \in l(G(t, x))$ for $t \in [0, T]$ and $x \in \mathbb{R}^n$. There is an $(T_{R^{n+1}}, \mathcal{T}_I)$-continuous selector $g$ of $G$ such that $l(f(t, x)) = l(g(t, x))$ for $t \in [0, T]$ and $x \in \mathbb{R}^n$, where $T_{R^{n+1}}$ denotes a norm topology in $\mathbb{R}^{n+1}$.

Proof. For every $(t, x) \in [0, T] \times \mathbb{R}^n$ we can select $u'_i \in G(t, x)$ such that $l(f(t, x)) = l(u'_i)$. Put $g(t, x) = u'_i$ for $t \in [0, T]$ and $x \in \mathbb{R}^n$. We have $q(g(t, x)) = q(f(t, x))$. By the $(T_{R^{n+1}}, \mathcal{T}_I)$-continuity of $q(f(\cdot))$ for every $u \in \mathcal{T}_I$ we have $(qf)^{-1}(V) \in T_{R^{n+1}}$. Then $g^{-1}(q^{-1}(V)) = f^{-1}(q^{-1}(V)) \in T_{R^{n+1}}$ for every $V \in \mathcal{T}_I$. Therefore for every $U \in T_I$ we have $g^{-1}(U) = f^{-1}(q^{-1}(V)) \in T_{R^{n+1}}$ because by the definition of $\mathcal{T}_I$ for every $U \in T_I$ there is $V \in \mathcal{T}_I$ such that $U = q^{-1}(V)$. □

We shall show that if $G : [0, T] \times \mathbb{R}^n \to \text{Cl}(\mathbb{R}^{n \times n})$ is such that $\det(u'_i) \neq 0$ for every $(t, x, u'_i) \in \text{Graph}(G)$ and a set-valued mapping $D : [0, T] \times \mathbb{R}^n \to \text{Cl}(\mathbb{R}^{n \times n})$ defined by $D(t, x) = l(G(t, x))$ for $t \in [0, T]$ and $x \in \mathbb{R}^n$ has a continuous selector $\sigma$ then there is an $(T_{R^{n+1}}, \mathcal{T}_I)$-continuous selector $g$ of $G$ such that $\sigma(t, x) = l(g(t, x))$ for $t \in [0, T]$ and $x \in \mathbb{R}^n$. The result will follow from the properties (see [12, pp. 81, 153]) of positive defined symmetric matrices. To present the proof of the result let us recall that a matrix $v \in \mathbb{R}^{n \times n}$ with elements $v_{ij}$ is said to be a down triangular matrix if $v_{ij} \neq 0$ for $i = 1, 2, \ldots, n$, and all elements of $v$ lying below its left–right diagonal are equal to zero. It can be proved (see [12, pp. 81–82]) that for every symmetric matrix $\sigma \in \mathbb{R}^{n \times n}$ of the range $r$ such that $d_k \neq 0$ for $k = 1, 2, \ldots, r$, with $d_k$ defined below the symmetric matrix $v = (v_{ij})_{n \times n}$ such that $\sigma = v \cdot v^T$ and elements $v_{ij}$ of such matrix $v$ are defined by

$$v_{ij} = \begin{cases} \dfrac{1}{\sqrt{d_{i}d_{j}}}\sigma \begin{pmatrix} 1 & 2 & \cdots & j-1 & i \\ 1 & 2 & \cdots & j-1 & j \\ & & \ddots & & \vdots \\ 0, & & & & 0 \end{pmatrix}, & j = 1, 2, \ldots, r, i = j, j+1, \ldots, n, \\ 0, & j = r + 1, r + 2, \ldots, n, \end{cases} \tag{6}$$

where $\sigma \begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix}$ denotes the $k$th order minor consisting of elements of $\sigma$ lying on the crossing of $k$ rows with indexes $i_1, \ldots, i_k$ and of $k$ columns with indexes $j_1, \ldots, j_k$ and $d_p = \sigma \begin{pmatrix} 1 & \cdots & p \\ 1 & \cdots & p \end{pmatrix}$ for $p = 1, 2, \ldots, r$.

Lemma 8. Let $G : [0, T] \times \mathbb{R}^n \to \text{Cl}(\mathbb{R}^{n \times n})$ be such that $\det(u'_i) \neq 0$ for every $(t, x, u'_i) \in \text{Graph}(G)$ and such that a set-valued mapping $D = l(G)$ has a continuous selector $\sigma$. Then there is an $(T_{R^{n+1}}, \mathcal{T}_I)$-continuous selector $g$ of $G$ such that $\sigma(t, x) = l(g(t, x))$ for $t \in [0, T]$ and $x \in \mathbb{R}^n$.

Proof. For every $t \in [0, T]$ and $x \in \mathbb{R}^n$ there is $u'_i \in G(t, x)$ such that $\sigma(t, x) = u'_i \cdot (u'_i)^T$ and $\det(u'_i) \neq 0$. Then (see [12, p. 153]) $\sigma(t, x)$ is symmetric positively defined for every $t \in [0, T]$ and $x \in \mathbb{R}^n$. Therefore, for every $t \in [0, T]$ and $x \in \mathbb{R}^n$ there is the down triangular matrix $v(t, x) = (v_{ij}(t, x))_{n \times n}$ such that $\sigma(t, x) = v(t, x) \cdot v^T(t, x)$ and such that all its elements $v_{ij}(t, x)$ are defined by (6) with $r = n$. By the continuity of $\sigma$ all its minors are continuous, too. Therefore, by (6) all elements $v_{ij}$ of $v$ are continuous on $[0, T] \times \mathbb{R}^n$. Then $v : [0, T] \times \mathbb{R}^n \to \mathbb{R}^{n \times n}$ is a continuous matrix such that $\sigma(t, x) = l(v(t, x))$ for $t \in [0, T]$ and $x \in \mathbb{R}^n$. By the definitions of $D$ and $\sigma$ we get $l(v(t, x)) \in l(G(t, x))$ for every $t \in [0, T]$ and $x \in \mathbb{R}^n$. Therefore, by virtue of Lemma 7 there is an $(T_{R^{n+1}}, \mathcal{T}_I)$-continuous selector $g$ of $G$ such that $\sigma(t, x) = l(g(t, x))$ for $t \in [0, T]$ and $x \in \mathbb{R}^n$. □
Immediately from Lemmata 6 and 8 we obtain the following theorems.

**Theorem 9.** Let $F : [0, T] \times \mathbb{R}^n \to \text{Cl}(\mathbb{R}^n)$ and $G : [0, T] \times \mathbb{R}^n \to \text{Cl}(\mathbb{R}^{n \times n})$ be l.s.c. and bounded set-valued mappings with convex and diagonally convex values, respectively. Assume $\det(u'_x) \neq 0$ for every $(t, x, u'_x) \in \text{Graph}(G)$ and let $u \in C^{1,2}([0, T] \times \mathbb{R}^n, \mathbb{R})$ be such that $u'_t(t, x) \in (\mathcal{L}_FGu(t, \cdot))(t, x)$ for $t \in [0, T]$ and $x \in \mathbb{R}^n$. Then for every $k = 1, 2, \ldots$ there are continuous and $(\mathcal{T}_{\mathbb{R}^n+1}, \mathcal{T})$-continuous selectors $f_k$ and $g_k$ of $F$ and $G$, respectively such that $|u'_t(t, x) - (\mathcal{L}_{f_k}g_ku(t, \cdot))(t, x)| \leq \frac{1}{k}$ for $t \in [0, T]$ and $x \in \mathbb{R}^n$.

**Proof.** Similarly as in the proof of Lemma 8 it can be verified that for every continuous selector $\sigma$ of $D$ there is a continuous down triangular matrix function $v : [0, T] \times \mathbb{R}^n \to \mathbb{R}^{n \times n}$ such that $\sigma(t, x) = l(v(t, x))$ for $t \in [0, T]$ and $x \in \mathbb{R}^n$. On the other hand by $\sigma(t, x) \in l(G(t, x))$ there is $u'_x \in G(t, x)$ such that $\sigma(t, x) = l(u'_x)$ for $t \in [0, T]$ and $x \in \mathbb{R}^n$. But every $u'_x \in G(t, x)$ is a down triangular matrix for $t \in [0, T]$ and $x \in \mathbb{R}^n$. Therefore the position defined and symmetric matrix $\sigma(t, x)$ has two representations $\sigma(t, x) = l(v(t, x))$ and $\sigma(t, x) = l(u'_x)$ for $t \in [0, T]$ and $x \in \mathbb{R}^n$. By virtue of (6) all elements matrices $v(t, x)$ and $u'_x$ are defined by the same formulas by elements of $\sigma(t, x)$. Therefore $v(t, x) = u'_x$ for every $t \in [0, T]$ and $x \in \mathbb{R}^n$. Taking $g(t, x) = v(t, x)$ for $t \in [0, T]$ and $x \in \mathbb{R}^n$ we get $g(t, x) \in G(t, x)$ and $\sigma(t, x) = l(g(t, x))$ for $t \in [0, T]$ and $x \in \mathbb{R}^n$. We also have $\sigma(t, x) = l(-v(t, x))$ for $t \in [0, T]$ and $x \in \mathbb{R}^n$ but $-g$ does not have to be a selector of $G$. Now by virtue of Lemmata 6 and 8 we can see that for every $k = 1, 2, \ldots$ there are continuous selectors $f_k$ and $g_k$ of $F$ and $G$, respectively such that $|u'_t(t, x) - (\mathcal{L}_{f_k}g_ku(t, \cdot))(t, x)| \leq \frac{1}{k}$ for $t \in [0, T]$ and $x \in \mathbb{R}^n$. $\square$

For a given set-valued mapping $G : [0, T] \times \mathbb{R}^n \to \text{Cl}(\mathbb{R}^{n \times n})$ and a continuous selector $\sigma$ of $D = l(G)$, we are interested in the existence of a continuous selector $g$ of $G$ that is a uniformly positive generator of $\sigma$.

We have the following results.

**Lemma 11.** Let $G : [0, T] \times \mathbb{R}^n \to \text{Cl}(\mathbb{R}^{n \times n})$ be such that a set-valued mapping $D = l(G)$ has a continuous selector $\sigma$ and that there exists a positive number $L > 0$ such that $\frac{l(u'_x)(u, u)}{\|u\|} \geq L$ for $(t, x, u'_x) \in \text{Graph}(G)$ and $u \in \mathbb{R}^n$ with $\|u\| \neq 0$, where $l(u'_x)(u, u)$ denotes a quadratic form on $\mathbb{R}^n$ with the matrix $l(u'_x) = u'_x \cdot (u'_x)^T$. Then there exists an $(\mathcal{T}_{\mathbb{R}^n+1}, \mathcal{T})$-continuous selector $g$ of $G$ that is a uniformly positive generator of $\sigma$.

**Proof.** Let us observe (see [12, p. 165]) that for every $(t, x, u'_x) \in \text{Graph}(G)$ a number

$$\lambda = \sup_{t \in [0, T]} \sup_{x \in \mathbb{R}^n} \min_{u \in \mathbb{R}^n, u \neq 0} \frac{l(u'_x)(u, u)}{\|u\|}$$

for $(t, x, u'_x) \in \text{Graph}(G)$.
is equal to \( \sup_{t \in [0, T]} \sup_{x \in \mathbb{R}^n} \inf_{1 \leq i \leq n} \lambda_i (u^i_x(t, x)^T) \), where \( \lambda_i (u^i_x(t, x)^T) \) denotes for \( i = 1, 2, \ldots, n \), the eigenvalues of the symmetric matrix \( u^i_x(t, x)^T \). Therefore a matrix function \([0, T] \times \mathbb{R}^n \ni (t, x) \rightarrow u^i_x(t, x)^T \in \mathbb{R}^{n \times n}\) is uniformly positive defined because \( \lambda \geq L > 0 \). Similarly as above for every \((t, x) \in [0, T] \times \mathbb{R}^n\) we can select \( u^i_x(t, x) \) such that \( \sigma(t, x) = l(u^i_x) \). Now by the statements presented above it follows that \( \sigma \) is uniformly positive defined. Then there exists a continuous matrix function \( v : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n} \) such that \( (t, x) = l(v(t, x)) \) for every \((t, x) \in [0, T] \times \mathbb{R}^n\). Hence by virtue of Lemma 7 there is an \((\mathcal{T}_{\mathbb{R}^{n+1}}, \mathcal{T})\)-continuous selector \( g \) of \( G \) such that \( l(v(t, x)) = l(g(t, x)) \) for \((t, x) \in [0, T] \times \mathbb{R}^n\). But \( \sigma(t, x) = l(g(t, x)) \) for \((t, x) \in [0, T] \times \mathbb{R}^n\) and \( \sigma \) is uniformly positive defined. Then \( g \) is an \((\mathcal{T}_{\mathbb{R}^{n+1}}, \mathcal{T})\)-continuous selector of \( G \) that is a uniformly positive generator of \( \sigma \).

Similarly as above we can also prove the following selection theorem.

**Theorem 12.** Let \( F : [0, T] \times \mathbb{R}^n \rightarrow \text{Cl}(\mathbb{R}^n) \) and \( G : [0, T] \times \mathbb{R}^n \rightarrow \text{Cl}(\mathbb{R}^{n \times n}) \) be l.s.c. and bounded set-valued mappings with convex and diagonally convex values, respectively. Assume \( G \) is such that \( u^i_x \) is a down triangle matrix for every \((t, x, u^i_x) \in \text{Graph}(G)\) and such that for every continuous selector \( \varphi \) of \( G \) there is a number \( L_\varphi > 0 \) such that \( \frac{l(\varphi(t, x))(u, \cdot)}{u^2} \geq L_\varphi \) for every \((t, x) \in [0, T] \times \mathbb{R}^n\) and \( u \in \mathbb{R}^n \) with \( \|u\| \neq 0 \). Let \( v \in C^{1,2}([0, T] \times \mathbb{R}^n, \mathbb{R}) \) be such that \( v'(t, x) \in (\mathcal{L}_{FG}v(t, \cdot))(t, x) \) for every \( t \in [0, T] \) and \( x \in \mathbb{R}^n \). Then for every \( k = 1, 2, \ldots \) there are continuous selectors \( f_k \) and \( g_k \) of \( F \) and \( G \), respectively such that \( |v'(t, x) - (\mathcal{L}_{fgk}v(t, \cdot))(t, x)| \leq \frac{1}{k} \) for \( t \in [0, T] \) and \( x \in \mathbb{R}^n \). Furthermore for every \( k = 1, 2, \ldots \) the symmetric matrix function \( \sigma_k = l(gk) \) is uniformly positive defined.

5. Stochastic differential inclusions and diffusion processes

Denote by \( \mathcal{G} \) a family of all l.s.c. bounded and diagonally convex values set-valued mappings \( G : [0, T] \times \mathbb{R}^n \rightarrow \text{Cl}(\mathbb{R}^{n \times m}) \) such that for every \( G_1, G_2 \in \mathcal{G} \) we have \( l(G_1) = l(G_2) \) and such that for every \( G \in \mathcal{G} \) and a continuous selector \( \sigma \) of \( D = l(G) \) there is a continuous selector \( g \) of \( G \) that is a uniformly positive generator for \( \sigma \). Let us observe that for every \( G, \tilde{G} \in \mathcal{G} \) one has \( \text{Graph}(G) \in [0, T] \times \mathbb{R}^n \times \text{Cl}(\mathbb{R}^{n \times m}) \) and \( \text{Graph}(\tilde{G}) \in [0, T] \times \mathbb{R}^n \times \text{Cl}(\mathbb{R}^{n \times m}) \) for positive integers \( m \) and \( \tilde{m} \). In fact the class \( \mathcal{G} \) contains set-valued mappings \( G \) and \( \tilde{G} \) with values in \( \mathbb{R}^{n \times m} \) and \( \mathbb{R}^{n \times \tilde{m}} \), respectively with arbitrarily taken integers \( m \) and \( \tilde{m} \) if \( l(G) = l(\tilde{G}) \) and the multifunctions \( D = l(G) \) possess the properties presented above. In particular \( G \in \mathcal{G} \) implies that \( G \) is diagonally convex-valued and that every continuous selector \( \sigma \) of \( D = l(G) \) is uniformly positive defined. Furthermore there is a continuous selector \( g \) of \( G \) such that \( \sigma = l(g) \). It is clear that \( \mathcal{G} \neq \emptyset \) because it contains all set-valued mappings \( R_i \)-equivalent to the set-valued mapping \( G : [0, T] \times \mathbb{R}^n \rightarrow \text{Cl}(\mathbb{R}^{n \times n}) \) satisfying the assumptions of Theorem 12.

We shall prove now the main results of the paper.

**Theorem 13.** Let \( F : [0, T] \times \mathbb{R}^n \rightarrow \text{Cl}(\mathbb{R}^n) \) be l.s.c. and bounded and let \( G \in \mathcal{G} \). Assume \( F \) has convex values. Then for every \( x \in \mathbb{R}^n \) there are a filtered probability space \((\Omega, \mathcal{F}, (\mathcal{F}_i)_{0 \leq i \leq T}, P)\) and an \((n + 1)\)-dimensional diffusion process \( Y_{(0, x)} = (t, X^T_{\lambda}(t))_{0 \leq t \leq T} \) on \((\Omega, \mathcal{F}, P)\) such that
(i) $X_x$ is a weak solution to a stochastic differential inclusion

$$x_t - x_s \in \int_s^t F(\tau, x_\tau) \, d\tau + \int_s^t G(\tau, x_\tau) \, dB_\tau \quad \text{for} \ 0 \leq s \leq t \leq T,$$

on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, P)$ satisfying $X_x(0) = x$ with (P.1).

(ii) for every $h \in C^2_0(\mathbb{R}^n)$ a function $u:[0, T] \times \mathbb{R}^n \to \mathbb{R}$ defined by $u(t, x) = E[h(X_x(t))]$ for $t \in [0, T]$ and $x \in \mathbb{R}^n$ belongs to $C^{1,2}([0, T] \times \mathbb{R}^n, \mathbb{R})$ and is a solution to a partial differential inclusion

$$u'_t(t, x) \in (\mathcal{L}_{FG}u(t, \cdot)) (t, x)$$

for $t \in [0, T]$ and $x \in \mathbb{R}^n$ satisfying the initial condition $u(0, x) = h(x)$ for $x \in \mathbb{R}^n$.

**Proof.** Let $f$ and $g$ be continuous selectors for $F$ and $G$, respectively such that the matrix function $\sigma = I(g)$ is uniformly positive defined. Such selectors exist by Michael’s continuous selection theorem (see [9, Theorem II.4.1]) and the properties of the family $G$. Consider now a stochastic differential equation

$$dx_t = f(t, x_t) \, dt + g(t, x_t) \, dB_t.$$  

By virtue of [6, Theorem IV.2.2] and the uniqueness theorem of Strook and Varadhan (see [17]) for every $x \in \mathbb{R}^n$ there exists exactly one weak solution $(X_x, B)$ to (8) on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, P)$ such that $X_x(0) = x$ with (P.1). Similarly as above we can verify that the unique in law process $Y_x(t) = (t, X_x(t))^T$ satisfies $dY_x(t) = \tilde{f}(Y_x(t)) \, dt + \tilde{g}(Y_x(t)) \, dB_t$ and $Y_x(0) = (0, x)$ P-a.s. for $t \in [0, T]$ where $\tilde{f} = (1, f)^T$ and $\tilde{g} = (0, g_1, \ldots, g_n)^T$, with $0, g_i \in \mathbb{R}^{1 \times m}$, where $0 = (0, \ldots, 0)$ and $g_i$ denotes the $i$th row of $g$ for $i = 1, \ldots, n$. Then by virtue of [6, Theorem IV.6.1] and [16, Theorem IV.6.1] it follows that $(Y_x(t))_{t \in [0, T]}$ is the diffusion process. Furthermore by Dynkin’s formula (see [16, Theorem 7.4.1]) it follows that the function $U(t, (0, x)) = E[(h \circ \Pi_{\mathbb{R}^n})(Y_x(t))]$ defined for $h \in C^2_0(\mathbb{R}^n)$ belongs to $C^{1,2}([0, T] \times \mathbb{R}^{n+1}, \mathbb{R})$ and satisfies $U'_t(t, (0, x)) = E[A_f \tilde{g}(h \circ \Pi_{\mathbb{R}^n})(Y_x(t))]$ for $t \in [0, T]$ and $x \in \mathbb{R}^n$. Similarly as in the proof of [16, Theorem 8.1.1] it follows that $U'_t(t, (0, x)) = (A_f \mathcal{L}_{g} U(t, \cdot))(t, (0, x))$ for $t \in [0, T]$ and $x \in \mathbb{R}^n$. We can easily verify that the function $u(t, x) = U(t, (0, x))$ satisfies $u(t, x) = Eh(X_x(t))$, $u'_t(t, x) = (\mathcal{L}_{fg} u(t, \cdot))(t, x)$ and $u(0, x) = h(x)$ for $t \in [0, T]$ and $x \in \mathbb{R}^n$. By the properties of $f$ and $g$ hence it follows

$$\begin{cases} u'_t(t, x) \in (\mathcal{L}_{FG}u(t, \cdot))(t, x) & \text{for} \ t \in [0, T], \ x \in \mathbb{R}^n, \\
0(t, x) = h(x) & \text{for} \ x \in \mathbb{R}^n. \end{cases} \quad \square$$

**Theorem 14.** Let $F : [0, T] \times \mathbb{R}^n \to \text{Cl}(\mathbb{R}^n)$ be continuous and bounded and let $G \in G$ be continuous. Assume $F$ and $G$ have convex values. For every bounded function $v \in C^{1,2}([0, T] \times \mathbb{R}^n, \mathbb{R})$ and $h \in C^2_0(\mathbb{R}^n)$ such that

$$\begin{cases} v'_t(t, x) \in (\mathcal{L}_{FG}v(t, \cdot))(t, x) & \text{for} \ t \in [0, T], \ x \in \mathbb{R}^n, \\
v(0, x) = h(x) & \text{for} \ x \in \mathbb{R}^n, \end{cases}$$

there is a weak solution $\tilde{X}_x$ to the stochastic differential inclusion (2) on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, \tilde{P})$ satisfying an initial condition $\tilde{X}_x(0) = x$ with (P.1) and such that $v(t, x) = E[h(X_x(t))]$ for $t \in [0, T]$ and $x \in \mathbb{R}^n$. 

**Proof.** Let $D(t, x) = I(G(t, x))$ for $(t, x) \in [0, T] \times \mathbb{R}^n$. By virtue of Lemma 6 for every $k = 1, 2, \ldots$ there are continuous selectors $f_k$ and $\sigma_k$ of $F$ and $D$, respectively such that $|v_k(t, x) - \gamma(v(t, x), f_k(t, x), \sigma_k(t, x))| \leq \frac{1}{k}$ for $(t, x) \in [0, T] \times \mathbb{R}^n$. By the properties of the family $G$ for every $k = 1, 2, \ldots$ there is a continuous selector $g_k$ of $G$ such that $\sigma_k = I(g_k)$ is uniformly positive defined. It is clear that $f_k$ and $g_k$ satisfy an inequality $|v_k(t, x) - (\mathcal{L}_{f_k g_k} v(t, \cdot))(t, x)| \leq \frac{1}{k}$ for every $(t, x) \in [0, T] \times \mathbb{R}^n$ and $k = 1, 2, \ldots$. By the properties of $f_k$ and $g_k$ for every $k = 1, 2, \ldots$ there is exactly one weak solution $(X^k_x, B^k)$ to the stochastic differential equation

$$dx_t = f_k(t, x_t)dt + g_k(t, x_t)dB_t$$

on a filtered probability space $(\Omega^k, \mathcal{F}^k, (\mathcal{F}^k_t)_{0 \leq t \leq T}, P^k)$ satisfying the initial condition $X^k_x(0) = x$ with $(P^k.1)$ It is clear that $(X^k_x)_{k=1}^{\infty}$ is a sequence of weak solutions to the stochastic differential inclusion (2) such that $X^k_x(0) = x$ with $(P^k.1)$. Then $X^k_x \in \mathcal{X}_x(F, G)$ for every $k = 1, 2, \ldots$. By virtue of Theorem 1 the set $\mathcal{X}_x(F, G)$ is sequentionally weakly compact with respect to the convergence in distribution. Therefore (see [6, Theorem I.2.7]) there are a subsequence $(n_k)_{k=1}^{\infty}$ of $(k)_{k=1}^{\infty}$, a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ and continuous stochastic processes $\tilde{X}^n_{n_k} \tilde{X}_x$ on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ such that $P(\tilde{X}^n_{n_k})^{-1} = P((\tilde{X}^n_{n_k})^{-1}$ for $k = 1, 2, \ldots,$ and $P(\tilde{X}^n_{n_k})^{-1} \rightarrow P(\tilde{X}_x)^{-1}$ weakly in $\mathcal{M}(C^{T}_{\mathbb{R}})$ as $k \rightarrow \infty$. Then $(X^k_x)_{k=1}^{\infty}$ converges in distribution to $\tilde{X}_x$ as $k \rightarrow \infty$ that is equivalent to $\tilde{E}[I(\tilde{X}^n_{n_k})] \rightarrow \tilde{E}[I(\tilde{X}_x)]$ for every $l \in C_b(C^{T}_{\mathbb{R}})$. By the weak compactness of $\mathcal{X}_x(F, G)$ it follows that $\tilde{X}_x \in \mathcal{X}_x(F, G)$ Therefore there is an $\tilde{F}_t$-Brownian motion $(\tilde{B}_t)_{0 \leq t \leq T}$ on the probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ such that $(\tilde{X}_x, \tilde{B})$ satisfies (2) and $\tilde{X}_x(0) = x$ with $(\tilde{P}.1)$.

Let $\sigma_r = \inf \{ t \in [0, T] : \tilde{X}_x(t) \notin K_r \}$ for $r = 1, 2, \ldots$, where $K_r = \{ y \in \mathbb{R}^n : |y| \leq r \}$. Since $\tilde{X}_x \in \mathcal{X}_x(F, G)$ then by [8, Theorem 4] there are $\tilde{F}_t$-nonanticipative selectors $\tilde{f}$ and $\tilde{g}$ of $(F \circ \tilde{X})$ and $(G \circ \tilde{X})$ such that

$$\tilde{X}(t) = x + \int_0^t \tilde{f}_r d\tau + \int_0^t \tilde{g}_r dB_r$$

for $t \in [0, T]$ with $(\tilde{P}.1)$. By the boundedness of $F$ and $G$ also $\tilde{f}$ and $\tilde{g}$ are bounded. Therefore, $\lim_{r \rightarrow \infty} \sigma_r = \infty$ with $(\tilde{P}.1)$. Then for every $t \in [0, T]$ we have

$$\lim_{r \rightarrow \infty} \int_{\sigma_r \leq t} v(t - \sigma_r, \tilde{X}_x(\sigma_r)) d\tilde{P} = 0.$$  

(11)

Let us observe that

$$E_{n_k}[I(X^n_{n_k})] = \int_{\mathbb{R}^n} l(x) P^n_{x}(dx) = \int_{\mathbb{R}^n} l(x) \tilde{P}^{n_k}_{x}(dx) = \tilde{E}[I(\tilde{X}^{n_k}_x)]$$

for every $k = 1, 2, \ldots$, where $P^n_{x} = P(X^n_{x})^{-1} = P(\tilde{X}^{n_k}_x)^{-1} = \tilde{P}^{n_k}_{x}$. Then for every $k = 1, 2, \ldots$, one has

$$E_{n_k}[v'_k(t, X^n_{n_k}(t)) - (\mathcal{L}_{f_k g_k} v(t, \cdot))(t, X^n_{n_k}(t))]$$

$$= \tilde{E}[v'_k(t, \tilde{X}^{n_k}_x(t)) - (\mathcal{L}_{f_k g_k} v(t, \cdot))(t, \tilde{X}^{n_k}_x(t))]$$
But \((X_x^{n_k}, B^{n_k})\) is a weak solution to (10) with \(k = n_k\) such that \(X_x^{n_k}(0) = x\) with \((P^{n_k}.1)\) for \(k = 1, 2, \ldots\). Therefore,

\[
X_x^{n_k}(t) = x + \int_0^t f_{n_k}(\tau, X_x^{n_k}(\tau))\,d\tau + \int_0^t g_{n_k}(\tau, X_x^{n_k}(\tau))\,dB^{n_k}_\tau,
\]

for \(t \in [0, T]\) with \((P^{n_k}.1)\) for \(k = 1, 2, \ldots\). Hence by Itô’s formula it follows

\[
E_{n_k}\left[ v\left( t_0 - t \wedge \sigma_r, X_x^{n_k}(t \wedge \sigma_r) \right) \right] - v(t_0, x)
\]

\[
= E_{n_k}\left[ \int_0^{t \wedge \sigma_r} \left\{ (\mathcal{L}_{f_{n_k} g_{n_k}} v(t_0 - s, \cdot))(s, X_x^{n_k}(s)) - \frac{\partial v}{\partial t}(t_0 - s, X_x^{n_k}(s)) \right\} \,ds \right]
\]

for every \(k, r = 1, 2, \ldots, t_0 \in (0, T)\) and \(0 \leq t \leq t_0\). Therefore, by the equality presented above for every \(k, r = 1, 2, \ldots\) we also have

\[
\tilde{E}\left[ v\left( t_0 - t \wedge \sigma_r, \tilde{X}_x^{n_k}(t \wedge \sigma_r) \right) \right] - v(t_0, x)
\]

\[
= \tilde{E}\left[ \int_0^{t \wedge \sigma_r} \left\{ (\mathcal{L}_{f_{n_k} g_{n_k}} v(t_0 - s, \cdot))(s, \tilde{X}_x^{n_k}(s)) - \frac{\partial v}{\partial t}(t_0 - s, \tilde{X}_x^{n_k}(s)) \right\} \,ds \right]
\]

Hence it follows

\[
\left| \tilde{E}\left[ v\left( t_0 - t \wedge \sigma_r, \tilde{X}_x^{n_k}(t \wedge \sigma_r) \right) \right] - v(t_0, x) \right| 
\]

\[
\leq \tilde{E}\left[ \int_0^{t \wedge \sigma_r} \left| (\mathcal{L}_{f_{n_k} g_{n_k}} v(t_0 - s, \cdot))(s, \tilde{X}_x^{n_k}(s)) - \frac{\partial v}{\partial t}(t_0 - s, \tilde{X}_x^{n_k}(s)) \right| \,ds \right]
\]

\[
\leq \tilde{E}\left[ \int_0^{t \wedge \sigma_r} \sup_{x \in \mathbb{R}^n} \left| (\mathcal{L}_{f_{n_k} g_{n_k}} v(t_0 - s, \cdot))(s, x) - \frac{\partial v}{\partial t}(t_0 - s, x) \right| \,ds \leq \frac{1}{n_k} \tilde{E}(t \wedge \sigma_r) \right) \tag{12}
\]

for every \(k, r = 1, 2, \ldots, t \in [0, t_0]\). Passing to the limit in (12) by \(k \to \infty\) we obtain

\[
\tilde{E}\left[ v\left( t_0 - t \wedge \sigma_r, \tilde{X}_x(t \wedge \sigma_r) \right) \right] = v(t_0, x)
\]

for \(0 \leq t \leq t_0\) and \(r = 1, 2, \ldots\). Then

\[
v(t_0, x) = \int_{[\sigma_r < t]} v(t_0 - \sigma_r, \tilde{X}_x(\sigma_r))\,d\tilde{P} + \int_{[\sigma_r > t]} v(t_0 - t, \tilde{X}_x(t))\,d\tilde{P}
\]

for \(0 \leq t \leq t_0\) and \(r = 1, 2, \ldots\). Hence by (11) we obtain

\[
v(t_0, x) = \lim_{r \to \infty} \int_{[\sigma_r > t]} v(t_0 - t, \tilde{X}_x(t))\,d\tilde{P} = \tilde{E}\left[ v(t_0 - t, \tilde{X}_x(t)) \right].
\]

Passing to the limit in the last equality by \(t \to t_0\) we get

\[
v(t_0, x) = \lim_{t \to t_0} \tilde{E}\left[ v(t_0 - t, \tilde{X}_x(t)) \right] = \tilde{E}\left[ v(0, \tilde{X}_x(t_0)) \right] = \tilde{E}\left[ h(\tilde{X}_x(t_0)) \right]
\]

for \(t_0 \in (0, T]\) and \(x \in \mathbb{R}^n\). It is clear that we have also \(v(0, x) = h(x) = \tilde{E}[h(\tilde{X}_x(0))]\) for \(x \in \mathbb{R}^n\). \(\square\)
In a similar way we can also prove the following set-valued Feynman’s–Kac type formulas.

**Theorem 15.** Let \( F : [0, T] \times \mathbb{R}^n \to \text{Cl}(\mathbb{R}^n) \) be l.s.c. and bounded and let \( G \in \mathcal{G} \). Assume \( F \) has convex values. Then for every \( x \in \mathbb{R}^n \) there are a filtered probability space \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, P)\) and an \((n+1)\)-dimensional diffusion process \( Y_x = (t, X^T_x(t))_{0 \leq t \leq T} \) such that

(i) \( Y_x \) is a weak solution to the stochastic differential inclusion (2) on \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, P)\) satisfying \( Y_x(0) = x \) with (P.1),

(ii) for every \( h \in C^2_0(\mathbb{R}^n) \) and \( c \in C_b([0, T] \times \mathbb{R}^n, \mathbb{R}) \) a function \( u : [0, T] \times \mathbb{R}^n \to \mathbb{R} \) defined by

\[
u(t, x) = E\left[\exp\left(\int_0^t c(t, X^T_x(\tau)) d\tau\right) h(X^T_x(t))\right]
\]

for \( t \in [0, T] \) and \( x \in \mathbb{R}^n \) belongs to \( C^{1,2}([0, T] \times \mathbb{R}^n, \mathbb{R}) \) and is a solution to a partial differential inclusion

\[
u'_t(t, x) \in (\mathcal{L}_F G u(t, \cdot))(t, x) + c(t, x)u(t, x)
\]

for \( t \in [0, T] \) and \( x \in \mathbb{R}^n \) satisfying an initial condition \( u(0, x) = h(x) \) for \( x \in \mathbb{R}^n \).

**Theorem 16.** Let \( F : [0, T] \times \mathbb{R}^n \to \text{Cl}(\mathbb{R}^n) \) be continuous and bounded and let \( G \in \mathcal{G} \) be continuous. Assume \( F \) and \( G \) have convex values. Let \( v \in C^{1,2}([0, T] \times \mathbb{R}^n, \mathbb{R}) \) be bounded and such that

\[
\begin{cases}
    v'_t(t, x) \in (\mathcal{L}_F G v(t, \cdot))(t, x) + c(t, x)v(t, x) & \text{for } t \in [0, T], x \in \mathbb{R}^n, \\
    v(0, x) = h(x) & \text{for } x \in \mathbb{R}^n,
\end{cases}
\]

with \( h \in C^2_0(\mathbb{R}^n) \) and \( c \in C_b([0, T] \times \mathbb{R}^n, \mathbb{R}) \). Then there exists a weak solution \( \tilde{X}_x \) to the stochastic differential inclusion (2) on a filtered probability space \((\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_t)_{0 \leq t \leq T}, \tilde{P})\) satisfying an initial condition \( \tilde{X}_x(0) = x \) with (P.1) and such that \( v(t, x) = \tilde{E}[\exp(\int_0^t c(t, \tilde{X}_x(\tau)) d\tau) \times h(\tilde{X}_x(t))] \) for \( t \in [0, T] \) and \( x \in \mathbb{R}^n \).

**Acknowledgments**

The author is deeply indebted to the anonymous referees of the paper for very valuable suggestions greatly improving the text.

**References**