International Journal of Difference Equations
ISSN 0973-6069, Volume 4, Number 1, pp. 27–41 (2009)
http://campus.mst.edu/ijde

Stochastic Differential Equations and Inclusions with Mean Derivatives relative to the Past

Svetlana V. Azarina
Voronezh State Technical University
Department of Highest Mathematics and
Physical and Mathematical Modelling
Moskovsky prosp. 14, 394026 Voronezh, Russia
azarina@math.vsu.ru

Yuri E. Gliklikh
Voronezh State University
Mathematics Faculty
Universitetskaya pl. 1, 394006 Voronezh, Russia
yeg@math.vsu.ru

Abstract

We introduce and investigate a new sort of stochastic equations and inclusions given in terms of mean derivatives defined with respect to conditional expectation relative to the "past" sigma-algebra of a process. Some existence of solution results are proved. A new type of approximations to an upper semi-continuous set-valued mapping with convex compact values, point-wise converging to a measurable selector, is constructed and applied to investigation of inclusions with the above-mentioned derivatives.

AMS Subject Classifications: 60H10, 34K50, 37H10, 47H04.
Keywords: Mean derivatives, "past" σ-algebra, stochastic differential equations, stochastic differential inclusions, approximations of set-valued mappings.

1 Introduction

The notion of mean derivatives was introduced by Edward Nelson (see [12–14]) for the needs of stochastic mechanics (a version of quantum mechanics). Then it was found

Received May 6, 2008; Accepted July 16, 2008
Communicated by Alexandra Rodkina

that equations with mean derivatives also arose as equations of motion in some other branches of mathematical physics (see, e.g., [9–11], etc.).

The key point of Nelson's idea is the use of conditional expectation in construction of mean derivatives (see the definitions below). Nelson suggested two versions of construction based on the conditional expectation either relative to the "past" or relative to the "present" σ-algebras (see [9]). For a Markov process they give the same result but for a nonmarkovian one they are different.

Note that the classical Nelson's forward mean derivatives give information only about the drift of stochastic process. In [2, 3] we modified a certain Nelson's idea and introduced a new mean derivative relative to the "present", called quadratic, that is responsible for the diffusion coefficient. This allowed us to prove that given Nelson's forward mean derivative and quadratic derivative (both relative to the "present"), under some natural conditions it was possible to recover the process. Some existence of solutions theorems for differential equations and inclusions with mean derivatives relative to the "present" were obtained in [2, 3].

The equations and inclusion with mean derivatives relative to the "past" is a natural replacement of those relative to the "present" in the case of nonmarkovian processes. But they require more complicated methods for their investigation. The main aim of this paper is to describe equations and inclusions with mean derivatives relative to the "past" and to prove some existence of solutions theorems for this case.

The structure of paper is as follows. In Section 2 we describe the mean derivatives (both Nelson's classical and quadratic) relative to the "past" and investigate their properties used below. In order to distinguish the derivatives relative to the "past" from those relative to the "present" as in [2, 3], we call the former $\mathcal{P}$-mean derivatives.

In Section 3 we introduce equations with $\mathcal{P}$-mean derivatives and prove a simple existence of solutions theorem. We show that under some rather strong assumptions the solution of equation exists in the class of Itô diffusion type processes. The material of this section forms the basis for main results on inclusions with $\mathcal{P}$-mean derivatives in Section 5.

Section 4 is devoted to a complicated technical statement on existence of some special single-valued continuous approximations to the set-valued mappings appearing in the right-hand sides of inclusions with $\mathcal{P}$-mean derivatives. Those approximations point-wise converge to a measurable selector of the set-valued mapping and possess some special measurability properties. The use of such approximations is one of key points in investigation of differential inclusions below.

In Section 5 we prove the main result of the paper, an existence of solutions theorem for inclusions with $\mathcal{P}$-mean derivatives, having upper semi-continuous convex-valued right-hand sides.

We refer the reader to [1, 5] for preliminary facts from the theory of set-valued mappings and to [8, 16] for those from stochastic analysis.

2 $\mathcal{P}$-mean Derivatives

Consider the n-dimensional vector space $\mathbb{R}^n$ and a stochastic process $\xi(t)$, $t \in [0, +\infty)$ with values in $\mathbb{R}^n$ given on a certain probability space $(\Omega, \mathcal{F}, \mathsf{P})$. We say that $\xi(t)$ is L^1 if the expectation $E(\xi(t))$ is well-posed at every t.

Denote by $\mathcal{P}_t^\xi$ the σ-subalgebra of $\mathcal{F}$ that is generated by preimages of Borel sets in $\mathbb{R}^n$ under all mappings $\xi(s) : \Omega \to \mathbb{R}^n$, $0 \le s \le t$. By $E(\cdot \mid \mathcal{P}_t^\xi)$ we denote the conditional expectation with respect to $\mathcal{P}_t^\xi$. According to Nelson [12–14] we call $\mathcal{P}_t^\xi$ the past of process $\xi(t)$.

Definition 2.1. The forward mean derivative relative to the past ($\mathcal{P}$-mean derivative) $D^{\mathcal{P}}\xi(t)$ of a process $\xi(t)$ at a time instant t is L^1-random element of the form

$$D^{\mathcal{P}}\xi(t) = \lim_{\triangle t \to +0} E\left(\frac{\xi(t + \triangle t) - \xi(t)}{\triangle t} \,\Big|\, \mathcal{P}_t^\xi \right), \tag{2.1}$$

where the limit is assumed to exist in L^1 and $\triangle t \to +0$ means that $\triangle t$ tends to 0 and $\triangle t > 0$.

Definition 2.2. The quadratic mean derivative relative to the past (quadratic $\mathcal{P}$-mean derivative) $D_2^{\mathcal{P}}\xi(t)$ of $\xi(t)$ at t is L^1-random element of the form

$$D_2^{\mathcal{P}}\xi(t) = \lim_{\triangle t \to +0} E\left(\frac{(\xi(t + \triangle t) - \xi(t)) \otimes (\xi(t + \triangle t) - \xi(t))}{\triangle t} \,\Big|\, \mathcal{P}_t^\xi \right), \tag{2.2}$$

where the limit is assumed to exist in L^1, $\triangle t \to +0$ means that $\triangle t$ tends to 0 and $\triangle t > 0$ and $\otimes$ denotes the tensor product in $\mathbb{R}^n$.

Note that here the tensor product of two vectors in $\mathbb{R}^n$ is the $n \times n$ matrix formed by products of every component of the first vector with every component of the second one. Note also that for column vectors $X, Y \in \mathbb{R}^n$ their tensor product $X \otimes Y$ equals the matrix product XY^* of column vector X and row vector Y^* (transposed column Y).

We denote by $L(\mathbb{R}^n, \mathbb{R}^n)$ the space of linear automorphisms in $\mathbb{R}^n$. Without loss of generality we shall consider points of $L(\mathbb{R}^n, \mathbb{R}^n)$ as $n \times n$ matrices. Recall that the space of such matrices is isomorphic to $\mathbb{R}^{n^2}$. Everywhere below for a set B in $\mathbb{R}^n$ or in $L(\mathbb{R}^n, \mathbb{R}^n)$ we use the norm introduced by usual formula $\|B\| = \sup\limits_{y \in B} \|y\|$. The norm in $\mathbb{R}^n$ is Euclidean, and the norm in $L(\mathbb{R}^n, \mathbb{R}^n)$ is Euclidean in $\mathbb{R}^{n^2}$.

In what follows, for the sake of simplicity of presentation, we shall deal with processes given on a certain finite time interval $t \in [0, T] \subset \mathbb{R}$.

Recall that a stochastic process $\xi(t)$ in $\mathbb{R}^n$, given on a certain probability space $(\Omega, \mathcal{F}, \mathsf{P})$, is called an Itô diffusion type process if there exist: a vector valued process $a(t)$ non-anticipative with respect to $\mathcal{P}_t^\xi$ and such that the Lebesgue integral $\int_0^t a(s)ds$

along sample paths is a.s. well-posed; a matrix-valued process $A(t) = (A_i^j(t))$ non-anticipative with respect to $\mathcal{P}_t^\xi$ and such that $\mathsf{P}\left\{\int_0^T (A_i^j(t))^2 dt < \infty\right\} = 1$ for all i,j; and a Wiener process $w(t)$, adapted to $\mathcal{P}_t^\xi$, such that $\xi(t) = \xi_0 + \int_0^t a(s)ds + \int_0^t A(s)dw(s)$. In particular, this means that the Itô integral $\int_0^t A(s)dw(s)$ is a martingale with respect to $\mathcal{P}_t^\xi$. For simplicity we deal with deterministic initial condition $\xi_0 \in \mathbb{R}^n$. Recall also that diffusion type processes do exist, say, as solutions of the so-called diffusion-type Itô equations (see, e.g., [8, Theorem III.2.4]).

Theorem 2.3. *For the above-mentioned Itô diffusion type process $\xi(t)$ the derivatives $D^{\mathcal{P}}\xi(t)$ and $D_2^{\mathcal{P}}\xi(t)$ exist and take the form $D^{\mathcal{P}}\xi(t) = a(t)$ and $D_2^{\mathcal{P}}\xi(t) = A(t)A^*(t)$ where $A^*(t)$ is the transposed matrix to $A(t)$.*

Proof. Note that $\xi(t + \triangle t) - \xi(t) = \int_t^{t+\triangle t} a(s)ds + \int_t^{t+\triangle t} A(s)dw(s)$. Since the Itô integral is a martingale with respect to $\mathcal{P}_t^\xi$, $E\left(\int_t^{t+\triangle t} A(s)dw(s) \mid \mathcal{P}_t^\xi\right) = 0$ and so we get

$$E(\xi(t + \triangle t) - \xi(t) \mid \mathcal{P}_t^\xi) = E\left(\int\limits_t^{t+\triangle t} a(t)dt \middle| \mathcal{P}_t^\xi\right) = \int\limits_t^{t+\triangle t} E(a(t) \mid \mathcal{P}_t^\xi)dt.$$

Applying formula (2.1) we obtain that $D^{\mathcal{P}}\xi(t) = E(a(t) \mid \mathcal{P}_t^\xi)$. Since $a(t)$ is measurable with respect to $\mathcal{P}_t^\xi$, $E(a(t) \mid \mathcal{P}_t^\xi) = a(t)$.

Taking into account the properties of Lebesgue and Itô integrals and calculating the tensor product as mentioned above, one can see that $(\xi(t+\triangle t)-\xi(t))\otimes(\xi(t+\triangle t)-\xi(t))$ is approximated by $(a(t) \otimes a(t))(\Delta t)^2 + (a(t)\Delta t) \otimes (A(t)\Delta w(t)) + (A(t)\Delta w(t)) \otimes (a(t)\Delta t) + A(t)A^*(t)\Delta t$. Application of formula (2.2) and of the fact that $A(t)A^*(t)$ is measurable with respect to $\mathcal{P}_t^\xi$, yields $D_2^{\mathcal{P}}\xi(t) = A(t)A^*(t)$. □

By $\mathrm{S}(n)$ we denote the linear space of symmetric $n \times n$ matrices that is a subspace in $L(\mathbb{R}^n, \mathbb{R}^n)$. The symbol $\mathrm{S}_+(n)$ denotes the set of positive definite symmetric $n \times n$ matrices that is a convex open set in $\mathrm{S}(n)$. Note that for each matrix from $\mathrm{S}_+(n)$ its trace is a positive real number. The closure of $\mathrm{S}_+(n)$, i.e., the set of positive semi-definite symmetric $n \times n$ matrices, is denoted by $\bar{\mathrm{S}}_+(n)$.

Note that for $A \in L(\mathbb{R}^n, \mathbb{R}^n)$ the matrix product AA^* is a symmetric positive semi-definite matrix. Thus, from Theorem 2.3 it follows that for an Itô diffusion type process $\xi(t)$ its derivative $D_2^{\mathcal{P}}\xi(t)$ takes values in $\bar{S}_+(n)$.

3 Equations with $\mathcal{P}$-mean Derivatives

In what follows, for the sake of simplicity of presentation, we shall deal with processes given on a certain finite time interval $t \in [0,T] \subset \mathbb{R}$.

Introduce $\tilde{\Omega} = C^0([0,T],\mathbb{R}^n)$ – the Banach space of continuous curves in $\mathbb{R}^n$ given on $[0,T]$, with usual uniform norm – and the σ-algebra $\tilde{\mathcal{F}}$ on $\tilde{\Omega}$ generated by cylinder sets. By $\mathcal{P}_t$ we denote the σ-subalgebra of $\mathcal{F}$ generated by cylinder sets with bases over $[0,t] \subset [0,T]$. Recall that $\tilde{\mathcal{F}}$ is the Borel σ-algebra on $\tilde{\Omega}$ (see [16]).

Let $a : [0,T] \times \tilde{\Omega} \to \mathbb{R}^n$ and $\alpha : [0,T] \times \tilde{\Omega} \to \bar{S}_+(n)$ be measurable mappings. The first problem is to find a stochastic process $\xi(t)$ whose forward and quadratic $\mathcal{P}$-mean derivatives at each t are $a(t,\xi(\cdot))$ and $\alpha(t,\xi(\cdot))$, respectively.

Definition 3.1. The equation with $\mathcal{P}$-mean derivatives is a system of the form

$$\begin{cases} D^{\mathcal{P}}\xi(t) = a(t,\xi(\cdot)), \\ D_2^{\mathcal{P}}\xi(t) = \alpha(t,\xi(\cdot)). \end{cases} \tag{3.1}$$

Definition 3.2. We say that equation (3.1) has a weak solution $\xi(t)$ if there exists a probability space $(\Omega,\mathcal{F},\mathsf{P})$ and a stochastic process $\xi(t)$, given on $(\Omega,\mathcal{F},\mathsf{P})$ and taking values in $\mathbb{R}^n$, such that equation (3.1) is fulfilled P-a.s.

For simplicity we deal with deterministic initial conditions only.

Let $B : [0,T] \times \tilde{\Omega} \to Z$ be a mapping to some metric space Z. Below we shall often suppose that such mappings with various spaces Z satisfy the following condition:

Condition 3.3. For each $t \in [0,T]$ from the fact that the curves $x_1(\cdot), x_2(\cdot) \in \tilde{\Omega}$ coincide for $0 \le s \le t$, it follows that $B(t,x_1(\cdot)) = B(t,x_2(\cdot))$.

Remark 3.4. Note that the fact that a mapping B satisfies Condition 3.3 is equivalent to the fact that B at each t is measurable with respect to Borel σ-algebra in Z and $\mathcal{P}_t$ in $\tilde{\Omega}$ (see [8]).

Lemma 3.5. *For a continuous (measurable and C^k-smooth for $k \ge 1$) mapping α : $[0,T] \times \tilde{\Omega} \to S_+(n)$ satisfying Condition 3.3, there exists a continuous (measurable, C^k-smooth, respectively) mapping $A : [0,T] \times \tilde{\Omega} \to L(\mathbb{R}^n,\mathbb{R}^n)$ that satisfies Condition 3.3 and such that $\alpha(t,x(\cdot)) = A(t,x(\cdot))A^*(t,x(\cdot))$ for each $(t,x(\cdot)) \in \mathbb{R} \times \tilde{\Omega}$.*

Proof. Since the symmetric matrices $\alpha(t,x(\cdot)) \in S_+(n)$ are positive definite, all diagonal minors of $\alpha(t,x(\cdot))$ are positive and, in particular, are not equal to zero. Then for $\alpha(t,x(\cdot))$ the Gauss decomposition is valid (see [18, Theorem II.9.3]), i.e., there exist unique triple of matrices: $\zeta(t,x(\cdot))$, a lower-triangle matrix with units on the diagonal, $z(t,x(\cdot))$, an upper-triangle matrix with units on the diagonal, and $\delta(t,x(\cdot))$, a diagonal matrix such that $\alpha(t,x(\cdot)) = \zeta(t,x(\cdot))\delta(t,x(\cdot))z(t,x(\cdot))$. In addition, the elements of matrices $\zeta(t,x(\cdot))$, $\delta(t,x(\cdot))$ and $z(t,x(\cdot))$ are rationally expressed via the elements of $\alpha(t,x(\cdot))$. Hence, if the matrices $\alpha(t,x(\cdot))$ are continuous (measurable, smooth) jointly in $t, x(\cdot)$, the matrices $\zeta(t, x(\cdot))$, $\delta(t, x(\cdot))$ and $z(t, x(\cdot))$ are also continuous (measurable, smooth, respectively) jointly in variables $t, x(\cdot)$. From the fact that $\alpha(t, x(\cdot))$ is a symmetric matrix, one can easily derive that $z(t, x(\cdot)) = \zeta^*(t, x(\cdot))$ (i.e., $z(t, x(\cdot))$ equals the transposed $\zeta(t, x(\cdot))$). Besides, the elements of diagonal matrix $\delta(t, x(\cdot))$ equal to diagonal minors of $\alpha(t, x(\cdot))$ and so they are positive. Thus the diagonal matrix $\sqrt{\delta(t, x(\cdot))}$ is well-posed: its diagonal contains the square roots of the corresponding diagonal elements of $\delta(t, x(\cdot))$. Consider the matrix $A(t, x(\cdot)) = \zeta(t, x(\cdot))\sqrt{\delta(t, x(\cdot))}$. By construction, $A(t, x(\cdot))$ is jointly continuous (measurable, smooth, respectively) in $t, x(\cdot)$ and

$$A(t, x(\cdot))A^*(t, x(\cdot)) = \zeta(t, x(\cdot))\delta(t, x(\cdot))z(t, x(\cdot)) = \alpha(t, x(\cdot)).$$

The fact that $\zeta(t, x(\cdot))$, $\sqrt{\delta(t, x(\cdot))}$ and so $A(t, x(\cdot))$ satisfy Condition 3.3, follows from the construction. $\square$

Theorem 3.6. *Let $a : [0, T] \times \tilde{\Omega} \to \mathbb{R}^n$ and $\alpha : [0, T] \times \tilde{\Omega} \to S_+(n)$ be jointly continuous in $t, x(\cdot)$ and satisfy Condition 3.3. Let also the following estimates take place:*

$$\operatorname{tr} \alpha(t, x(\cdot)) < K_1(1 + \|x(\cdot)\|)^2, \tag{3.2}$$

$$\|a(t, x(\cdot))\| < K_2(1 + \|x(\cdot)\|). \tag{3.3}$$

Then for every initial condition $\xi_0 \in \mathbb{R}^n$, equation (3.1) *has a weak solution that is well-defined on the entire interval* $[0, T]$.

Proof. Note that $\alpha(t, x(\cdot))$ satisfies the hypothesis of Lemma 3.5 and so there exists continuous $A(t, x(\cdot))$ such that $A(t, x(\cdot))A^*(t, x(\cdot)) = \alpha(t, x(\cdot))$ and $A(t, x(\cdot))$ satisfies Condition 3.3. Immediately from the definition of trace in this case it follows that $\operatorname{tr} \alpha(t, x(\cdot))$ equals the sum of squares of all elements of matrix $A(t, x(\cdot))$, i.e., it is the square of Euclidean norm in $L(\mathbb{R}^n, \mathbb{R}^n)$. Since in the finite dimensional vector space all norms are equivalent, from estimate (3.2) it follows that $\|A(t, x(\cdot))\| < K_3(1 + \|x(\cdot)\|)$ for some $K_3 > 0$. Recall that $a(t, x(\cdot))$ is continuous and satisfies Condition 3.3 and estimate (3.3). Under all these conditions, by [8, Theorem III.2.4] there exists a weak solution $\xi(t)$ of diffusion type stochastic differential equation

$$\xi(t) = \xi_0 + \int_0^t a(s, \xi(\cdot))ds + \int_0^t A(s, \xi(\cdot))dw(s),$$

that is a diffusion type process, well-defined on the entire interval $[0, T]$. From Theorem 2.3 it follows that $\xi(t)$ a.s. satisfies (3.1). $\square$

Remark 3.7. A more general existence result where $\alpha(t, x(\cdot))$ may take values in $\bar{S}_+(n)$, is obtained in Theorem 5.4 below. Its proof follows the scheme of that for Theorem 5.2, an existence of solutions theorem for inclusions with $\mathcal{P}$-mean derivatives.

4 A Technical Result on Set-Valued Mappings

Lemma 4.1. *Specify an arbitrary sequence of positive numbers $\varepsilon_k \to 0$ as $k \to \infty$. Let $\boldsymbol{B}$ be an upper semi-continuous set-valued mapping with compact convex values sending $[0,T] \times \tilde{\Omega}$ to a finite-dimensional Euclidean vector space Y and satisfying Condition 3.3. Then there exists a sequence of continuous single-valued mappings $B_k : [0,T] \times \tilde{\Omega} \to Y$ with the following properties:*

(*i*) *each B_k satisfies Condition 3.3;*

(*ii*) *the sequence B_k point-wise converges to a selector of $\boldsymbol{B}$ that is measurable with respect to Borel σ-algebra in Y and the product σ-algebra of Borel one on $[0,T]$ and $\tilde{\mathcal{F}}$ on $\tilde{\Omega}$;*

(*iii*) *at each $(t, x(\cdot)) \in [0,T] \times \tilde{\Omega}$ the inequality $\|B_k(t, x(\cdot))\| \leq \|\boldsymbol{B}(t, x(\cdot))\|$ holds for all k;*

(*iv*) *if $\boldsymbol{B}$ takes values in a closed convex set $\Xi \subset Y$, the values of all B_k belong to Ξ.*

Proof. In this proof we combine and modify ideas used in the proofs of B.D. Gel'man's result [6, Theorem 2] and our result [4, Theorem 2].

For $t \in [0,T]$ introduce the mapping $f_t : \tilde{\Omega} \to \tilde{\Omega}$ by the formula

$$f_t x(\cdot) = \begin{cases} x(s) & \text{if} \quad 0 \leq s \leq t \\ x(t) & \text{if} \quad t \leq s \leq T. \end{cases} \tag{4.1}$$

Obviously $f_t x(\cdot)$ is continuous jointly in $t \in [0,T]$ and $x(\cdot) \in \tilde{\Omega}$. Since $\boldsymbol{B}$ satisfies Condition 3.3, $\boldsymbol{B}(t, x(\cdot)) = \boldsymbol{B}(t, f_t x(\cdot))$ for each $x(\cdot) \in \tilde{\Omega}$ and $t \in [0,T]$.

Specify an element ε_k from the sequence. Since $\boldsymbol{B}$ is upper semi-continuous, for every $(t, x(\cdot)) \in [0,T] \times \tilde{\Omega}$ there exists $\delta_k(t,x) > 0$ such that for every $(t^*, x^*(\cdot))$ from the $\delta_k(t,x)$ neighbourhood of $(t, x(\cdot))$ the set $\boldsymbol{B}(t^*, x^*(\cdot))$ is contained in the $\frac{\varepsilon_k}{2}$-neighbourhood of the set $\boldsymbol{B}(t, x(\cdot))$. Without loss of generality we can suppose $0 < \delta_k(t,x) < \varepsilon_k$ for every $(t, x(\cdot))$. Consider the $\frac{\delta_k(t,x)}{4}$-neighbourhood of $(t, x(\cdot))$ in $[0,T] \times \tilde{\Omega}$ and construct the open covering of $[0,T] \times \tilde{\Omega}$ by such neighbourhoods for all $(t, x(\cdot))$. Since $[0,T] \times \tilde{\Omega}$ is paracompact, there exists a locally finite refinement $\{V_j^k\}$ of this covering. Without loss of generality we can consider each V_j^k as an $\eta_k(t_j^k, x_j^k)$-neighbourhood of a certain $(t_j^k, x_j^k(\cdot))$ where by construction the radius $\eta_k(t_j^k, x_j^k) \leq \frac{\delta_k(t_j^k, x_j^k)}{4}$.

Consider a continuous partition of unity $\{\varphi_j^k\}$ adapted to $\{V_j^k\}$ and introduce the set-valued mapping $\Phi_k(t, x(\cdot)) = \sum_j \varphi_j^k(t, x(\cdot)) \overline{\mathrm{co}} \boldsymbol{B}(V_j^k)$ where $\overline{\mathrm{co}}$ denotes the convex closure. Since $\boldsymbol{B}(t, x(\cdot))$ is upper semi-continuous and has compact values, without loss of generality we can suppose $\delta_k(t,x)$ to be such that the images $\boldsymbol{B}(V_j^k)$ are bounded in Y and so the sets $\overline{\mathrm{co}}\boldsymbol{B}(V_j^k)$ are compact. Denote by $\overline{\Phi}_k(t,x(\cdot))$ the closure of $\Phi_k(t,x(\cdot))$. Then one can easily see that $\overline{\Phi}_k : [0,T]\times\tilde{\Omega}\to Y$ is a Hausdorff continuous set-valued mapping with compact convex values.

Introduce $\Psi_k : [0,T]\times\tilde{\Omega}\to Y$ by formula $\Psi_k(t,x(\cdot)) = \Phi_k(t, f_t x(\cdot))$ and consider the set-valued mapping $\overline{\Psi}_k(t,x(\cdot))$. Since f_t is continuous, every $\overline{\Psi}_k$ is a Hausdorff continuous set-valued mapping with compact convex values and by construction it satisfies Condition 3.3.

The couple $(t, f_t x(\cdot))$ belongs to a finite collection of neighbourhoods $V_{j_i}^k$ with centers at $(t_{j_i}^k, x_{j_i}^k(\cdot))$, $i=1,\dots,n$ and so by construction $\boldsymbol{B}(t,x(\cdot)) = \boldsymbol{B}(t,f_t x(\cdot)) \subset \boldsymbol{B}(V_{j_i}^k)$ for each i. Hence $\boldsymbol{B}(t,x(\cdot)) = \boldsymbol{B}(t,f_t x(\cdot)) \subset \Psi_k(t,x(\cdot))$ for every couple $(t,x(\cdot))$.

Let l be the number from the collection of indices j_i as above such that $\eta_k(t_l^k, x_l^k)$ takes the greatest value among $\eta_k(t_{j_i}^k, x_{j_i}^k)$. Then all $(t_{j_i}^k, x_{j_i}^k(\cdot))$ are contained in the $2\eta_k(t_l^k,x_l^k)$-neighbourhood of $(t_l^k, x_l^k(\cdot))$ and so every $V_{j_i}^k$ is contained in $3\eta_k(t_l^k,x_l^k)$-neighbourhood of $(t_l^k,x_l^k(\cdot))$ that is contained in $\delta_k(t_l^k,x_l^k(\cdot))$-neighbourhood of (t_l^k,x_l^k) by construction. Hence, also by construction, $\Psi_k(t,x(\cdot))$ belongs to the $\frac{\varepsilon_k}{2}$-neighbourhood of $\boldsymbol{B}(t_l^k,x_l^k(\cdot))$. Since both $\Psi_k(t,x(\cdot))$ and $\boldsymbol{B}(t_l^k,x_l^k(\cdot))$ are convex, this means that $\overline{\Psi}_k(t,x(\cdot))$ also belongs to the $\frac{\varepsilon_k}{2}$-neighbourhood of $\boldsymbol{B}(t_l^k,x_l^k(\cdot))$. Notice that this is true for each k.

Since $\boldsymbol{B}(t,x(\cdot)) \subset \Psi_k(t,x(\cdot)) \subset \overline{\Psi}_k(t,x(\cdot))$, for the Hausdorff submetric $\bar{H}$ we have

$$\bar{H}(\boldsymbol{B}(t,x(\cdot)), \overline{\Psi}_k(t,x(\cdot))) = 0.$$

Hence for the Hausdorff metric H we obtain that

$$H(\overline{\Psi}_k(t,x(\cdot)), \boldsymbol{B}(t,x(\cdot))) = \bar{H}(\overline{\Psi}_k(t,x(\cdot)), \boldsymbol{B}(t,x(\cdot))).$$

Since $\varepsilon_k \to 0$, for $(t,x(\cdot))$ there exists an integer $\theta = \theta(t,x(\cdot)) > 0$ such that $\varepsilon_{k+\theta} < \delta_k(t,x(\cdot))$. Without loss of generality we can suppose that $\theta \geq 1$.

Thus $\boldsymbol{B}(t_l^{k+\theta}, x_l^{k+\theta}(\cdot))$ belongs to the $\frac{\varepsilon_k}{2}$-neighbourhood of $\boldsymbol{B}(t,x(\cdot))$ and so

$$\bar{H}(\boldsymbol{B}(t_l^{k+\theta}, x_l^{k+\theta}(\cdot)), \boldsymbol{B}(t,x(\cdot))) < \frac{\varepsilon_k}{2}.$$

Since $\overline{\Psi}_{k+\theta}(t,x(\cdot))$ belongs to the $\frac{\varepsilon_{k+\theta}}{2}$-neighbourhood of $\boldsymbol{B}(t_l^{k+\theta}, x_l^{k+\theta}(\cdot))$ (see above), we obtain that

$$\bar{H}(\overline{\Psi}_{k+\theta}(t,x(\cdot)), \boldsymbol{B}(t_l^{k+\theta}, x_l^{k+\theta)}(\cdot))) < \frac{\varepsilon_{k+\theta}}{2}.$$

Thus

$$\begin{aligned} H(\overline{\Psi}_{k+\theta}(t,x(\cdot)),\boldsymbol{B}(t,x(\cdot))) &= \bar{H}(\overline{\Psi}_{k+\theta}(t,x(\cdot)),\boldsymbol{B}(t,x(\cdot))) \\ \leq\ & \bar{H}(\overline{\Psi}_{k+\theta}(t,x(\cdot)),\boldsymbol{B}(t_l^{k+\theta},x_l^{k+\theta}(\cdot))) + \bar{H}(\boldsymbol{B}(t_l^{k+\theta},x_l^{k+\theta}(\cdot)),\boldsymbol{B}(t,x(\cdot))) \\ <\ & \frac{\varepsilon_{k+\theta}}{2} + \frac{\varepsilon_k}{2} < \varepsilon_k. \end{aligned}$$

So, at each $(t,x(\cdot))$ we have that $H(\overline{\Psi}_k(t,x(\cdot)),\boldsymbol{B}(t,x(\cdot))) \to 0$ as $k \to \infty$ and $\boldsymbol{B}(t,x(\cdot)) \subset \Psi_k(t,x(\cdot))$ for all k.

Consider the minimal selector $B_k(t,x(\cdot))$ of $\overline{\Psi}_k(t,x(\cdot))$, i.e., $B_k(t,x(\cdot))$ is the closest to origin point in $\overline{\Psi}_i(t,x(\cdot))$. We refer the reader to [1] for complete description of minimal selectors. In particular, it is shown there that minimal selectors in our situation are continuous. By construction all B_k satisfy Condition 3.3.

By construction the minimal selectors $B_k(t,x(\cdot))$ of $\overline{\Psi}_k(t,x(\cdot))$ point-wise converge to the minimal selector $B(t,x(\cdot))$ of $\boldsymbol{B}(t,x(\cdot))$ as $k \to \infty$ since at any $(t,x(\cdot))$ we have that $H(\overline{\Psi}_k(t,x(\cdot)),\boldsymbol{B}(t,x(\cdot))) \to 0$ as $k \to \infty$ and $\boldsymbol{B}(t,x(\cdot)) \subset \Psi_k(t,x(\cdot))$ for all k (see above). It is a well-known fact that the point-wise limit B of the sequence of continuous mappings B_k is measurable with respect to Borel σ-algebras in Y and in $[0,T] \times \tilde{\Omega}$ (see [15]). The latter coincides with the product σ-algebra of Borel one on $[0,T]$ and $\tilde{\mathcal{F}}$ on $\tilde{\Omega}$ (see [16]). Properties (iii) and (iv) immediately follow from the construction. □

Remark 4.2. Note that unlike $\bar{\Psi}_k(t,x(\cdot))$, the set-valued mapping $\bar{\Phi}_k(t,x(\cdot))$ may not satisfy Condition 3.3 since two different curves $x_1(\cdot)$ and $x_2(\cdot)$ coinciding on $[0,t]$, may have different neighbourhoods V_j^k, to which they belong, and so the values $\bar{\Phi}_k(t,x_1(\cdot))$ and $\bar{\Phi}_k(t,x_2(\cdot))$ may be different. On the other hand, it follows from [6] that $\bar{\Phi}_k$ is an ε_k-approximation of $\boldsymbol{B}$ while it is not true for $\bar{\Psi}_k$.

5 Differential Inclusions with $\mathcal{P}$-mean Derivatives

Consider set-valued mappings $\boldsymbol{a}(t,x)$ and $\boldsymbol{\alpha}(t,x)$ sending $[0,T] \times \tilde{\Omega}$ to $\mathbb{R}^n$ and $\overline{S}_+(n)$, respectively, and satisfying Condition 3.3. The differential inclusion with forward $\mathcal{P}$-mean derivatives is a system of the form

$$\begin{cases} D^{\mathcal{P}}\xi(t) \in \boldsymbol{a}(t,\xi(\cdot)), \\ D_2^{\mathcal{P}}\xi(t) \in \boldsymbol{\alpha}(t,\xi(\cdot)). \end{cases} \tag{5.1}$$

Definition 5.1. We say that inclusion (5.1) has a weak solution with initial condition $\xi_0 \in \mathbb{R}^n$ if there exists a probability space and a stochastic process $\xi(t)$ given on it and taking values in $\mathbb{R}^n$, such that $\xi(0) = \xi_0$ and a.s. $\xi(t)$ satisfies inclusion (5.1).

As well as in Section 3 we deal with deterministic initial conditions only.

If, say, $\boldsymbol{a}(t,x)$ and $\boldsymbol{\alpha}(t,x)$ are lower semi-continuous and have closed convex values, then by Michael's theorem they have continuous selectors $a(t,x(\cdot))$ and $\alpha(t,x(\cdot))$, respectively. If those selectors satisfy conditions of Theorem 3.6, the weak solution of (3.1) with coefficients $a(t,x(\cdot))$ and $\alpha(t,x(\cdot))$ that exists by Theorem 3.6, is obviously a weak solution of (5.1).

The main result of this paper is the following existence theorem for the case where $\boldsymbol{a}(t,x)$ and $\boldsymbol{\alpha}(t,x)$ are upper semi-continuous.

Theorem 5.2. *Let $\boldsymbol{\alpha}(t,x) : [0,T]\times\tilde{\Omega} \to \overline{S}_+(n)$ be an upper semi-continuous set-valued mapping with closed convex values that satisfies Condition 3.3 and let for every $\alpha(t,x(\cdot)) \in \boldsymbol{\alpha}(t,x(\cdot))$ the estimate*

$$\operatorname{tr}\alpha(t,x(\cdot)) < K_1(1+\|x(\cdot)\|)^2 \tag{5.2}$$

hold for some $K_1 > 0$.

Let $\boldsymbol{a}(t,x(\cdot))$ be an upper semi-continuous set-valued mapping from $[0,T]\times\tilde{\Omega}$ to $\mathbb{R}^n$ with closed convex values that satisfies Condition 3.3 and the estimate

$$\|\boldsymbol{a}(t,x(\cdot))\| < K_2(1+\|x(\cdot)\|) \tag{5.3}$$

for some $K_2 > 0$.

Then for any initial condition $\xi(0)\in\mathbb{R}^n$, inclusion (5.1) *has a weak solution.*

Proof. Choose a sequence of positive numbers $\varepsilon_k \to 0$. The set-valued mapping $\boldsymbol{a}(t,x(\cdot))$ satisfies the conditions of Lemma 4.1 and so there exists a sequence of continuous single-valued mappings $a_k : [0,T]\times\tilde{\Omega} \to \mathbb{R}^n$ that point-wise converges to a certain measurable selector $a(t,x(\cdot))$ of $\boldsymbol{a}(t,x(\cdot))$ and every $a_k(t,x(\cdot))$ satisfies Condition 3.3 and the estimate

$$\|a_k(t,x(\cdot))\| < K_2(1+\|x(\cdot)\|). \tag{5.4}$$

The mapping $\boldsymbol{\alpha}(t,x(\cdot))$ that takes values in the closed convex set $\bar{S}_+(n)$ in the space of all symmetric $n\times n$ matrices, also satisfies the conditions of Lemma 4.1 and so there exists a sequence of continuous single-valued mappings $\tilde{\alpha}_k : [0,T]\times\tilde{\Omega} \to \bar{S}_+(n)$ that point-wise converges to a measurable selector $\alpha(t,x(\cdot))$ of $\boldsymbol{\alpha}(t,x(\cdot))$ and every $\tilde{\alpha}_k(t,x(\cdot))$ satisfies Condition 3.3 and the estimate

$$\operatorname{tr}\tilde{\alpha}_k(t,x(\cdot)) < K_1(1+\|x(\cdot)\|)^2. \tag{5.5}$$

Create another sequence $\alpha_k(t,x(\cdot)) = \tilde{\alpha}_k(t,x(\cdot)) + \varepsilon_k I$ where I is the unit matrix, that evidently point-wise converges to $\alpha(t,x(\cdot))$ as well. All mappings $\alpha_k(t,x(\cdot))$ are continuous, satisfy Condition 3.3 and estimate (5.5) – at least for k large enough – and in addition they all take values in the open set $S_+(n)$ of positive definite symmetric matrices. Thus by Lemma 3.5 for every $\alpha_k(t,x(\cdot))$ there exists continuous $A_k : [0,T]\times\tilde{\Omega} :\to L(\mathbb{R}^n,\mathbb{R}^n)$ such that $\alpha_k(t,x(\cdot)) = A_k(t,x(\cdot))A_k^*(t,x(\cdot))$ and all $A(t,x(\cdot))$ satisfy Condition 3.3.

As well as in Theorem 3.3, immediately from the definition of trace in this case it follows that $\operatorname{tr}\alpha_k(t,x(\cdot))$ equals the sum of squares of all elements of matrix $A_k(t,x(\cdot))$, i.e., it is the square of Euclidean norm of $A_k(t,x(\cdot))$ in $L(\mathbb{R}^n,\mathbb{R}^n)$. Hence, from (5.5) it follows that

$$\|A_k(t,x(\cdot))\| < \sqrt{K_1}(1+\|x(\cdot)\|). \tag{5.6}$$

Thus from (5.4) and (5.6) it follows that for each k the couple $(a_k(t,x(\cdot)),A_k(t,x(\cdot)))$ satisfies the so-called Itô condition

$$\|a_k(t,x(\cdot))\| + \|A_k(t,x(\cdot))\| < K(1+\|x(\cdot)\|) \tag{5.7}$$

with a certain $K>0$ the same for all k.

Consider the sequence of diffusion type Itô stochastic differential equations

$$\xi_k(t) = \xi_0 + \int_0^t a_k(s,\xi_k(\cdot))ds + \int_0^t A_k(s,\xi_k(\cdot))dw(s). \tag{5.8}$$

Since their coefficients are continuous, satisfy Condition 3.3 and estimate (5.7) with the same K, by [8, Theorem III.2.4] they all have weak solutions $\xi_k(t)$, well-posed on the entire interval $[0,T]$, and the set of measures μ_k generated by $\xi_k(t)$ on $(\tilde{\Omega},\tilde{\mathcal{F}})$, is weakly compact (see [8, Corollary to Lemma III.2.2]). Hence we can choose a subsequence (we keep notation μ_k for this subsequence) that weakly converges to a certain probability measure μ. Denote by $\xi(t)$ the coordinate process on probability space $(\tilde{\Omega},\tilde{\mathcal{F}},\mu)$ (recall: this means that $\xi(t,x(\cdot)) = x(t)$).

Show that $\xi(t)$ is a solution we are looking for. First of all note that $\mathcal{P}_t$ is the “past” σ-algebra of $\xi(t)$.

Denote by λ the normalized Lebesgue measure on $[0,T]$. Introduce measures ν_k on $(\tilde{\Omega},\tilde{\mathcal{F}})$ by the relations $d\nu_k = (1+\|x(\cdot)\|)d\mu_k$. It is a well known fact (see, e.g., [8] or [4, Lemma 4]) that ν_k weakly converge to the measure ν defined by relation $d\nu = (1+\|x(\cdot)\|)d\mu$.

As $a_i(t,x(\cdot))$ converge as $i\to\infty$ to $a(t,x(\cdot))$ point-wise, it converges a.s. with respect to all $\lambda\times\mu_k$, and so the functions $\dfrac{a_i(t,x(\cdot))}{1+\|x(\cdot)\|}$ converge to $\dfrac{a(t,x(\cdot))}{1+\|x(\cdot)\|}$ a.s. with respect to all $\lambda\times\nu_k$. Specify $\delta>0$. By Egorov’s theorem (see, e.g., [17]) for any k there exists a subset $\tilde{K}^k_\delta \subset [0,T]\times\tilde{\Omega}$ such that $(\lambda\times\nu_k)(\tilde{K}^k_\delta) > 1-\delta$, and the sequence $\dfrac{a_i(t,x(\cdot))}{1+\|x(\cdot)\|}$ converges to $\dfrac{a(t,x(\cdot))}{1+\|x(\cdot)\|}$ uniformly on $\tilde{K}^k_\delta$. Introduce $(\tilde{K}_\delta = \bigcup_{i=0}^{\infty}\tilde{K}^k_\delta)$. The sequence $\dfrac{a_i(t,x(\cdot))}{1+\|x(\cdot)\|}$ converges to $\dfrac{a(t,x(\cdot))}{1+\|x(\cdot)\|}$ uniformly on $\tilde{K}_\delta$ and $(\lambda\times\nu_k)(\tilde{K}_\delta) > 1-\delta$ for all $k=0,\ldots,\infty$.

Notice that $a(t,x(\cdot))$ is continuous on a set of full measure $\lambda\times\nu$ on $[0,T]\times\tilde{\Omega}$. Indeed, consider a sequence $\delta_k\to 0$ and the corresponding sequence $\tilde{K}_{\delta_k}$ from Egorov’s theorem. By the above construction $a(t,x(\cdot))$ is a uniform limit of continuous functions on each $\tilde{K}_{\delta_k}$. Thus it is continuous on each $\tilde{K}_{\delta_k}$ and so on every finite union $\bigcup_{i=1}^{n}\tilde{K}_{\delta_i}$. Evidently $\lim_{n\to\infty}(\lambda\times\nu)(\bigcup_{i=1}^{n}\tilde{K}_{\delta_i}) = (\lambda\times\nu)([0,T]\times\tilde{\Omega})$.

Hence $\dfrac{a(t,x(\cdot))}{1+\|x(\cdot)\|}$ is continuous on a set of full measure $\lambda\times\nu$ on $[0,T]\times\tilde{\Omega}$.

Let $g_t(x(\cdot))$ be a bounded (say, $|g_t(x(\cdot))| < \Theta$ for all $x(\cdot)\in\tilde{\Omega}$) and continuous $\mathcal{P}_t$-measurable function on $\tilde{\Omega}$.

Because of the above uniform convergence on $\tilde{K}_\delta$ for all k and boundedness of g_t we get that for k large enough

$$\left\|\int_{\tilde{K}_\delta}\left(\int_t^{t+\Delta t}(a_k(\tau,x(\cdot)) - a(\tau,x(\cdot)))d\tau\right)g_t(x(\cdot))d\mu_k\right\|$$
$$= \left\|\int_{\tilde{K}_\delta}\left(\int_t^{t+\Delta t}\frac{a_k(\tau,x(\cdot)) - a(\tau,x(\cdot))}{1+\|x(\cdot)\|}d\tau\right)g_t(x(\cdot))d\nu_k\right\| < \delta.$$

Since $(\lambda\times\mu_k)(\tilde{K}_\delta) > 1-\delta$ for all k, $\left\|\dfrac{a_k(t,x(\cdot)) - a(t,x(\cdot))}{1+\|x(\cdot)\|}\right\| < Q$ for all k and $|g_t(x(\cdot))| < \Theta$ (see above), we get

$$\left\|\int_{\tilde{\Omega}\setminus\tilde{K}_\delta}\left(\int_t^{t+\Delta t}(a_k(\tau,x(\cdot)) - a(\tau,x(\cdot)))d\tau\right)g_t(x(\cdot))d\mu_k\right\|$$
$$= \left\|\int_{\tilde{\Omega}\setminus\tilde{K}_\delta}\left(\int_t^{t+\Delta t}\frac{a_k(\tau,x(\cdot)) - a(\tau,x(\cdot))}{1+\|x(\cdot)\|}d\tau\right)g_t(x(\cdot))d\nu_k\right\| < 2Q\Theta\delta.$$

From the fact that δ is an arbitrary positive number it follows that

$$\lim_{k\to\infty}\int_{\tilde{\Omega}}\left(\int_t^{t+\Delta t}a_k(\tau,x(\cdot))d\tau - \int_t^{t+\Delta t}a(\tau,x(\cdot))d\tau\right)g_t(x(\cdot))d\mu_k = 0.$$

The function $\dfrac{a(t,x(\cdot))}{1+\|x(\cdot)\|}$ is $\lambda\times\nu$-a.s. continuous (see above) and bounded on $[0,T]\times\tilde{\Omega}$. Hence by [7, Lemma in Section VI.4] from the weak convergence of ν_k to ν it follows that

$$\begin{aligned}
&\lim_{k\to\infty}\int_{\tilde{\Omega}}\left(\int_t^{t+\Delta t}a(\tau,x(\cdot))d\tau\right)g_t(x(\cdot))d\mu_k \\
&= \lim_{k\to\infty}\int_{\tilde{\Omega}}\left(\int_t^{t+\Delta t}\frac{a(\tau,x(\cdot))}{1+\|x(\cdot)\|}d\tau\right)g_t(x(\cdot))d\nu_k \\
&= \int_{\tilde{\Omega}}\left(\int_t^{t+\Delta t}\frac{a(\tau,x(\cdot))}{1+\|x(\cdot)\|}d\tau\right)g_t(x(\cdot))d\nu \\
&= \int_{\tilde{\Omega}}\left(\int_t^{t+\Delta t}a(\tau,x(\cdot))d\tau\right)g_t(x(\cdot))d\mu.
\end{aligned} \tag{5.9}$$

Obviously

$$\lim_{k\to\infty}\int_{\tilde\Omega}(x(t+\Delta t)-x(t))d\mu_k=\lim_{k\to\infty}\int_{\tilde\Omega}\frac{x(t+\Delta t)-x(t)}{1+\|x(\cdot)\|}d\nu_k$$
$$=\int_{\tilde\Omega}\frac{x(t+\Delta t)-x(t)}{1+\|x(\cdot)\|}d\nu=\int_{\tilde\Omega}(x(t+\Delta t)-x(t))d\mu. \tag{5.10}$$

Notice that

$$\int_{\tilde\Omega}\left(\Big[x(t+\Delta t)-x(t)\Big]-\int_t^{t+\Delta t}a_k(\tau,x(\cdot))d\tau\right)g_t(x(\cdot))d\mu_k=0 \tag{5.11}$$

since

$$\int_{\tilde\Omega}[x(t+\Delta t)-x(t)]g_t(x(\cdot))d\mu_k=E\left[(\xi_k(t+\Delta t)-\xi_k(t))g_t(\xi_k(t))\right],$$
$$\int_{\tilde\Omega}\left(\int_t^{t+\Delta t}a_k(\tau,x(\cdot))d\tau\right)g_t(x(\cdot))d\mu_k=E\Big[\left(\int_t^{t+\Delta t}a_k(\tau,\xi_k(\tau))d\tau\right)g_t(\xi_k(t))\Big]$$

and $\xi_k(t)$ is a solution of (5.8). Formulae (5.9), (5.10) and (5.11) yield

$$\int_{\tilde\Omega}\left(\Big[x(t+\Delta t)-x(t)\Big]-\int_t^{t+\Delta t}a(s,x(\cdot))ds\right)g_t(x(\cdot))d\mu=0.$$

Since g_t is an arbitrary continuous bounded function measurable with respect to $\mathcal{P}_t$, the last relation is equivalent to

$$E\left(\Big[\xi(t+\Delta t)-\xi(t)\Big]-\int_t^{t+\Delta t}a(s,\xi(\cdot))ds\,\Big|\,\mathcal{P}_t\right)=0. \tag{5.12}$$

From (5.12) it evidently follows that

$$D^{\mathcal{P}}\xi(t)=a(t,\xi(\cdot))\subset \boldsymbol{a}(t,\xi(\cdot)) \tag{5.13}$$

and that the process $\xi(t)-\int_0^t a(s,\xi(\cdot))ds$ is a martingale on $(\tilde\Omega,\tilde{\mathcal{F}},\mu)$ with respect to $\mathcal{P}_t$.

Now turn to $A_k(t,x(\cdot))$. Recall that $\alpha_k(t,x(\cdot))=A_k(t,x(\cdot))A_k^*(t,x(\cdot))$ point-wise converge to $\alpha(t,x(\cdot))$, a measurable selector of $\boldsymbol{\alpha}(t,x(\cdot))$. In complete analogy with the above arguments one can show that

$$\int_{\tilde\Omega}\left(\Big[(x(t+\Delta t)-x(t))\otimes(x(t+\Delta t)-x(t))\Big]-\int_t^{t+\Delta t}\alpha(s,x(\cdot))ds\right)g_t(x(\cdot))d\mu=0 \tag{5.14}$$

with g_t as above. Relation (5.14) is equivalent to

$$E\left(\left[(\xi(t+\Delta t)-\xi(t))\otimes(\xi(t+\Delta t)-\xi(t))\right]-\int\limits_t^{t+\Delta t}\alpha(s,\xi(\cdot)))ds\,\Big|\,\mathcal{P}_t\right)=0$$

from which it evidently follows that

$$D_2^{\mathcal{P}}\xi(t)=\alpha(t,\xi(\cdot))\subset\boldsymbol{\alpha}(t,\xi(\cdot)) \tag{5.15}$$

and that the process $[\xi(t)\otimes\xi(t)]-\int_0^t\alpha(t,\xi(\cdot))dt$ is a martingale on $(\tilde{\Omega},\tilde{\mathcal{F}},\mu)$ with respect to $\mathcal{P}_t$. Relations (5.13) and (5.15) mean that $\xi(t)$ is a solution of (5.1) that we are looking for. □

Remark 5.3. From (5.12) it evidently follows that the solution $\xi(t)$ of (5.1) obtained in Theorem 5.2, is a semi-martingale with respect to $\mathcal{P}_t$ since $\xi(t)-\int_0^t a(s,\xi(\cdot))ds$ is a martingale with respect to $\mathcal{P}_t$.

Theorem 5.4. *Let $a(t,x(\cdot))$ and $\alpha(t,x(\cdot))$ be as in Theorem 3.6 but α sends $[0,T]\times\tilde{\Omega}$ to $\bar{S}_+(n)$ instead of $S_+(n)$. Then for every initial condition $\xi_0\in\mathbb{R}^n$ equation* (3.1) *has a weak solution that is well-defined on the entire interval $[0,T]$.*

Indeed, we can construct a sequence of continuous single-valued mappings $\alpha_k=\alpha+\varepsilon_k I:[0,T]\times\tilde{\Omega}\to S_+(n)$ satisfying Condition 3.3, that converge to α. Then the proof of Theorem 5.4 follows the same scheme as that of Theorem 5.2.

Acknowledgements

This research is supported in part by RFBR Grants 07–01–00137 and 08–01–00155. We express our thanks to the anonymous referees for very much useful advice.

References

[1] Aubin J.-P., Cellina A., Differential Inclusions. Set-valued maps and viability theory, Berlin, Springer-Verlag, 1984.

[2] Azarina S.V., Gliklikh Yu.E., Differential equations and inclusions with forward mean derivatives in $\mathbb{R}^n$, Proceedings of Voronezh State University, 2006, No. 2., 138–146.

[3] Azarina S.V., Gliklikh Yu.E., Differential inclusions with mean derivatives, Dynamic Systems and Applications, 2007, Vol.16, No. 1., 49–71.

[4] Azarina S.V., Gliklikh Yu.E. and Obukhovskiĭ A.V., Solvability of Langevin differential inclusions with set-valued diffusion terms on Riemannian manifolds, Applicable Analysis, 2007, Vol. 86, No. 9., 1105–1116.

[5] Borisovich Yu.G., Gelman B.D., Myshkis A.D., Obukhovskii V.V., Introduction to the theory of multi-valued mappings and differential inclusions, Moscow: KomKniga (URSS), 2005 (in Russian).

[6] Gel'man B.D., Continuous approximations of set-valued mappings and fixed points, Mathematical Notices, 2005, Vol. 78, No. 2., 212–222.

[7] Gihman I.I., Skorohod A.V., Theory of Stochastic Processes. Vol.1, New York, Springer–Verlag, 1974.

[8] Gihman I.I., Skorohod A.V., Theory of Stochastic Processes. Vol.3, New York, Springer–Verlag, 1979.

[9] Gliklikh Yu.E., Global Analysis in Mathematical Physics. Geometric and Stochastic Methods, New York, Springer–Verlag, 1997.

[10] Gliklikh Yu.E., Deterministic viscous hydrodynamics via stochastic processes on groups of diffeomorphisms, Probabilistic Metohds in Fluids (I.M. Davis et al., eds), Singapore, World Scientific, 2003, 179–190.

[11] He X., A probabilistic method for Navier–Stokes vorticies, J. Appl. Probab., 2001, 1059–1066.

[12] Nelson E., Derivation of the Schrödinger equation from Newtonian mechanics, Phys. Reviews, 1966, Vol. 150, No. 4, 1079–1085.

[13] Nelson E., Dynamical theory of Brownian motion, Princeton, Princeton University Press, 1967.

[14] Nelson E., Quantum fluctuations, Princeton, Princeton University Press, 1985.

[15] Parthasarathy K.R., Introduction to Probability and Measure, New York, Springer–Verlag, 1978.

[16] Shiryaev A.N., Probability, New York, Springer–Verlag, 1984.

[17] Yosida Y. Functional Analysis, Berlin, Springer–Verlag, 1965.

[18] Zhelobenko D.P., Compact Lie groups and their representations, Providence, RI, Amer. Math. Soc., 1973.