Preface

Stochastic control theory is a relatively young branch of mathematics. The beginning of its intensive development falls in the late 1950s and early 1960s. During that period an extensive literature appeared on optimal stochastic control using the quadratic performance criterion (see references in Wonham [76]). At the same time, Girsanov [25] and Howard [26] made the first steps in constructing a general theory, based on Bellman's technique of dynamic programming, developed by him somewhat earlier [4].

Two types of engineering problems engendered two different parts of stochastic control theory. Problems of the first type are associated with multistep decision making in discrete time, and are treated in the theory of discrete stochastic dynamic programming. For more on this theory, we note in addition to the work of Howard and Bellman, mentioned above, the books by Derman [8], Mine and Osaki [55], and Dynkin and Yushkevich [12].

Another class of engineering problems which encouraged the development of the theory of stochastic control involves time continuous control of a dynamic system in the presence of random noise. The case where the system is described by a differential equation and the noise is modeled as a time continuous random process is the core of the optimal control theory of diffusion processes. This book deals with this latter theory.

The mathematical theory of the evolution of a system usually begins with a differential equation of the form

$$\dot{x}_t = f(t,x_t)$$

with respect to the vector of parameters x of such a system. If the function $f(t,x)$ can be measured or completely defined, no stochastic theory is needed. However, it is needed if $f(t,x)$ varies randomly in time or if the errors of measuring this vector cannot be neglected. In this case $f(t,x)$ is, as a rule, representable as $b(t,x) + \sigma(t,x)\dot{\xi}_t$ where b is a vector, σ is a matrix, and ξ_t is a random vector process. Then

$$\dot{x}_t = b(t,x_t) + \sigma(t,x_t)\dot{\xi}_t. \tag{1}$$

It is convenient to write the equation in the integral form

$$x_t = x_0 + \int_0^t b(t,x_s)\,ds + \int_0^t \sigma(s,x_s)\,d\xi_s, \tag{2}$$

where x_0 is the vector of the initial state of the system. We explain why Eq. (2) is preferable to Eq. (1). Usually, one tries to choose the vector of parameters x_t of the system in such a way that the knowledge of them at time t enables one to predict the probabilistic behavior of the system after time t with the same certainty (or uncertainty) to the same extent as would knowledge of the entire prior trajectory x_s $(s \leq t)$. Such a choice of parameters is convenient because the vector x_t contains all the essential information about the system. It turns out that if x_t has this property, it can be proved under rather general conditions that the process ξ_t in (2) can be taken to be a Brownian motion process or, in other words, a Wiener process $\mathbf{w}_t$. The derivative of ξ_t is then the so-called "white noise," but, strictly speaking, $\dot{\xi}_t$ unfortunately cannot be defined and, in addition, Eq. (1) has no immediate meaning. However, Eq. (2) does make sense, if the second integral in (2) is defined as an Ito stochastic integral.

It is common to say that the process x_t satisfying Eq. (2) is a diffusion process. If, in addition, the coefficients b, σ of Eq. (2) depend also on some control parameters, we have a "controlled diffusion process."

The main subject matter of the book having been outlined, we now indicate how some parts of optimal control theory are related to the contents of the book.

Formally, the theory of deterministic control systems can be viewed as a special case of the theory of stochastic control. However, it has its own unique characteristics, different from those of stochastic control, and is not considered here. We mention only a few books in the enormous literature on the theory of deterministic control systems: Pontryagin, Boltyansky, Gamkrelidze, and Mishchenko [60] and Krassovsky and Subbotin [27].

A considerable number of works on controlled diffusion processes deal with control problems of linear systems of type (2) with a quadratic performance criterion. Besides Wonham [76] mentioned above, we can also mention Astrom [2] and Bucy and Joseph [7] as well as the literature cited in those books. We note that the control of such systems necessitates the construction of the so-called Kalman–Bucy filters. For the problems of the application of filtering theory to control it is appropriate to mention Lipster and Shiryayev [51].

Since the theory of linear control systems with quadratic performance index is represented well in the literature, we shall not discuss it here.

Control techniques often involve rules for stopping the process. A general and rather sophisticated theory of optimal stopping rules for Markov chains and Markov processes, developed by many authors, is described by Shiryayev [69]. In our book, problems of optimal stopping also receive considerable attention. We consider such problems for controlled processes with the help of the method of randomized stopping. It must be admitted, however, that our theory is rather crude compared to the general theory presented in [69] because of the fact that in the special case of controlled diffusion processes, imposing on the system only simply verifiable (and therefore crude) restrictions, we attempt to obtain strong assertions on the validity of the Bellman equation for the payoff function.

Concluding the first part of the Preface, we emphasize that in general the main aim of the book is to prove the validity of the Bellman differential equations for payoff functions, as well as to develop (with the aid of such equations) rules for constructing control strategies which are close to optimal for controlled diffusion processes.

A few remarks on the structure of the book may be helpful. The literature cited so far is not directly relevant to our discussion. References to the literature of more direct relevance to the subject of the book are given in the course of the presentation of the material, and also in the notes at the end of each chapter.

We have discussed only the main features of the subject of our investigation. For more detail, we recommend Section 1, of Chapter 1, as well as the introductions to Chapters 1–6.

The text of the book includes theorems, lemmas, and definitions, numeration of which is carried out throughout according to a single system in each section. Thus, the invoking of Theorem 3.1.5 means the invoking of the assertions numbered 5 in Section 1 in Chapter 3. In Chapter 3, Theorem 3.1.5 is referred to as Theorem 1.5, and in Section 1, simply as Theorem 5. The formulas are numbered in a similar way.

The initial constants appearing in the assumptions are, as a rule, denoted by K_i, δ_i. The constants in the assertions and in the proofs are denoted by the letter N with or without numerical subscripts. In the latter case it is assumed that in each new formula this constant is generally speaking unique to the formula and is to be distinguished from the previous constants. If we write $N = N(K_i, \delta_i, \ldots)$, this means that N depends only on what is inside the parentheses. The discussion of the material in each section is carried out under the same assumptions listed at the start of the section. Occasionally, in order to avoid the cumbersome formulation of lemmas and theorems, additional assumptions are given prior to the lemmas and theorems rather than in them.

Reading the book requires familiarity with the fundamentals of stochastic integral theory. Some material on this theory is presented in Appendix 1. The Bellman equations which we shall investigate are related to nonlinear partial differential equations. We note in this connection that we do not assume the reader to be familiar with the results related to differential equation theory.

In conclusion, I wish to express my deep gratitude to A. N. Shiryayev and all participants of the seminar at the Department of Control Probability of the Interdepartmental Laboratory of Statistical Methods of the Moscow State University for their assistance in our work in this book, and for their useful criticism of the manuscript.

N. V. Krylov

Auxiliary Propositions 2

1. Notation and Definitions

In addition to the notation given on pages xi and xii we shall use the following:

T is a nonnegative number, and the interval $[0,T]$ is interpreted as an interval of time; the points on this interval are, as a rule, denoted by t, s.

D denotes an open set in Euclidean space, $\bar{D}$ the closure of D, and ∂D the boundary of D.

Q denotes an open set in E_{d+1}; the points of Q are expressed as (t,x) where $t \in E_1, x \in E_d$. $\partial' Q$ denotes the parabolic boundary of Q (see Section 4.5).

$$S_R = \{x \in E_d : |x| < R\}, \qquad C_{T,R} = (0,T) \times S_R, \qquad C_R = C_{\infty,R},$$
$$H_T = (0,T) \times E_d.$$

If ν is a countably additive set function, then $|\nu|$ is the variation of ν, $\nu_+ = \frac{1}{2}(|\nu| + \nu)$ is the positive part of ν, and $\nu_- = \frac{1}{2}(|\nu| - \nu)$ is the negative part of ν.

If Γ denotes a measurable set in Euclidean space, meas Γ is the Lebesgue measure of this set.

For $p \geq 1$ $\mathscr{L}_p(\Gamma)$ denotes a set of real-valued Borel functions $f(x)$ on Γ such that

$$\|f\|_{p,\Gamma} \equiv \left(\int_\Gamma |f(x)|^p \, dx\right)^{1/p} < \infty.$$

In the cases where the middle expression is equal to infinity, we continue to denote it by $\|f\|_{p,\Gamma}$ as before. In general, we admit infinite values for various integrals (and mathematical expectations) of measurable functions. These values are considered to be defined if either the positive part or the negative part of the function has a finite integral. In this case the integral is assumed to be equal to $+\infty$ ($-\infty$) if the integral of the positive (negative) part of the function is infinite.

For any (possibly, nonmeasurable) function $f(x)$ on Γ we define an exterior norm in $\mathscr{L}_p(\Gamma)$, using the formula

$$]\|f\|_{p,\Gamma} = \inf \|h\|_{p,\Gamma},$$

where the lower bound is taken over the set of all Borel functions $h(x)$ on Γ such that $|f| \leq h$ on Γ. We shall use the fact that the exterior norm satisfies the triangle inequality: $]\|f_1 + f_2\|_{p,\Gamma} \leq]\|f_1\|_{p,\Gamma} +]\|f_2\|_{p,\Gamma}$. Also, we shall use the fact that if $]\|f_n\|_{p,\Gamma} \to 0$ as $n \to \infty$, there is a subsequence $\{n'\}$ for which $f_{n'}(x) \to 0$ as $n' \to \infty$ (Γ-a.s.).

$B(\Gamma)$ denotes the set of bounded Borel functions on Γ with the norm

$$\|f\|_{B(\Gamma)} = \sup_{x \in \Gamma} |f(x)|.$$

$C(\Gamma)$ denotes the set of continuous (possibly, unbounded) functions on Γ.

f *is a smooth function* means that f is infinitely differentiable. We say that f has compact support in a region D if it vanishes outside some compact subset of D.

$C_0^\infty(D)$ denotes the set of all smooth functions with compact support in the region D.

$$f_{(y)}(t,x) \equiv \frac{1}{|y|} \sum_i y^i f_{x^i}(t,x) \quad \text{if } |y| \neq 0; \qquad f_{(0)}(t,x) \equiv 0;$$

$$f_{(y_1)(y_2)}(t,x) \equiv \frac{1}{|y_1| \cdot |y_2|} \sum_{i,j} y_1^i y_2^j f_{x^i x^j}(t,x) \quad \text{if } |y_1| \cdot |y_2| \neq 0;$$

$$f_{(y_1)(y_2)}(t,x) \equiv 0 \quad \text{if } |y_1| \cdot |y_2| = 0.$$

We introduce $f_{(y_1) \cdots (y_n)}$. These elements are derivatives of $f(t,x)$ along spacial directions. The time derivative is always expressed as $(\partial/\partial t) f(t,x)$.

$C^2(\bar{D})$ denotes the set of functions $u(x)$ twice continuously differentiable in $\bar{D}$ (i.e., twice continuously differentiable in D and such that $u(x)$ as well as all first and second derivatives of $u(x)$ have extensions continuous in $\bar{D}$).

$C^{1,2}(\bar{Q})$ denotes the set of functions $u(t,x)$ twice continuously differentiable in x and once continuously differentiable in t in $\bar{Q}$.

Let D be a bounded region in E_d, and let $u(x)$ be a function in $\bar{D}$. We write $u \in W^2(D)$ if there exists a sequence of functions $u^n \in C^2(\bar{D})$ such that

$$\|u - u^n\|_{B(\bar{D})} \to 0, \qquad \|u^n - u^m\|_{W^2(D)} \to 0 \tag{1}$$

as $n, m \to \infty$, where

$$\|f\|_{W^2(D)} \equiv \sum_{i,j=1}^{d} \|f_{x^i x^j}\|_{d,D} + \sum_{i=1}^{d} \|f_{x^i}\|_{d,D} + \|f\|_{B(\bar{D})}.$$

Under the first condition of (1) and due to the continuity property of u^n, the functions in $W^2(D)$ are continuous in $\bar{D}$. The second condition in (1) implies that the sequences $u^n_{x^i}$, $u^n_{x^ix^j}$ are fundamental in $\mathscr{L}_d(D)$. Hence there exist (Borel) functions $u_i, u_{ij} \in \mathscr{L}_d(D)$, to which $u^n_{x^i}$, $u^n_{x^ix^j}$ converge in $\mathscr{L}_d(D)$. These sequences $u^n_{x^i}$, $u^n_{x^ix^j}$ converge weakly as well to the functions given above. In particular, assuming $\varphi \in C_0^\infty(D)$, and integrating by parts, we obtain

$$\int_D \varphi u^n_{x^i}\,dx = -\int_D \varphi_{x^i} u^n\,dx,$$

Letting $n \to \infty$, we obtain

$$\int_D \varphi u_i\,dx = -\int_D \varphi_{x^i} u\,dx. \tag{2}$$

1. Definition. Let $D \subset E_d$, let v and h be Borel functions locally summable in D, and let $l_1, \ldots, l_n \in E_d$. The function h is said to be a generalized derivative (in the region D) of the function v of order n in the $l_1, \ldots, l_n$ directions and this function h is denoted by $v_{(l_1)\cdots(l_n)}$ if for each $\varphi \in C_0^\infty(D)$

$$\int_D \varphi(x)h(x)\,dx = (-1)^n \int_D v(x)\varphi_{(l_1)\cdots(l_n)}\,dx.$$

In the case where the l_i direction coincides with the direction of the r_ith coordinate vector, the above function is expressed in terms of $v_{x^{r_1}\cdots x^{r_n}} = v_{(l_1)\cdots(l_n)}$.

The properties of a generalized derivative are well known (see [57, 71, 72]. We shall list below only those properties which we use frequently, without proving them. Note first that a generalized derivative can be defined uniquely almost everywhere.

Equation (2) shows that $u_i = u_{x^i}$ in the sense of Definition 1. Similarly, $u_{ij} = u_{x^ix^j}$. Therefore, the functions $u \in W^2(D)$ have generalized derivatives up to and including the second order. Furthermore, these derivatives belong to $\mathscr{L}_d(D)$. We assume that the values of first and second derivatives of each function $u \in W^2(D)$ are fixed at each point. By construction, for the sequence u^n entering (1),

$$\|u^n_{x^i} - u_{x^i}\|_{d,D} \to 0, \qquad \|u^n_{x^ix^j} - u_{x^ix^j}\|_{d,D} \to 0.$$

The set of functions $W^2(D)$ introduced resembles the well-known Sobolev space $W^2_d(D)$ (see [46, 71, 72]). If the boundary of the region D is sufficiently regular, for example, it is once continuously differentiable; Sobolev's theorem on imbedding (see [46, 47]) shows that, in fact, $W^2(D) = W^2_d(D)$. In this case $u \in W^2(D)$ if and only if u is continuous in $\bar{D}$, has generalized derivatives up to and including the second order, and, furthermore, these derivatives are summable in D to the power d.

It is seen that if the function u is once continuously differentiable in D, its ordinary first derivatives coincide with its first generalized derivatives (almost everywhere). It turns out (a corollory of Fubini's theorem) that, for example, a generalized derivative u_{x^1} exists in the region D if for almost all $(x_0^2,\dots,x_0^d)$ the function $u(x^1,x_0^2,\dots,x_0^d)$ is absolutely continuous in x^1 on $\{x^1:(x^1,x_0^2,\dots,x_0^d)\in D\}$ and its usual derivative with respect to x^1 is locally summable in D. The converse is also true. However, we ought to replace then the function u by a function equivalent with respect to Lebesgue measure. It is well known that if for almost all $(x_0^{i+1},\dots,x_0^d)$ the function $u(x^1,\dots,x^i,x_0^{i+1},\dots,x_0^d)$ has a generalized derivative on $\{(x^1,\dots,x^i):(x^1,\dots,x^i,x_0^{i+1},\dots,x_0^d)\in D\}$ and, in addition, this derivative is locally summable in D, u will have a generalized derivative in D.

Using the notion of weak convergence, we can easily prove that if the functions φ, v^n $(n=0,1,2,\dots)$ are uniformly bounded in D, $v^n\to v^0$ (D-a.s.), for some $l_1,\dots,l_k$ for $n\geq 1$ the generalized derivatives $v^n_{(l_1)\dots(l_k)}$ exist, and $|v^n_{(l_1)\dots(l_k)}|\leq\varphi$ (D-a.s.), the generalized derivative $v^0_{(l_1)\dots(l_k)}$ also exists, $|v^0_{(l_1)\dots(l_k)}|\leq\varphi$ (D-a.s.), and

$$v^n_{(l_1)\dots(l_k)}\to v^0_{(l_1)\dots(l_k)}$$

weakly in $\mathscr{L}_2$ in any bounded subset of the region D.

In many cases, one needs to "mollify" functions to be smooth. We shall do this in a standard manner. Let $\zeta(x)$, $\zeta_1(t)$, $\zeta(t,x)\equiv\zeta_1(t)\zeta(x)$ be nonnegative, infinitely differentiable functions of the arguments $x\in E_d$, $t\in E_1$, equal to zero for $|x|>1$, $|t|>1$ and such that

$$\int_{E_d}\zeta(x)\,dx=1,\qquad\int_{-\infty}^{\infty}dt\int_{E_d}\zeta(t,x)\,dx=1.$$

For $\varepsilon\neq 0$ and the functions $u(x)$, $u(t,x)$ locally summable in E_d, $E_1\times E_d$, let

$$u^{(\varepsilon)}(x)=\varepsilon^{-d}\zeta\left(\frac{x}{\varepsilon}\right)*u(x)\quad\text{(convolution with respect to } x),$$

$$u^{(0,\varepsilon)}(t,x)=\varepsilon^{-d}\zeta\left(\frac{x}{\varepsilon}\right)*u(t,x)\quad\text{(convolution with respect to } x),$$

$$u^{(\varepsilon)}(t,x)=\varepsilon^{-(d+1)}\zeta\left(\frac{t}{\varepsilon},\frac{x}{\varepsilon}\right)*u(t,x)\quad\text{(convolution with respect to } (t,x)).$$

The functions $u^{(\varepsilon)}(x)$, $u^{(0,\varepsilon)}(t,x)$,$u^{(\varepsilon)}(t,x)$ are said to be mean functions of the functions $u(x)$, $u(t,x)$. It is a well-known fact (see [10, 71]) that $u^{(\varepsilon)}\to u$ as $\varepsilon\to 0$:

a. at each Lebesgue point of the function u, therefore almost everywhere;
b. at each continuity point of the function u; uniformly in each bounded region, if u is continuous;
c. in the norm $\mathscr{L}_p(D)$ if $u\in\mathscr{L}_p(D)$ and in computing the convolution of $u^{(\varepsilon)}$ the function u is assumed to be equal to zero outside D.

Furthermore, $u^{(\varepsilon)}$ is infinitely differentiable. If a generalized derivative $u_{(l)}$ exists in E_d, then $[u_{(l)}]^{(\varepsilon)}=[u^{(\varepsilon)}]_{(l)}$. Finally, for $p\geq 1$

$$\|u^{(\varepsilon)}\|_{p,E_d}\leq\|u\|_{p,E_d},\qquad\|u^{(\varepsilon)}\|_{B(E_d)}\leq\|u\|_{B(E_d)}.$$

Considering the functions $u^{(\varepsilon)}$, we prove that the generalized derivative u_{x^1} of the function $u(x)$ continuous in D does not exceed a constant N_1 almost everywhere if and only if the function $u(x)$ satisfies in D the Lipschitz condition with respect to x^1 having this constant, that is, if for any points $x_1, x_2 \in D$ such that an interval with the end points x_1, x_2 lies in D and $x_1^i = x_2^i$ $(i = 2, \ldots, d)$, the inequality $|u(x_1) - u(x_2)| \leq N_1|x_1 - x_2|$ can be satisfied. It turns out that if a bounded function σ has a bounded generalized derivative, σ^2 has as well a generalized derivative, and one can use usual formulas to find this generalized derivative.

In addition to the space $W^2(D)$ we need spaces $\bar{W}^2(D)$, $W^{1,2}(Q)$, and $\bar{W}^{1,2}(Q)$, which are introduced for bounded regions D, Q in a way similar to the way $W^2(D)$ was, starting from sets of functions $C^2(\bar{D})$, $C^{1,2}(\bar{Q})$, and $C^{1,2}(\bar{Q})$, respectively, and using the norms

$$\|f\|_{\bar{W}^2(D)} = \|f\|_{W^2(D)} + \sum_{i=1}^{d} \|f_{x^i}\|_{2d,D},$$

$$\|f\|_{W^{1,2}(Q)} = \left\|\frac{\partial}{\partial t} f\right\|_{d+1,Q} + \sum_{i,j=1}^{d} \|f_{x^i x^j}\|_{d+1,Q} + \sum_{i=1}^{d} \|f_{x^i}\|_{d+1,Q} + \|f\|_{B(\bar{Q})},$$

$$\|f\|_{\bar{W}^{1,2}(Q)} = \|f\|_{W^{1,2}(Q)} + \sum_{i=1}^{d} \|f_{x^i}\|_{2(d+1),Q}.$$

For proving existence of generalized derivatives of a payoff function another notion proves to be useful.

2. Definition. Let a function $u(x)$ be given, and let it be locally summable in a region D. Let $v(\Gamma)$ be a function of a set Γ which is definite, σ-additive, and finite on the σ-algebra of Borel subsets of each bounded region $D' \subset \bar{D}' \subset D$. We say that the set function v on D is a generalized derivative of the function u in the $l_1, \ldots, l_k$ directions, and we write

$$v(dx) = u_{(l_1)\ldots(l_k)}(x)\,(dx), \tag{3}$$

if for each function $\varphi \in C_0^\infty(D)$,

$$\int_D u\varphi_{(l_1)\ldots(l_k)}\,dx = (-1)^k \int_D \varphi v(dx). \tag{4}$$

The generalized derivative $(\partial/\partial t)u(t,x)(dt\,dx)$ for the function $u(t,x)$ locally summable in the region Q can be found in a similar way.

The definitions given above immediately imply the following properties. It is easily seen that there exists only one function $v(dx)$ satisfying (4) for all $\varphi \in C_0^\infty(D)$. If the function $u_{(l_1)\ldots(l_k)}(x)$ exists, which is a generalized derivative of u in the $l_1, \ldots, l_k$ directions in the sense of Definition 1, assuming that $v(dx) = u_{(l_1)\ldots(l_k)}(x)\,dx$, we obtain in an obvious manner a set function

v, being the generalized derivative of u in the $l_1, \ldots, l_k$ directions in the sense of Definition 2.

Conversely, if the set function v in Definition 2 is absolutely continuous with respect to Lebesgue measure, its Radon–Nikodym derivative will satisfy Definition 1 in conjunction with (4). Therefore, this Radon–Nikodym derivative is the generalized derivative $u_{(l_1)\cdots(l_k)}(x)$. This fact justifies the notation of (3). In the case where the direction l_i coincides with the direction of the r_ith coordinate vector, we shall write

$$u_{(l_1)\cdots(l_k)}(x)(dx) = u_{x^{r_1}\cdots x^{r_k}}(x)(dx).$$

Using the uniqueness property of a generalized derivative, we easily prove that if the derivatives $u_{(l_1)\cdots(l_k)}(x)(dx)$ for some k exist for all $l_1, \ldots, l_k$, then

$$u_{(l_1)\cdots(l_k)}(x)(dx) = \frac{1}{|l_1|\cdots|l_k|}\sum_{r_1,\ldots,r_k} u_{x^{r_1}\cdots x^{r_k}}(x)(dx)l_1^{r_1}\cdots l_k^{r_k}$$

for $|l_1|\cdots|l_k| \neq 0$. Further, if the derivatives $u_{(l)(l)}(x)(dx)$ exist for all l, all the derivatives $u_{(l_1)(l_2)}(x)(dx)$ exist as well. In this case, if $|l_1|\cdot|l_2| \neq 0$, then

$$\begin{aligned} u_{(l_1)(l_2)}(x)(dx) = \frac{1}{4|l_1|\cdot|l_2|}&[(l_1 + l_2)^2 u_{(l_1+l_2)(l_1+l_2)}(x)(dx) \\ &- (l_1 - l_2)^2 u_{(l_1-l_2)(l_1-l_2)}(x)(dx)]. \end{aligned}$$

In fact, using Definition 2 we easily prove that the right side of this formula satisfies Definition 2 for $k = 2$.

Theorem V of [67, Chapter 1, §1] constitutes the main tool enabling us to prove the existence of $u_{(l_1)\cdots(l_k)}(x)(dx)$. In accord with this theorem from [67], the nonnegative generalized function is a measure. Regarding

$$\int u\varphi_{(l_1)\cdots(l_k)}(dx) - (-1)^k \int \varphi v(dx) \tag{5}$$

as a generalized function, we have the following.

3. Lemma. *Let $u(x)$, $v(\Gamma)$ be the same as those in the first two propositions of Definition 2. For each nonnegative $\varphi \in C_0^\infty(D)$ let the expression (5) be nonnegative. Then there exists a generalized derivative $u_{(l_1)\cdots(l_k)}$ in the sense of Definition 2. In this case, inside D*

$$(-1)^k u_{(l_1)\cdots(l_k)}(x)(dx) \geq (-1)^k v(dx),$$

that is, for all bounded Borel $\Gamma \subset \bar{\Gamma} \subset D$

$$(-1)^k u_{(l_1)\cdots(l_k)}(\Gamma) \geq (-1)^k v(\Gamma).$$

To conclude the discussion in this section we summarize more or less conventional agreements and notation.

$(\mathbf{w}_t, \mathscr{F}_t)$ is a Wiener process (see Appendix 1).

$\mathscr{F}_\tau$ is the σ-algebra consisting of all those sets A for which the set $A \cap \{\tau \leq t\} \in \mathscr{F}_t$ for all t.

$\mathfrak{M}(t)$ denotes the set of all Markov (with respect to $\{\mathscr{F}_t\}$) times τ not exceeding t (see Appendix 1).

$C([0,T],E_d)$ denotes a Banach space of continuous functions on $[0,T]$ with range in E_d, $\mathscr{N}_t$ the smallest σ-algebra of the subsets of $C([0,T],E_d)$ which contains all sets of the form

$$\{x_{[0,T]} \in C([0,T],E_d): x_s \in \Gamma\},$$

where $s \leq t$, Γ denotes a Borel subset of E_d.

l.i.m. reads the mean square limit.

ess sup reads the essential upper bound (with respect to the measure which is implied).

$$\inf \varnothing = \infty, \qquad f(x_\tau) \equiv f(x_\tau)\chi_{\tau < \infty}.$$

When we speak about measurable functions (sets), we mean, as a rule, Borel functions (sets). The words "nonnegative," "nonpositive," "it does not increase," "it does not decrease," mean the same as the words "positive," "negative," "it decreases," "it increases," respectively.

Finally,

$$\Delta = \sum_{i=1}^{d} \frac{\partial^2}{\partial (x^i)^2}$$

denotes the Laplace operator. The operators L^α, $F[u]$, $F_1[u]$, used in Chapters 4–6 are defined in the introductory section in Chapter 4.

2. Estimates of the Distribution of a Stochastic Integral in a Bounded Region

Let A be a set of pairs (σ,b), where σ is a matrix of dimension $d \times d_1$ and b is a d-dimensional vector. We assume that a random process $(\sigma_t,b_t) \in A$ for all (ω,t), and that the process

$$x_t = x_0 + \int_0^t \sigma_s \, d\mathrm{w}_s + \int_0^t b_s \, ds$$

is defined.

We shall see further that in stochastic control, estimates of the form

$$\mathsf{M} \int_0^{\tau_D} |f(t,x_t)| \, dt \leq N \|f\|_{p,Q} \tag{1}$$

play an essential role, in (1) f is an arbitrary Borel function, τ_D is the first exit time of x_t from the region D, and $Q = (0,\infty) \times D$. A crucial fact here is that the constant N does not depend on a specified process (σ_t,b_t), but is given instead by the set A. In this section, our objective is to deduce a few versions of the estimate (1).

We assume that D is a bounded region in E_d, x_0 is a fixed point of D, an integer $d_1 \geq d$, $(\mathbf{w}_t,\mathscr{F}_t)$ is a d_1-dimensional Wiener process, $\sigma_t(\omega)$ is a matrix of dimension $d \times d_1$, $b_t(\omega)$ is a d-dimensional vector, and $c_t(\omega)$, $r_t(\omega)$ are nonnegative numbers. Assume in addition that σ_t, b_t, c_t, r_t are progressively measurable with respect to $\{\mathscr{F}_t\}$ and that they are bounded functions of (t,ω). Let $a_t = \frac{1}{2}\sigma_t\sigma_t^*$.

Next, let p be a fixed number, $p \geq d$, and let

$$y_{s,t} = \int_s^t r_u\,du, \qquad \varphi_{s,t} = \int_s^t c_u\,du, \qquad \psi_t = c_t^{1-[(d+1)/(p+1)]}(r_t \det a_t)^{1/(p+1)}.$$

One should keep in mind that for $p = d$ the expression $c_t^{(p-d)/(p+1)}$ is equal to unity even if $c_t = 0$; therefore $\psi_t = (r_t \det a_t)^{1/(d+1)}$ for $p = d$.

1. Definition. A nonnegative function $F(c,a)$ defined on the set of all nonnegative numbers c as well as nonnegative definite symmetric matrices a of dimension $d \times d$ is said to be regular if for each $\varepsilon > 0$ there is a constant $k(\varepsilon)$ such that for all c, a and unit vectors λ

$$F(c,a) \leq \varepsilon \operatorname{tr} a + k(\varepsilon)[c + (a\lambda,\lambda)].$$

2. Theorem. *Assume that $|b_t| \leq F(c_t,a_t)$ for all (t,ω) for some regular function $F(c,a)$. There exist constants N_1,N_2 depending only on d, the function $F(c,a)$ and the diameter of the region D, and such that for all $s \geq 0$, Borel $f(t,x)$ and $g(x)$, on a set $\{\tau_D \geq s\}$, almost surely*

$$\mathsf{M}\left\{\int_s^{\tau_D} e^{-\varphi_{s,t}}\psi_t|f(y_{s,t},x_t)|\,dt\,\big|\,\mathscr{F}_s\right\} \leq N_1\|f\|_{p+1,Q}, \tag{2}$$

$$\mathsf{M}\left\{\int_s^{\tau_D} e^{-\varphi_{s,t}}c_t^{1-(d/p)}(\det a_t)^{1/p}|g(x_t)|\,dt\,\big|\,\mathscr{F}_s\right\} \leq N_s\|g\|_{p,D}. \tag{3}$$

Before proving our theorem, we discuss the assertions of the theorem and give examples of regular functions. Note that the left sides of the inequalities (2) and (3) make sense because of the measurability requirements.

It is seen that the function $F(c,a) \equiv c$ is regular. Next, in conjunction with Young's inequality,

$$xy = \left(\frac{x}{\varepsilon}\right)(\varepsilon y) \leq \frac{\varepsilon^q y^q}{q} + \frac{x^p}{\varepsilon^p p},$$

if $x,y \geq 0$, $p^{-1} + q^{-1} = 1$. Hence for $\alpha \in (0,1)$, $\varepsilon \in (0,1)$

$$c^\alpha(\operatorname{tr} a)^{1-\alpha} \leq \varepsilon(1-\alpha)\operatorname{tr} a + \alpha\varepsilon^{1-(1/\alpha)}c \leq \varepsilon \operatorname{tr} a + \varepsilon^{1-(1/\alpha)}c.$$

Therefore, $c^\alpha(\operatorname{tr} a)^{1-\alpha}$ is a regular function for $\alpha \in (0,1)$.

We show that the function $(\det a)^{1/d}$ not depending on c is regular. Let $\mu_1 \leq \mu_2 \leq \cdots \leq \mu_d$ be eigenvalues of a matrix a. We know that $\mu_1 \leq (a\lambda,\lambda)$ if $|\lambda| = 1$. Further, $\det a = \mu_1\mu_2 \cdots \mu_d$, $\operatorname{tr} a = \mu_1 + \mu_2 + \cdots + \mu_d$. From this, in conjunction with the Young's inequality, we have

$$(\det a)^{1/d} = \mu_1^{1/d}(\mu_2 \cdots \mu_d)^{1/d} \leq \frac{1}{d}\varepsilon^{-(d-1)}\mu_1 + \frac{d-1}{d}(\varepsilon\mu_2 \cdots \varepsilon\mu_d)^{1/(d-1)}$$

$$\leq \frac{1}{d}\varepsilon^{-(d-1)}(a\lambda,\lambda) + \frac{1}{d}\varepsilon(\mu_2 + \cdots + \mu_d)$$

$$\leq (\varepsilon \operatorname{tr} a + \varepsilon^{-(d-1)}(a\lambda,\lambda).$$

Using the regular functions given above, we can construct many other regular functions, noting that a linear combination with positive coefficients of regular functions is a regular function.

The function tr a is the limit of regular functions $c^{\alpha}(\operatorname{tr} a)^{1-\alpha}$ as $\alpha \downarrow 0$. However, for $d \geq 2$ the function $\operatorname{tr} a$ is not regular. To prove this, we suggest the reader should consider

3. Exercise

For $p = d$, $c_t \equiv 0$, $s = 0$, $g \equiv 1$ it follows from (3) that

$$\mathsf{M} \int_0^{\tau_D} (\det a_t)^{1/d}\, dt \leq N_2\, (\operatorname{meas} D)^{1/d}. \tag{4}$$

From the statement of Theorem 2 we take $D = S_R$, $F(c,a) = K \operatorname{tr} a$, with $K > R^{-1}$.

It is required to prove that for $d \geq 2$ there exists no constant N_2 depending only on d, K, R, for which (4) can be satisfied.

This exercise illustrates the fact that the requirement $|b_t| \leq F(c_t,a_t)$, where F is a regular function, is essential. In contrast to this requirement, we can weaken considerably the assumption about boundedness of σ, b, c, r. For example, considering instead of the process x_t, $y_{r,t}$ the processes

$$\bar{x}_t = x_0 + \int_0^t \chi_{u<\tau_D}\sigma_u\, dw_u + \int_0^t \chi_{u<\tau_D} b_u\, du, \qquad \bar{y}_{s,t} = \int_s^t \chi_{u<\tau_D} r_u\, du,$$

where τ_D is the time of first departure of x_t from D, and noting that $x_t = \bar{x}_t$, $y_{s,t} = \bar{y}_{s,t}$ for $t < \tau_D$, we immediately establish the assertion of Theorem 2 in the case where $\chi_{t<\tau_D}\sigma_t$, $\chi_{t<\tau_D}b_t$, $\chi_{t<\tau_D}c_t$, $\chi_{t<\tau_D}r_t$ are bounded functions of (t,ω).

We think that the case where $s = 0$, $r_t \equiv 1$, $p = d$ is the most important particular case of Theorem 2. It is easily seen, in fact, that the proof of our theorem follows generally from the particular case indicated. The formal proof is rather difficult, however. It should be noted that according to our approach to the proof of the theorem, assuming $s \neq 0$, $r_t \not\equiv 1$ makes the proving of estimates for $s = 0$, $r_t \equiv 1$ essentially easier. In the future, it will be convenient to use the following weakened version of the assertions of Theorem 2.

4. Theorem. *Let τ be a Markov time (with respect to $\{\mathscr{F}_t\}$), not exceeding τ_D. Also, let there exist constants K, $\delta > 0$ such that for all $t < \tau(\omega)$, $\lambda \in E_d$*

$$|b_t(\omega)| \leq K, \qquad \sum_{i,j=1}^{d} a_t^{ij}(\omega)\lambda^i\lambda^j \geq \delta|\lambda|^2.$$

Then there exists a constant N depending only on d, K, δ, and the diameter of the region D such that for all $s \geq 0$ and Borel $f(t,x)$ and $g(x)$ on the set $\{s \leq \tau\}$, almost surely

$$\mathsf{M}\left\{\int_s^\tau |f(t,x_t)|\,dt \,\middle|\, \mathscr{F}_s\right\} \leq N\|f\|_{d+1,Q},$$

$$\mathsf{M}\left\{\int_s^\tau |g(x_t)|\,dt \,\middle|\, \mathscr{F}_s\right\} \leq N\|g\|_{d,D}.$$

This theorem follows immediately from Theorem 2 for $r_t \equiv 1$, $c_t \equiv 0$, $p = d$. In fact, we have

$$x_{t\wedge\tau} = x_0 + \int_0^t \chi_{u<\tau}\sigma_u\,d\mathbf{w}_u + \int_0^t \chi_{u<\tau}b_u\,du,$$

$$\mathsf{M}\left\{\int_s^\tau |f(t,x_t)|\,dt \,\middle|\, \mathscr{F}_s\right\} \leq \delta^{-d/(d+1)}\mathsf{M}\left\{\int_s^{\tau_D} e^{-\varphi_{s,t}}\chi_{t<\tau}\psi_t|f(s+y_{s,t},x_t)|\,dt \,\middle|\, \mathscr{F}_s\right\}$$

since $e^{-\varphi_{s,t}}\psi_t = (\det a_t)^{1/(d+1)}$ and $\det a_t$, which is equal to the product of eigenvalues of the matrix a_t for $t \leq \tau$, is not smaller than δ^d. Furthermore, $|\chi_{t<\tau}b_t| \leq K\delta^{-1}(\det \chi_{t<\tau}a_t)^{1/d}$, the function $F(c,a) = K\delta^{-1}(\det a)^{1/d}$ is regular and, in addition, $\{s \leq \tau_D\} \supset \{s \leq \tau\}$.

Next, in order to prove Theorem 2, we need three lemmas.

5. Lemma. *Let $|b_t| \leq F(c_t,a_t)$ for all (t,ω) for some regular function $F(c,a)$. There exists a constant N depending only on the function $F(c,a)$ and the diameter of the region D such that on the set $\{\tau_D \geq s\}$ almost surely*

$$\mathsf{M}\left\{\int_s^{\tau_D} e^{-\varphi_{s,t}}|b_t|\,dt \,\middle|\, \mathscr{F}_s\right\} \leq \mathsf{M}\left\{\int_s^{\tau_D} e^{-\varphi_{s,t}}F(c_t,a_t)\,dt \,\middle|\, \mathscr{F}_s\right\} \leq N.$$

Proof. We can assume without loss of generality that $x_0 = 0$. We denote by R the diameter of the region D and set $u(x) = \beta - \operatorname{ch}\alpha|x|$ for $\alpha > 0$, $\beta > \operatorname{ch}(\alpha R)$. We note that $u(x)$ is twice continuously differentiable and $u(x) \geq 0$ for $x \in D$. Applying Ito's formula to $e^{-\varphi_{s,t}}u(x_t)$, we have for $t \geq s$ on the set $\{\tau_D \geq s\}$ that

$$u(x_s) = \mathsf{M}\left\{\int_s^{t\wedge\tau_D} e^{-\varphi_{s,r}}[c_r u(x_r) - L^{\sigma_r,b_r}u(x_r)]\,dr + e^{-\varphi_{s,t\wedge\tau_D}}u(x_{t\wedge\tau_D}) \,\middle|\, \mathscr{F}_s\right\}$$

$$\geq \mathsf{M}\left\{\int_s^{t\wedge\tau_D} e^{-\varphi_{s,r}}[c_r u(x_r) - L^{\sigma_r,b_r}u(x_r)]\,dr \,\middle|\, \mathscr{F}_s\right\} \qquad \text{(a.s.)}$$

Assume that for all $x \in D$, $r \geq 0$

$$c_r u(x) - L^{\sigma_r, b_r} u(x) \geq F(c_r, a_r). \tag{5}$$

Then

$$\beta \geq u(x_s) \geq \mathsf{M}\left\{\int_s^{t \wedge \tau_u} e^{-\varphi_{s,r}} F(c_r, a_r)\, dr \,\middle|\, \mathscr{F}_s\right\},$$

which proves the assertion of the lemma as $t \to \infty$, with the aid of Fatou's lemma.

Therefore, it remains only to choose constants α, β such that (5) is satisfied, assuming obviously that $x \neq 0$. For simplicity of notation, we shall not write the subscript r in c_r, σ_r, a_r, b_r. In addition, let $\lambda = x/|x|$, $\rho = |x|$. A simple computation shows that

$$\begin{aligned} I &\equiv (1 + \alpha \operatorname{sh} \alpha |x|)^{-1}[cu(x) - L^{\sigma,b}u(x) - F(c,a)] \\ &= (1 + \alpha \operatorname{sh} \alpha\rho)^{-1}\{c(\beta - \operatorname{ch} \alpha\rho) + \alpha \operatorname{sh} \alpha\rho(b,\lambda) \\ &\quad + \alpha^2 \operatorname{ch} \alpha\rho(a\lambda,\lambda) + \frac{\alpha}{\rho} \operatorname{sh} \alpha\rho[\operatorname{tr} a - (a\lambda,\lambda)] - F(c,a)\} \\ &\geq c\frac{\beta - \operatorname{ch} \alpha\rho}{1 + \alpha \operatorname{sh} \alpha\beta} + (a\lambda,\lambda)\frac{\alpha^2 \operatorname{ch} \alpha\rho}{1 + \alpha \operatorname{sh} \alpha\rho} \\ &\quad + \frac{\alpha \operatorname{sh} \alpha\rho}{1 + \alpha \operatorname{sh} \alpha\rho}\frac{1}{\rho}[\operatorname{tr} a - (a\lambda,\lambda)] - F(c,a). \end{aligned} \tag{6}$$

We note that $\operatorname{ch} \alpha\rho \geq 1, \operatorname{ch} \alpha\rho \geq \operatorname{sh} \alpha\rho$, $\alpha \operatorname{sh} \alpha\rho \geq \alpha^2\rho$ and for $x \in D$ the number $\rho \leq R$. Hence

$$\frac{\beta - \operatorname{ch} \alpha\rho}{1 + \alpha \operatorname{sh} \alpha\rho} \geq \frac{\beta - \operatorname{ch} \alpha R}{1 + \alpha \operatorname{sh} \alpha R}, \qquad \frac{\alpha^2 \operatorname{ch} \alpha\rho}{1 + \alpha \operatorname{sh} \alpha\rho} \geq \frac{\alpha^2 \operatorname{ch} \alpha\rho}{\operatorname{ch} \alpha\rho + \alpha \operatorname{ch} \alpha\rho} = \frac{\alpha^2}{1 + \alpha},$$

$$\frac{1}{\rho}\frac{\alpha \operatorname{sh} \alpha\rho}{1 + \alpha \operatorname{sh} \alpha\rho} \geq \frac{1}{\rho}\frac{\alpha^2\rho}{1 + \alpha^2\rho} \geq \frac{\alpha^2}{1 + \alpha^2 R}.$$

Therefore, it follows from (6) that

$$I \geq c\frac{\beta - \operatorname{ch} \alpha R}{1 + \alpha \operatorname{sh} \alpha R} + (a\lambda,\lambda)\frac{\alpha^2}{1 + \alpha} + \frac{\alpha^2}{1 + \alpha^2 R}[\operatorname{tr} a - (a\lambda,\lambda)] - F(c,a).$$

We recall that $F(c,a)$ is a regular function. Also, we fix some $\varepsilon < 1/R$ and choose α large enough that $\alpha^2/(1 + \alpha^2 R) > \varepsilon$, $\alpha^2/(1 + \alpha) \geq k(\varepsilon) + \varepsilon$. Next, we take a number β so large that

$$\frac{\beta - \operatorname{ch} \alpha R}{1 + \alpha \operatorname{sh} \alpha R} \geq k(\varepsilon).$$

Then $I \geq k(\varepsilon)[c + (a\lambda,\lambda)] + \varepsilon \operatorname{tr} a - F(c,a) \geq 0$, thus proving the lemma. □

6. Corollary. *Let $G(c,a)$ be a regular function. There exists a constant N depending only on $F(c,a)$, $G(c,a)$ and the diameter of the region D such that*

$$\mathsf{M}\left\{\int_s^{\tau_D} e^{-\varphi_{s,t}}G(c_t,a_t)\,dt\,\Big|\,\mathscr{F}_s\right\} \leq N \qquad (\{\tau_D \geq s\}\text{-a.s.}).$$

In fact, let $F_1(c,a) = F(c,a) + G(c,a)$. Then $|b_t| \leq F_1(c_t,a_t)$, $G(c_t,a_t) \leq F_1(c_t,a_t)$, and the assertion of our lemma is proved for $F_1(c,a)$.

7. Lemma. *Let $R > 0, h(t,x) \geq 0, h \in \mathscr{L}_{d+1}(C_R), h(t,x) = 0$ for $t \leq 0, h(t,x) = 0$ for $|x| \geq R$. Then on $(-\infty,\infty) \times E_d$ there exists a bounded function $z(t,x) \leq 0$ equal to zero for $t < 0$ and such that for all sufficiently small $\varepsilon > 0$ and non-negative definite symmetric matrices $a = (a^{ij})$ on a cylinder C_R.*

$$N(d)(\det a)^{1/(d+1)}h^{(\varepsilon)} \leq -\frac{\partial}{\partial t}z^{(\varepsilon)} + \sum_{i,j=1}^{d} a^{ij}z^{(\varepsilon)}_{x^ix^j},$$

where $N(d) > 0$. Furthermore, if the vector b and the number c are such that $|b| \leq (R/2)c$, then on the same set $\sum_{i=1}^d b^i z^{(\varepsilon)}_{x^i} \geq cz^{(\varepsilon)}$, if ε is sufficiently small. Finally, for all $t \geq 0$, $x \in E_d$

$$|z(t,x)|^{d+1} \leq N(d,R)\int_{S_R}\int_0^t h^{d+1}(s,y)\,ds\,dy$$

This lemma is proved in [42] by geometric arguments.

8. Lemma. *Let $|b_t| \leq F(c_t,a_t)$ for all (t,ω) for a regular function $F(c,a)$. There exists a constant N depending only on d, $F(c,a)$, and the diameter of D, and such that for all $s \geq 0$, $f(t,x)$ on a set $\{\tau_D \geq s\}$, almost surely*

$$\mathsf{M}\left\{\int_s^{\tau_D}\exp\left\{-\int_s^t c_u\,du\right\}(r_t\det a_t)^{1/(d+1)}\left|f\left(\int_s^t r_u\,du,x_t\right)\right|dt\,\Big|\,\mathscr{F}_s\right\} \leq N\|f\|_{d+1,Q}.$$

In other words, the inequality (2) holds for $p = d$.

PROOF. Let us use the notation introduced above:

$$\varphi_{s,t} = \int_s^t c_u\,du, \qquad \psi_t = (r_t\det a_t)^{1/(d+1)}, \qquad y_{s,t} = \int_s^t r_u\,du.$$

We denote by R the diameter of D and we consider without loss of generality that $x_0 = 0$. In this case $D \subset S_R$. Also, we assume that τ_R is the first exit time of x_t from S_R. It is seen that $\tau_R \geq \tau_D$.

Suppose that we have proved the inequality

$$\mathsf{M}\left\{\int_s^{\tau_R} e^{-\varphi_{s,t}}\psi_t|f(y_{s,t},x_t)|\,dt\,\Big|\,\mathscr{F}_s\right\} \leq N\|f\|_{d+1,C_R} \tag{7}$$

($\{\tau_R \geq s\}$-a.s.) for arbitrary s, f, where $N = N(d,F,R)$. Furthermore, taking in (7) the function f equal to zero for $x \notin D$, we obtain

$$\mathsf{M}\left\{\int_s^{\tau_D} e^{-\varphi_{s,t}}\psi_t|f(y_{s,t},x_t)|\,dt\,|\mathscr{F}_s\right\} \le \mathsf{M}\left\{\int_s^{\tau_R} e^{-\varphi_{s,t}}\psi_t|f(y_{s,t},x_t)|\,dt\,|\mathscr{F}_s\right\}$$
$$\le N\|f\|_{d+1,C_R} = N\|f\|_{d+1,Q}$$

($\{\tau_R \ge s\}$-a.s.) and, *a fortiriori*, ($\{\tau_D \ge s\}$-a.s.).

It suffices therefore to prove (7). Usual reasoning (using, for example, the results given in [54, Chapter 1, §2]) shows that it suffices to prove (7) only for bounded continuous nonnegative $f(t,x)$. Noting in addition that by Fatou's lemma, for such a function

$$\mathsf{M}\left\{\int_s^{\tau_R} e^{-\varphi_{s,t}}\psi_t f(y_{s,t},x_t)\,dt\,|\mathscr{F}_s\right\} \le \lim_{\varepsilon\downarrow 0} \mathsf{M}\left\{\int_s^{\tau_R} e^{-\varphi_{s,t}}[(r_t+\varepsilon)\det a_t]^{1/(d+1)}\right.$$
$$\left.\times f\left(\int_s^t (r_u+\varepsilon)\,du, x_t\right)dt\,|\mathscr{F}_s\right\} \qquad \text{(a.s.)},$$

we conclude that it is enough to consider the case where $r_t(\omega) > 0$ for all (t,ω).

We fix $T > 0$ and assume that $h(y,x) = f(T-y,x)$ for $0 < y < T$, $x \in S_R$, and $h = 0$ in all the remaining cases. Using Lemma 7, we find an appropriate function z. Let $\tau = \tau_{T,R}$ be the first exit time of a process $(y_{s,t},x_t)$ considered for $t \ge s$ from a set $[0,T) \times S_R$.

We apply Ito's formula to the expression $e^{-\varphi_{s,t}}z^{(\varepsilon)}(T-y_{s,t},x_t)$ for $\varepsilon > 0$, $t \ge s$. Then

$$-z^{(\varepsilon)}(T,x_s) = \mathsf{M}\left\{\int_s^{t\wedge\tau}\left[-r_u\frac{\partial}{\partial t}z^{(\varepsilon)}(T-y_{s,u},x_u)\right.\right.$$
$$\left.-c_u z^{(\varepsilon)}(T-y_{s,u},x_u) + L^{\sigma_u,b_u}z^{(\varepsilon)}(T-y_{s,u},x_u)\right]e^{-\varphi_{s,t}}\,du$$
$$\left.-e^{-\varphi_{s,t\wedge\tau}}z^{(\varepsilon)}(T-y_{s,t\wedge\tau},x_{t\wedge\tau})\,|\mathscr{F}_s\right\} \qquad \text{(a.s.)}.$$

Using the properties of $z^{(\varepsilon)}$ for small $\varepsilon > 0$, we find

$$-r_u\frac{\partial}{\partial t}z^{(\varepsilon)} + L^{\sigma_u,b_u}z^{(\varepsilon)} = r_u\left[-\frac{\partial}{\partial t}z^{(\varepsilon)} + \sum_{i,j=1}^d \frac{1}{r_u}a_u^{ij}z^{(\varepsilon)}_{x^ix^j}\right] + \sum_{i=1}^d b_u^i z^{(\varepsilon)}_{x^i}$$
$$\ge N(d)\psi_u h^{(\varepsilon)} + \frac{2}{R}|b_u|z^{(\varepsilon)}.$$

Furthermore, $z^{(\varepsilon)} \le 0$. Hence

$$-z^{(\varepsilon)}(T,x_s) \ge \mathsf{M}\left\{\int_s^{t\wedge\tau} e^{-\varphi_{s,u}}\left[N(d)\psi_u h^{(\varepsilon)}(T-y_{s.u},x_u)\right.\right.$$
$$\left.\left.+\frac{2}{R}|b_u|z^{(\varepsilon)}(T-y_{s,u},x_u)\right]du\,|\mathscr{F}_s\right\} \qquad \text{(a.s.)},$$

in which we carry the term containing $z^{(\varepsilon)}$ from the right side to the left side. Also, we use the estimate $|z^{(\varepsilon)}| \le \sup_{t,x}|z| \le N\|h\|_{d+1,C_R} \le N\|f\|_{d+1,C_R}$:

$$N(d,R)\|f\|_{d+1,C_R}\left(1 + \mathsf{M}\left\{\int_s^{t\wedge\tau} e^{-\varphi_{s,u}}|b_u|\,du\,|\,\mathscr{F}_s\right\}\right)$$
$$\ge \mathsf{M}\left\{\int_s^{t\wedge\tau} e^{-\varphi_{s,u}}\psi_u h^{(\varepsilon)}(T - y_{s,u}, x_u)\,du\,|\,\mathscr{F}_s\right\} \qquad \text{(a.s.)},$$

where $y_{s,u} \in (0,T)$ for $u \in (s,\tau)$ by virtue of the condition $r_t > 0$, and in addition, $x_u \in S_R$; hence the function h is continuous at a point $(T - y_{s,u}, x_u)$ and $h(T - y_{s,u}, x_u) = f(y_{s,u}, x_u)$. Letting ε to zero in the last inequality, we obtain, using Fatou's lemma,

$$N(d,R)\|f\|_{d+1,C_R}\left(1 + \mathsf{M}\left\{\int_s^{t\wedge\tau} e^{-\varphi_{s,u}}|b_u|\,du\,|\,\mathscr{F}_s\right\}\right)$$
$$\ge \mathsf{M}\left\{\int_s^{t\wedge\tau} e^{-\varphi_{s,u}}\psi_u f(y_{s,u}, x_u)\,du\,|\,\mathscr{F}_s\right\} \qquad \text{(a.s.)}.$$

Further, on the set $\{\tau_R \ge s\}$ it is seen that $\tau \le \tau_R$. Therefore, by Lemma 5,

$$\chi_{\tau_R \ge s}\mathsf{M}\left\{\int_s^{t\wedge\tau} e^{-\varphi_{s,u}}|b_u|\,du\,|\,\mathscr{F}_s\right\}$$
$$= \mathsf{M}\left\{\chi_{\tau_R \ge s}\int_s^{t\wedge\tau} e^{-\varphi_{s,u}}|b_u|\,du\,|\,\mathscr{F}_s\right\}$$
$$\le \chi_{\tau_R \ge s}\mathsf{M}\left\{\int_s^{\tau_R} e^{-\varphi_{s,u}}|b_u|\,du\,|\,\mathscr{F}_s\right\} \le N(F,R) \qquad \text{(a.s.)}.$$

Finally, on the set $\{\tau_R \ge s\}$ for all $T > 0$, $t > s$, we obtain

$$\mathsf{M}\left\{\int_s^{t\wedge\tau_{T,R}} e^{-\varphi_{s,u}}\psi_u f(y_{s,u}, x_u)\,du\,|\,\mathscr{F}_s\right\} \le N(d,F,R)\|f\|_{d+1,C_R} \qquad \text{(a.s.)}.$$

It remains only to let first $t \to \infty$, second $T \to \infty$, and then to use Fatou's lemma as well as the fact that obviously $\tau_{T,R} \to \tau_R$ as $T \to \infty$ on the set $\{\tau_R \ge s\}$. We have thus proved the lemma. □

9. Proof of Theorem 2. We note first that it suffices to prove Theorem 2 only for $p = d$. In fact, for $p > d$ in accord with Hölder's inequality, for example,

$$\mathsf{M}\left\{\int_s^{\tau_D} e^{-\varphi_{s,t}} c_t^{1-(d/p)}(\det a_t)^{1/p}|g(x_t)|\,dt\,|\,\mathscr{F}_s\right\}$$
$$\le \left(\mathsf{M}\left\{\int_s^{\tau_D} e^{-\varphi_{s,t}}(\det a_t)^{1/d}|g(x_t)|^{p/d}\,dt\,|\,\mathscr{F}_s\right\}\right)^{p/d}\left(\mathsf{M}\left\{\int_s^{\tau_D} e^{-\varphi_{s,t}} c_t\,dt\,|\,\mathscr{F}_s\right\}\right)^{1-(d/p)}$$

In this case, $\int_s^{\tau_D} e^{-\varphi_{s,t}} c_t\,dt = 1 - e^{-\varphi_{s,\tau_D}} \le 1$, and if we have proved the theorem for $p = d$, the first factor does not exceed

$$[N(d,F,D)\|g^{p/d}\|_{d,D}]^{d/p} = N^{d/p}(d,F,D)\|g\|_{p,D} \le (N(d,F,D)+1)\|g\|_{p,D}.$$

The inequality (2) was proved for $p = d$ in Lemma 8. Therefore, it suffices to prove that

$$\mathsf{M}\left\{\int_s^{\tau_D} e^{-\varphi_{s,t}}(\det a_t)^{1/d}|g(x_t)|\,dt \,\Big|\, \mathscr{F}_s\right\} \le N(d,F,D)\|g\|_{d,D}$$

($\{\tau_D \le s\}$-a.s.) for all g. We can consider without loss of generality that g is a nonnegative bounded function. In this case, since $(\det a)^{1/d}$ is a regular function

$$v = \sup_{s\ge 0}\operatorname*{ess\,sup}_{\omega} \chi_{\tau_D\ge s}\mathsf{M}\left\{\int_s^{\tau_D}(\det a_t)^{1/d}g(x_t)e^{-\varphi_{s,t}}\,dt \,\Big|\, \mathscr{F}_s\right\}$$

is finite by Corollary 6. If $v = 0$, we have nothing to prove. Hence we assume that $v > 0$.

Using Fubini's theorem or integrating by parts, we obtain for any numbers $t_1 < t_2$ and nonnegative functions $h(t)$, $r(t)$ that

$$\int_{t_1}^{t_2} h(t)\,dt = \int_{t_1}^{t_2} h(t)\exp\left\{-\int_{t_1}^{t} r(u)\,du\right\}dt$$
$$+ \int_{t_1}^{t_2}\exp\left\{-\int_{t_1}^{t} r(u)\,du\right\}r(t)\left(\int_t^{t_2} h(u)\,du\right)dt.$$

From this for $s \ge 0$, $A \in \mathscr{F}_s$, $r_t = (1/v)(g(x_t)(\det a_t)^{1/d}$, $h_t = (\det a_t)^{1/d}g(x_t)$, we find

$$\mathsf{M}\chi_{A,\tau_D\ge s}\int_s^{\tau_D} h_t e^{-\varphi_{s,t}}\,dt$$
$$= \mathsf{M}\chi_{A,\tau_D\ge s}\int_s^{\tau_D} h_t\exp\left\{-\varphi_{s,t} - \int_s^t r_u\,du\right\}dt$$
$$+ \mathsf{M}\chi_{A,\tau_D\ge s}\int_s^{\tau_D}\exp\left\{-\int_s^t r_u\,du\right\}r_t\left(\int_t^{\tau_D} h_u e^{-\varphi_{s,u}}\,du\right)dt,$$

where the last term is equal to

$$\int_s^{\infty}\left[\mathsf{M}\chi_{A,\tau_D\ge s}\exp\left\{-\int_s^t r_u\,du - \varphi_{s,t}\right\}r_t\chi_{\tau_D\ge t}\left(\int_t^{\tau_D} h_u e^{-\varphi_{t,u}}\,du\right)\right]dt$$
$$= \int_s^{\infty}\left[\mathsf{M}\chi_{A,\tau_D\ge s}\exp\left\{-\int_s^t r_u\,du - \varphi_{s,t}\right\}r_t\chi_{\tau_D\ge t}\mathsf{M}\left\{\int_t^{\tau_D} h_u e^{-\varphi_{t,u}}\,du \,\Big|\, \mathscr{F}_t\right\}\right]dt$$
$$\le \int_s^{\infty}\mathsf{M}\chi_{A,\tau_D\ge s}\exp\left\{-\int_s^t r_u\,du - \varphi_{s,t}\right\}\chi_{\tau_D\ge t}r_t v\,dt$$
$$= \mathsf{M}\chi_{A,\tau_D\ge s}\int_s^{\tau_D} h_t\,exp\left\{-\int_s^t r_u\,du - \varphi_{s,t}\right\}dt.$$

Therefore,

$$\mathsf{M}\chi_{A,\tau_D\geq s}\int_s^{\tau_D} h_t e^{-\varphi_{s,t}}\,dt$$
$$\leq 2\mathsf{M}\chi_{A,\tau_D\geq s}\int_s^{\tau_D} h_t\exp\left\{-\int_s^t r_u\,du-\varphi_{s,t}\right\}dt$$
$$=2v^{1/(d+1)}\mathsf{M}\chi_{A,\tau_D\geq s}\int_s^{\tau_D}(r_t\det a_t)^{1/(d+1)}f\left(\int_s^t r_u\,du,x_t\right)e^{-\varphi_{s,t}}\,dt,$$

where $f(t,x)=e^{-t}g^{d/(d+1)}(x)$. Consequently, by Lemma 8,

$$\mathsf{M}\chi_{A,\tau_D\geq s}\int_s^{\tau_D} h_t e^{-\varphi_{s,t}}\,dt\leq N\|f\|_{d+1,Q}v^{1/(d+1)}\mathsf{P}\{A,\tau_D\geq s\}$$
$$\leq N\|g\|_{d,D}^{d/(d+1)}v^{1/(d+1)}\mathsf{P}\{A,\tau_D\geq s\},$$

where the constants N (which differ from one another) depend only on d, the function $F(c,a)$, and the diameter of D. The last inequality is equivalent to the fact that $\{\tau_D\geq s\}$-a.s.)

$$\mathsf{M}\left\{\int_s^{\tau_D}(\det a_t)^{1/d}g(x_t)e^{-\varphi_{s,t}}\,dt\,\middle|\,\mathscr{F}_s\right\}=\mathsf{M}\left\{\int_s^{\tau_D}h_t e^{-\varphi_{s,t}}\,dt\,\middle|\,\mathscr{F}_s\right\}$$
$$\leq N\|g\|_{d,D}^{d/(d+1)}v^{1/(d+1)}$$

From this, taking the upper bounds, we find

$$v\leq N\|g\|_{d,D}^{d/(d+1)}v^{1/(d+1)},\qquad v^{d(d+1)}\leq N\|g\|_{d,D}^{d/(d+1)}$$

and $v\leq N\|g\|_{d,D}$, thus completing the proof of Theorem 2. □

10. Remark. Let $\delta>0$. The function $F(c,a)$ is said to be δ-regular if for *some* $\varepsilon\in(0,\delta)$ there is a constant $k(\varepsilon)$ such that for all c, a, and unit vectors λ

$$F(c,a)\leq\varepsilon\operatorname{tr}a+k(\varepsilon)[c+(a\lambda,\lambda)].$$

In the sense of the above definition, the function which is δ-regular for all $\delta>0$, is a regular function.

Repeating almost word for word the proofs of Lemmas 5 and 8 and the proof of Theorem 2, we convince ourselves that if the region D belongs to a circle of radius R, $|b_t|\leq F(c_t,a_t)$ for all (t,ω) and if $F(c,a)$ is an R^{-1}-regular function, there exist constants N_1,N_2 depending only on d, $F(c,a)$, and R such that the inequalities (2) and (3) are satisfied.

11. Exercise

Let $d\geq 2$, $D=S_R$, $\varepsilon>0$. Give an example illustrating the $(R^{-1}+\varepsilon)$-regular function $F(c,a)$ for which the assertions of Theorem 2 do not hold. (*Hint*: See Exercise 3.)

12. Exercise

Let $z^{(\varepsilon)}$ be the function from Lemma 7. Prove that for sufficiently small ε the function $z^{(\varepsilon)}(t,x)$ decreases in t and is convex downward with respect to x on the cylinder C_R.

3. Estimates of the Distribution of a Stochastic Integral in the Whole Space

In this section[1] we shall estimate expressions of the form $\mathsf{M}\int_0^\infty |f(t,x_t)|\,dt$ using the $\mathscr{L}_p$-norm of f, that is, we extend the estimates from Section 2.2 to the case $D = E_d$.

We use in this section the assumptions and notation introduced at the beginning of Section 2.2. Furthermore, let

$$\varphi_t = \varphi_{0,t} = \int_0^t c_u\,du, \qquad y_t = y_{0,t} = \int_0^t r_u\,du.$$

Throughout this section we shall have two numbers $K_1, K_2 > 0$ fixed and assume permanently that

$$|b_t(\omega)| \le K_1 c_t(\omega), \qquad \operatorname{tr} a_t(\omega) \le K_2 c_t(\omega)$$

for all (t,ω). Note immediately that under this condition $|b_t|$ does not exceed the regular function $F(c_t,a_t) \equiv K_1 c_t$.

First we prove a version of Theorem 2.2.

1. Lemma. *Let $R > 0$, let τ be a Markov time with respect to $\{\mathscr{F}_t\}$, and let $\tau_R = \inf\{t \ge \tau : |x_t - x_\tau| \ge R\}$.[2] Then there exists a constant $N = N(d,K_1,R)$ such that for any Borel $f(t,x)$*

$$\mathsf{M}\left\{\int_\tau^{\tau_R} e^{-\varphi_t}\psi_t |f(y_t,x_t)|\,dt \,\Big|\, \mathscr{F}_\tau\right\}$$

$$\le e^{-\varphi_\tau}\chi_{\tau<\infty} N\left(\int_{y_\tau}^\infty \int |f(t,x)|^{p+1}\,dx\,dt\right)^{1/(p+1)} \qquad \text{(a.s.).}$$

PROOF. First, let τ be nonrandom finite. For $t \ge 0$ we set $\mathscr{F}'_t = \mathscr{F}_{\tau+t}$, $\mathbf{w}'_t = \mathbf{w}_{\tau+t} - \mathbf{w}_\tau$,

$$\sigma'_t = \sigma_{\tau+t}, \qquad b'_t = b_{\tau+t}, \qquad c'_t = c_{\tau+t}, \qquad r'_t = r_{\tau+t},$$

$$\psi'_t = \psi_{\tau+t}, \qquad x'_t = \int_0^t \sigma'_u\,d\mathbf{w}'_u + \int_0^t b'_u\,du = x_{\tau+t} - x_\tau,$$

$$y'_t = \int_0^t r'_u\,du, \qquad \varphi'_t = \int_0^t c'_u\,du;$$

τ' is the first exit time of the process x'_t from S_R. It is then seen that

$$e^{\varphi_\tau}\mathsf{M}\left\{\int_\tau^{\tau_R} e^{-\varphi_t}\psi_t |f(y_t,x_t)|\,dt \,\Big|\, \mathscr{F}_\tau\right\}$$

$$= \mathsf{M}\left\{\int_0^{\tau'} e^{-\varphi'_t}\psi'_t |f(y'_t + y_\tau, x'_t + x_\tau)|\,dt \,\Big|\, \mathscr{F}'_0\right\} \qquad \text{(a.s.).}$$

[1] Also, see Theorem 4.1.8.
[2] $\inf \phi = \infty$.

Furthermore, $(w'_t, \mathscr{F}'_t)$ is a Wiener process. In addition, by Theorem 2.2

$$\mathsf{M}\left\{\int_0^{\tau'} e^{-\varphi'_t}\psi'_t|f(y'_t + y, x'_t + x)|\,dt\,|\,\mathscr{F}'_0\right\}$$
$$\le N\left(\int_0^{\infty}\int|f(t + y, z)|^{p+1}\,dz\,dt\right)^{1/(p+1)} \qquad \text{(a.s.)}$$

for any $x \in E_d$, $y \ge 0$. In order to prove our lemma for the constant τ, it remains to replace y, x by the $\mathscr{F}'_0$-measurable variables y_τ, x_τ in the last inequality. To do as indicated, we let $\kappa_n(t) = (k + 1)/2^n$ for $t \in (k/2^n, (k + 1)/2^n]$, $\kappa_n(x) = \kappa_n(x^1, \dots, x^d) = (\kappa_n(x), \dots, \kappa_n(x^d))$. Note that $\kappa_n(t) \downarrow t$ for all $t \in (-\infty, \infty)$, $\kappa_n(x) \to x$ for all $x \in E_d$.

From the very start, we can consider without loss of generality that f is a continuous nonnegative function. We denote by Γ_n^1, Γ_n^d the sets of values of the functions $\kappa_n(t)$, $\kappa_n(x)$ respectively. Using Fatou's lemma, we obtain for the function f mentioned,

$$\mathsf{M}\left\{\int_0^{\tau'} e^{-\varphi'_t}\psi'_t f(y'_t + y_\tau, x'_t + x_\tau)\,dt\,|\,\mathscr{F}'_0\right\}$$
$$\le \varliminf_{n\to\infty} \mathsf{M}\left\{\int_0^{\tau'} e^{-\varphi'_t}\psi'_t f(y'_t + \kappa_n(y_\tau), x'_t + \kappa_n(x_\tau))\,dt\,|\,\mathscr{F}'_0\right\}$$
$$= \varliminf_{n\to\infty} \sum_{y\in\Gamma_n^1}\sum_{x\in\Gamma_n^d} \mathsf{M}\left\{\int_0^{\tau'} e^{-\varphi'_t}\psi'_t f(y'_t + y_\tau, x'_t + x)\,dt\,|\,\mathscr{F}'_0\right\}\chi_{\kappa_n(y_\tau)=y,\kappa_n(x_\tau)=x}$$
$$\le N \varliminf_{n\to\infty}\left(\int_{\kappa_n(y_\tau)}\int f^{p+1}(t,x)\,dx\,dt\right)^{1/(p+1)}$$
$$\le N\left(\int_{y_\tau}^{\infty}\int f^{p+1}(t,x)\,dx\,dt\right)^{1/(p+1)}$$

Further, we prove the lemma in the general case. Taking $A \in \mathscr{F}_\tau$ and setting $\tau^n = \kappa_n(\tau)$,

$$\tau_R^n = \inf\{t \ge \tau^n : |x_t - x_{\tau^n}| \ge R\},$$

we can easily see that

$$\tau^n \downarrow \tau, \qquad \varliminf_{n\to\infty} \tau_R^n \ge \tau_R,$$

$$\chi_{\tau<\infty} \varliminf_{n\to\infty} \int_{\tau^n}^{\tau_R^n} e^{-\varphi_t}\psi_t|f(y_t,x_t)|\,dt \ge \int_\tau^{\tau_R} e^{-\varphi_t}\psi_t|f(y_t,x_t)|\,dt,$$

and that for $s \in \Gamma_n^1$ the set

$$\{A, \tau^n = s\} = \left\{A, \tau \in \left(s - \frac{1}{2^n}, s\right]\right\} \in \mathscr{F}_s.$$

Therefore, in accord with what has been proved,

$$\begin{aligned}
&\mathsf{M}\chi_A \int_\tau^{\tau_R} e^{-\varphi_t}\psi_t|f(y_t,x_t)|\,dt \\
&\leq \varliminf_{n\to\infty} \mathsf{M}\chi_{A,\tau<\infty} \int_{\tau^n}^{\tau_R^n} e^{-\varphi_t}\psi_t|f(y_t,x_t)|\,dt \\
&= \varliminf_{n\to\infty} \sum_{s\in\Gamma_n^1} \mathsf{M}\chi_{A,\tau^n=s} \int_s^{\tau_R^n} e^{-\varphi_t}\psi_t|f(y_t,x_t)|\,dt \\
&= \varliminf_{n\to\infty} \sum_{s\in\Gamma_n^1} \mathsf{M}\chi_{A,\tau^n=s}\mathsf{M}\left\{\int_s^{\tau_R^n} e^{-\varphi_t}\psi_t|f(y_t,x_t)|\,dt \,\Big|\, \mathscr{F}_s\right\} \\
&\leq N \varliminf_{n\to\infty} \mathsf{M}\chi_{A,\tau^n<\infty}e^{-\varphi_{\tau^n}}\left(\int_{y_{\tau^n}}^\infty \int |f(t,x)|^{p+1}\,dx\,dt\right)^{1/(p+1)} \\
&= N\mathsf{M}\chi_{A,\tau<\infty}e^{-\varphi_\tau}\left(\int_{y_\tau}^\infty \int |f(t,x)|^{p+1}\,dx\,dt\right)^{1/(p+1)},
\end{aligned}$$

thus completing the proof of our lemma. □

2. Lemma. *As in the preceding lemma, we introduce R, τ, τ_R. Also, we denote by $\mu = \mu(\lambda)$ the positive root of the equation $\lambda - \mu K_1 - \mu^2 K_2 = 0$ for $\lambda > 0$. Then*

$$\mathsf{M}\{\chi_{\tau_R<\infty}e^{-\lambda\varphi_{\tau_R}}|\mathscr{F}_\tau\} \leq \frac{1}{\operatorname{ch}\mu R}e^{-\lambda\varphi_\tau}\chi_{\tau<\infty} \qquad \text{(a.s.).}$$

Proof. Let $\pi(x) = \operatorname{ch}\mu|x|$. Simple computations show that

$$\begin{aligned}
\lambda c_t\pi(x) - L^{\sigma_t,b_t}\pi(x) &= \lambda c_t \operatorname{ch}\mu|x| - \mu \operatorname{sh}\mu|x|\left(b_t,\frac{x}{|x|}\right) \\
&\quad - \mu^2 \operatorname{ch}\mu|x|\frac{(a_tx,x)}{|x|^2} - \mu^2\frac{1}{\mu|x|}\operatorname{sh}\mu|x|\left[\operatorname{tr} a_t - \frac{(a_tx,x)}{|x|^2}\right].
\end{aligned}$$

Taking advantage of the fact that $\operatorname{sh}\mu|x| \leq \operatorname{ch}\mu|x|$, $\operatorname{sh}\mu|x| \leq \mu|x|\operatorname{ch}\mu|x|$, we obviously obtain

$$\lambda c_t\pi(x) - L^{\sigma_t,b_t}\pi(x) \geq c_t\operatorname{ch}\mu|x|(\lambda - \mu K_1 - \mu^2 K_2) = 0.$$

Further, using Ito's formula applied to $e^{-\lambda\varphi_t}\pi(x_t + x)$, we have from the last inequality that

$$\begin{aligned}
e^{-\lambda\varphi_{t\wedge\tau}}\pi(x_{t\wedge\tau} + x) &= \mathsf{M}\left\{\int_{t\wedge\tau}^{t\wedge\tau_R} e^{-\lambda\varphi_u}[\lambda c_u\pi(x_u + x)\right. \\
&\quad \left. - L^{\sigma_u,b_u}\pi(x_u + x)]\,du + e^{-\lambda\varphi_{t\wedge\tau_R}}\pi(x_{t\wedge\tau_R} + x)\,|\,\mathscr{F}_{t\wedge\tau}\right\} \\
&\geq \mathsf{M}\{e^{-\lambda\varphi_{\tau_R}}\pi(x_{t\wedge\tau_R} + x)|\mathscr{F}_{t\wedge\tau}\}.
\end{aligned}$$

Using the continuity property of $\pi(x)$, we replace x with a variable $(-x_{t\wedge\tau})$ in the last inequality. Then

$$e^{-\lambda\varphi_{t\wedge\tau}} \geq \mathsf{M}\{e^{-\lambda\varphi_{\tau_R}}\pi(x_{t\wedge\tau_R} - x_{t\wedge\tau})\,|\,\mathscr{F}_{t\wedge\tau}\},$$

which yields for $A \in \mathscr{F}_\tau$

$$\begin{aligned}\mathsf{M}\chi_{A,\tau<\infty}e^{-\lambda\varphi_\tau} &= \lim_{t\to\infty} \mathsf{M}\chi_{A,\tau\leq t}e^{-\lambda\varphi_{t\wedge\tau}} \\ &\geq \varliminf_{t\to\infty} \mathsf{M}\chi_{A,\tau\leq t}e^{-\lambda\varphi_{\tau_R}}\pi(x_{t\wedge\tau_R} - x_{t\wedge\tau}) \\ &\geq \mathsf{M}\chi_A e^{-\lambda\varphi_{\tau_R}}\pi(x_{\tau_R} - x_\tau)\chi_{\tau_R<\infty} = \operatorname{ch}\mu R\,\mathsf{M}\chi_A e^{-\lambda\varphi_{\tau_R}}\chi_{\tau_R<\infty}. \quad \square\end{aligned}$$

We have proved the lemma; further, we shall prove the main theorem of Section 2.3.

3. Theorem. *There exist constants $N_i = N_i(d,K_1,K_2)$ $(i = 1,2)$ such that for all Markov times τ and Borel functions $f(t,x)$, $g(x)$*

$$\mathsf{M}\left\{\int_\tau^\infty e^{-\varphi_t}\psi_t|f(y_t,x_t)|\,dt\,\Big|\,\mathscr{F}_\tau\right\}$$

$$\leq N_1 e^{-\varphi_\tau}\left(\int_{y_\tau}^\infty \int |f(t,x)|^{p+1}\,dx\,dt\right)^{1/(p+1)} \qquad \text{(a.s.)},$$

$$\mathsf{M}\left\{\int_\tau^\infty e^{-\varphi_t}c_t^{1-(d/p)}(\det a_t)^{1/p}|g(x_t)|dt\,\Big|\,\mathscr{F}_\tau\right\} \leq N_2 e^{-\varphi_\tau}\|g\|_{p,H_d} \qquad \text{(a.s.)}.$$

PROOF. We regard f, g as nonnegative bounded functions and in addition, we introduce the Markov times recursively as follows:

$$\tau^0 = \tau,$$
$$\tau^{n+1} = \inf\{t \geq \tau^n : |x_t - x_{\tau^n}| \geq 1\}.$$

Note that by Lemma 2,

$$\begin{aligned}\mathsf{M}\{\chi_{\tau^{n+1}<\infty}e^{-\varphi_{\tau^{n+1}}}|\mathscr{F}_\tau\} &= \mathsf{M}\{\mathsf{M}\{\chi_{\tau^{n+1}<\infty}e^{-\varphi_{\tau^{n+1}}}|\mathscr{F}_{\tau^n}\}|\mathscr{F}_\tau\} \\ &\leq \frac{1}{\operatorname{ch}\mu}\mathsf{M}\{\chi_{\tau^n<\infty}e^{-\varphi_{\tau^n}}|\mathscr{F}_\tau\} \\ &\leq \left(\frac{1}{\operatorname{ch}\mu}\right)^{n+1} e^{-\varphi_\tau}\chi_{\tau<\infty} \qquad \text{(a.s.)},\end{aligned}$$

where μ is the positive root of the equation $1 - \mu K_1 - \mu^2 K_2 = 0$.

It is seen that τ^n increase as n increases; the variables

$$\chi_{\tau^n<\infty}e^{-\varphi_{\tau^n}}$$

decrease as n increases. The estimate given shows that as $n \to \infty$

$$\mathsf{M}\chi_{\tau^n<\infty}e^{-\varphi_{\tau^n}} \to 0, \qquad \chi_{\tau^n<\infty}e^{-\varphi_{\tau^n}} \to 0 \qquad \text{(a.s.)}.$$

Due to the boundedness of the function $c_t(\omega)$ we immediately have that $\tau^n \to \infty$ (a.s.) as $n \to \infty$.

Therefore, using Lemma 1, we obtain (a.s.)

$$\begin{aligned}
&\mathsf{M}\left\{\int_\tau^\infty e^{-\varphi_t}\psi_t f(y_t,x_t)\,dt \,\middle|\, \mathscr{F}_\tau\right\} \\
&= \sum_{n=0}^\infty \mathsf{M}\left\{\int_{\tau^n}^{\tau^{n+1}} \,\middle|\, \mathscr{F}_\tau\right\} \\
&= \sum_{n=0}^\infty \mathsf{M}\left\{\mathsf{M}\left\{\int_{\tau^n}^{\tau^{n+1}} e^{-\varphi_t}\psi_t f(y_t,x_t)\,dt \,\middle|\, \mathscr{F}_{\tau_n}\right\} \,\middle|\, \mathscr{F}_\tau\right\} \\
&\le N(d,K_1) \sum_{n=0}^\infty \mathsf{M}\left\{e^{-\varphi_{\tau^n}}\chi_{\tau^n<\infty}\left(\int_{y_{\tau^n}}^\infty \int f^{p+1}(t,x)\,dx\,dt\right)^{1/(p+1)} \,\middle|\, \mathscr{F}_t\right\} \\
&\le N\left(\int_{y_\tau}^\infty \int f^{p+1}(t,x)\,dx\,dt\right)^{1/(p+1)} \sum_{n=0}^\infty \mathsf{M}\{e^{-\varphi_{\tau n}}\chi_{\tau^n<\infty} \mid \mathscr{F}_\tau\} \\
&\le N\,\frac{\operatorname{ch}\mu}{\operatorname{ch}\mu - 1}\,e^{-\varphi_\tau}\left(\int_{y_\tau}^\infty \int f^{p+1}(t,x)\,dx\,dt\right)^{1/(p+1)}.
\end{aligned}$$

Having proved the first assertion of the theorem, we proceed to proving the second.

To this end, we use the same technique as in 2.9. The function g is bounded and

$$c_t^{1-(d/p)}(\det a_t)^{1/p} \le c_t^{1-(d/p)}(\operatorname{tr} a_t)^{d/p} \le K_2^{d/p} c_t.$$

Hence

$$\int^\infty e^{-\varphi_{\tau,t}} c_t^{1-(d/p)}(\det a_t)^{1/p} g(x_t)\,dt \le N \int_\tau^\infty e^{-\varphi_{\tau,t}} c_t\,dt = N(1 - e^{-\varphi_{\tau,\infty}}) \le N,$$

and the number

$$v = \sup_\tau \operatorname{ess\,sup}_\omega \mathsf{M}\left\{\int_\tau^\infty e^{-\varphi_{\tau,t}} c_t^{1-(d/p)}(\det a_t)^{1/p} g(x_t)\,dt \,\middle|\, \mathscr{F}_\tau\right.$$

is finite. We assume that $v > 0$, and that $r_t = (1/v)c_t^{1-(d/p)}(\det a_t)^{1/p}g(x_t)$, $h_t = c_t^{1-(d/p)}(\det a_t)^{1/p}g(x_t)$. Using Fubini's theorem, we obtain

$$\begin{aligned}
&\mathsf{M}\left\{\int_\tau^\infty e^{-\varphi_{\tau,t}} h_t\,dt \,\middle|\, \mathscr{F}_\tau\right\} \\
&= \mathsf{M}\left\{\int_\tau^\infty h_t \exp\left\{-\varphi_{\tau,t} - \int_\tau^t r_u\,du\right\} dt \,\middle|\, \mathscr{F}_\tau\right\} \\
&\quad + \mathsf{M}\left\{\int_\tau^\infty r_t \exp\left\{-\varphi_{\tau,t} - \int_\tau^t r_u\,du\right\}\left(\int_t^\infty h_u e^{-\varphi_{t,u}}\,du\right) dt \,\middle|\, \mathscr{F}_\tau\right\} \qquad \text{(a.s.)},
\end{aligned}$$

from which it follows, as in 2.9, that

$$\mathsf{M}\left\{\int_\tau^\infty e^{-\varphi_{\tau,t}} h_t\,dt \,\middle|\, \mathscr{F}_\tau\right\} \le 2\mathsf{M}\left\{\int_\tau^\infty h_t \exp\left\{-\varphi_{\tau,t} - \int_\tau^t r_u\,du\right\} dt \,\middle|\, \mathscr{F}_\tau\right\}.$$

Noting that the last expression equals zero on a set $\{\tau = \infty\}$ we transform it into

$$\exp\left\{\int_0^\tau r_u\,du + \varphi_\tau\right\}\chi_{\tau<\infty} 2v^{1/(p+1)}$$
$$\times \mathsf{M}\left\{\int_\tau^\infty e^{-\varphi_t} c_t^{(p-d)/(p+1)}(r_t \det a_t)^{1/(p+1)} f\left(\int_0^t r_u du, x_t\right) dt \,\middle|\, \mathscr{F}_\tau\right\},$$

where $f(t,x) = e^{-t} g^{p/(p+1)}(x)$. Therefore, according to the first assertion of the theorem

$$\mathsf{M}\left\{\int_\tau^\infty e^{-\varphi_{\tau,t}} h_t\,dt \,\middle|\, \mathscr{F}_\tau\right\} \le N_1 v^{1/(p+1)} e^{y_\tau}\chi_{\tau<\infty}\left(\int_{y_\tau}^\infty \int e^{-(p+1)t} g^p(x)\,dx\,dt\right)^{1/(p+1)}$$
$$\le N_1 v^{1/(p+1)} \|g\|_{p,E_d}^{p/(p+1)}\left(\frac{1}{p+1}\right)^{1/(p+1)} \qquad \text{(a.s.)}.$$

Consequently,

$$v \le N_1 v^{1/(p+1)} \|g\|_{p,E_d}^{p/(p+1)}, \qquad v \le N_1^{1+(1/p)}\|g\|_{p,E_d} \le (1 + N_1^2)\|g\|_{p,E_d},$$

which is equivalent to the second assertion, thus completing the proof of our theorem. □

We give one essential particular case of the theorem proved above.

4. Theorem. *Let* $K_3, K_4 < \infty$, $\lambda > 0$, $\delta > 0$, $s \ge 0$, *for all* $t \ge s$, $\omega \in \Omega$, $\xi \in E_d$

$$|b_t(\omega)| \le K_3, \qquad \sum_{i=1}^d a_t^{ii}(\omega) \le K_4, \qquad \sum_{i,j=1}^d a_t^{ij}(\omega)\xi^i\xi^j \ge \delta|\xi|^2. \tag{1}$$

There exist constants $N_i = N_i(d,p,\lambda,\delta,K_3,K_4)$ $(i = 1,2)$ *such that for all Borel functions* $f(t,x)$, $g(x)$

$$\mathsf{M}\int_s^\infty e^{-\lambda t}|f(t,x_t)|\,dt \le N_1\|f\|_{p+1,H_\infty}, \tag{2}$$

$$\mathsf{M}\int_s^\infty e^{-\lambda t}|g(x_t)|\,dt \le N_2\|g\|_{p,E_d}. \tag{3}$$

This theorem follows from the preceding theorem. In fact, for example, let $r_t = 1$, $c_t = \lambda$ for $t \ge s$, $K_1 = K_3/\lambda$, $K_2 = K_4/\lambda$. Then $|b_t| \le K_1 c_t$, $\operatorname{tr} a_t \le K_2 c_t$ for $t \ge s$. For $t < s$, let us take c_t such that the above inequalities still hold, noting that $(\det a_t)^{1/(p+1)} \ge \delta^{d/(p+1)}$. Therefore

$$\lambda^{(p-d)/(p+1)}\delta^{d/(p+1)}\mathsf{M}\int_s^\infty e^{-\lambda t}|f(t,x_t)|\,dt$$
$$\le e^{-\lambda s}\mathsf{M}\int_s^\infty \exp\left\{-\int_s^t c_u\,du\right\} c_t^{(p-d)/(p+1)}(r_t \det a_t)^{1/(p+1)}|f(y_t,x_t)|\,dt$$
$$\le \mathsf{M}\int_s^\infty e^{-\varphi_{s,t}}\psi_t|f(y_t,x_t)|\,dt.$$

5. Exercise

We replace the third inequality in (1) so that $\det a_t \geq \delta$, and we preserve the first two inequalities. Using the self-scaling property of a Wiener process, and also, using the fact that in (3) $g(x)$ can be replaced by $g(cx)$, prove

$$\mathsf{M}\int_0^\infty e^{-\lambda t}|g(x_t)|\,dt \leq \frac{1}{\sqrt{\lambda}}\delta^{-1/d}K_4^{1/2}N\left(d,\frac{K_2}{\sqrt{\lambda K_4}}\right)\|g\|_{d,E_d},$$

where $N(d,K_3)$ is a finite function nondecreasing with respect to K_3.

4. Limit Behavior of Some Functions

Theorems 6 and 7 are most crucial for the discussion in this section. We shall use them in Chapter 4 in deducing the Bellman equation. However, we use only Corollary 8 in the case of uniform nondegenerate controlled processes. In this regard, we note that the assertion of Corollary 8 follows obviously from intuitive considerations since the lower bound with respect to $\alpha \in \mathfrak{B}(s,x)$ which appears in the assertion of Corollary 8 is the lower bound with respect to a set of uniform nondegenerate diffusion processes with bounded coefficients (see Definition of $\mathfrak{B}(s,x)$ prior to Theorem 5).

We fix the integer d. Also, let the number $p \geq d$ and the numbers $K_1 > 0$, $K_2 > 0$, $K_3 > 0$. We denote by α an arbitrary set of the form

$$(\Omega,\mathscr{F},\mathsf{P},d_1,\mathbf{w}_t,\mathscr{F}_t,\sigma_t,b_t,c_t,r_t), \tag{1}$$

where $(\Omega,\mathscr{F},\mathsf{P})$ is a probability space, the integer $d_1 \geq d$, $(\mathbf{w}_t,\mathscr{F}_t)$ is a d_1-dimensional Wiener process on $(\Omega,\mathscr{F},\mathsf{P})$, $\sigma_t = \sigma_t(\omega)$ is a matrix of dimension $d \times d_1$, $b_t = b_t(\omega)$ is a d-dimensional vector, $c_t = c_t(\omega)$, $r_t = r_t(\omega)$ are non-negative numbers, and σ_t, b_t, c_t, r_t are progressively measurable with respect to $\{\mathscr{F}_t\}$ and are bounded functions of (t,ω) for $t \geq 0$, $\omega \in \Omega$. In the case where the set (1) is written as α, we write $\Omega = \Omega^\alpha$, $\mathscr{F} = \mathscr{F}^\alpha$, etc.

Denote by $\mathfrak{A}(K_1,K_2,K_3)$ the set of all sets α satisfying the conditions

$$|b_t^\alpha| \leq K_1 c_t^\alpha, \qquad \operatorname{tr}\tfrac{1}{2}\sigma_t^\alpha|\sigma_t^\alpha|^* \leq K_2 c_t^\alpha, \qquad r_t^\alpha \leq K_3 c_t^\alpha$$

for all (t,ω). For $x \in E_d$, $\alpha \in \mathfrak{A}(K_1,K_2,K_3)$, let

$$x_t^{\alpha,x} = x + \int_0^t \sigma_u^\alpha\,d\mathbf{w}_u^\alpha + \int_0^t b_u^\alpha\,du,$$

$$y_t^{\alpha,s} = s + \int_0^t r_u^\alpha\,du, \qquad \varphi_t^\alpha = \int_0^t c_u^\alpha\,du, \qquad a_t^\alpha = \tfrac{1}{2}\sigma_t^\alpha[\sigma_t^\alpha]^*,$$

$$\psi_t^\alpha = (c_t^\alpha)^{(p-d)/(p+1)}(r_t^\alpha \det a_t^\alpha)^{1/(p+1)}.$$

As usual, for $p = d$, $\psi_t^\alpha = (r_t^\alpha \det a_t^\alpha)^{1/(d+1)}$.

For the Borel function $f(t,y)$, $s \in (-\infty,\infty)$, $x \in E_d$ let

$$v(s,x) = v(f,s,x) = v(K_1,K_2,K_3,f,s,x)$$
$$= \sup_{\alpha \in \mathfrak{A}(K_1,K_2,K_3)} \mathsf{M}^\alpha \int_0^\infty e^{-\varphi_t^\alpha} \psi_t^\alpha f(y_t^{\alpha,s},x_t^{\alpha,x})\,dt,$$

where M^α denotes the integration over Ω^α with respect to a measure P^α. In addition to the elements mentioned, we shall use the elements given prior to Theorem 5.

1. Theorem. *Let $f \in \mathscr{L}_{p+1}(E_{d+1})$. Then $v(s,x)$ is a continuous function of (s,x) on E_{d+1}, and, furthermore,*

$$|v(s,x)| \le N(d,K_1,K_2)\left(\int_s^\infty \int |f(t,y)|^{p+1}\,dy\,dt\right)^{1/(p+1)}$$

Proof. Since $|b_t^\alpha| \le K_1 c_t^\alpha$, $\operatorname{tr} \alpha_t^\alpha \le K_2 c_t^\alpha$, the estimate of v follows from Theorem 3.3. In this case, we can take $N(d,K_1,K_2) = N_1(d,K_1,K_2)$, where N_1 is the constant given in Theorem 3.3.

Further, we note that for any families of numbers h_1^α, h_2^α,

$$\left|\sup_\alpha h_1^\alpha - \sup_\alpha h_2^\alpha\right| \le \sup_\alpha |h_1^\alpha - h_2^\alpha|.$$

Hence

$$|v(s_1,x_1) - v(s_2,x_2)| \le \sup_\alpha \mathsf{M}^\alpha \int_0^\infty e^{-\varphi_t^\alpha}\psi_t^\alpha |f(y_t^{\alpha,s_1},x_t^{\alpha,x_1}) - f(y_t^{\alpha,s_2},x_t^{\alpha,x_2})|\,dt.$$

If $f(t,x)$ is a smooth function of (t,x), with compact support, then

$$|f(y_t^{\alpha,s_1},x_t^{\alpha,x_1}) - f(y_t^{\alpha,s_2},x_t^{\alpha,x_2})|$$
$$\le \sup_{(t,x)}\left(|\operatorname{grad}_x f(t,x)| + \left|\frac{\partial f(t,x)}{\partial t}\right|\right)(|y_t^{\alpha,s_1} - y_t^{\alpha,s_2}| + |x_t^{\alpha,x_1} - x_t^{\alpha,x_2}|)$$
$$= N(|s_1 - s_2| + |x_1 - x_2|).$$

Morever,

$$\psi_t^\alpha \le (c_t^\alpha)^{(p-d)/(p+1)}(K_2 c_t^\alpha t d^{-d}(\operatorname{tr} a^\alpha)^d)^{1/(p+1)} \le K_3^{1/(p+1)} K_2^{d/(p+1)} c_t^\alpha.$$

Therefore

$$\int_0^\infty e^{-\varphi_t^\alpha}\psi_t^\alpha\,dt \le K_3^{1/(p+1)}K_2^{d/(p+1)} \int_0^\infty e^{-\varphi_t^\alpha}\,d\varphi_t^\alpha \le K^{1/(p+1)}K_2^{d/(p+1)}.$$

Consequently, we have $|v(s_1,x_1) - v(s_2,x_2)| \le N(|s_1 - s_2| + |x_1 - x_2|)$ for $f(t,x)$, with v being a continuous function.

If f is an arbitrary function in $\mathscr{L}_{p+1}(E_{d+1})$, we take a sequence of smooth functions f_n with compact support so that $\|f - f_n\|_{p+1,E_{d+1}} \to 0$. Using the property of the magnitude of the difference between the upper bounds, which we used before, we obtain

$$|v(f,s,x) - v(f_n,s,x)| \le v(|f - f_n|,s,x) \le N\|f - f_n\|_{p+1,E_{d+1}}.$$

This implies that the continuous functions $v(f_n,s,x)$ converge to $v(f,s,x)$ uniformly in E_{d+1}. Therefore, $v(f,s,x)$ is continuous, thus proving the theorem. □

The continuity property of $v(s,x)$ implies the measurability of this function. For investigating the integrability property of $v(s,x)$ we need the following lemma.

2. Lemma. *Let $R > 0$, let $\tau_R^{\alpha,x}$ be the time of first entry of a process $x_t^{\alpha,x}$ into a set $\bar{S}_R$, let γ^α be a random variable on Ω^α, $\gamma^\alpha \ge \tau_R^{\alpha,x}$ and let ε be the positive root of the equation $K_2\varepsilon^2 + K_1\varepsilon - 1 = 0$. Then, for all t_1, s*

$$\mathsf{M}^\alpha \chi_{\gamma^\alpha < \infty} e^{-\varphi_{\gamma^\alpha}^\alpha} \le e^{\varepsilon R - \varepsilon|x|},$$

$$\mathsf{M}^\alpha \chi_{t_1 \le y_{\gamma^\alpha}^{\alpha,s}, \gamma^\alpha < \infty} e^{-\varphi_{\gamma^\alpha}^\alpha} \le \exp\left\{\frac{\varepsilon}{2} R - \frac{\varepsilon}{2}|x| - \frac{1}{2K_3}(t_1 - s)\right\}.$$

PROOF. We fix α, x. For the sake of simplicity we do not write the superscripts α, x. In addition, we write $y_{\gamma^\alpha}^{\alpha,s} = s + y_{\gamma^\alpha}^{\alpha,0}$ as $s + y_\gamma$.

The first assertion of the lemma is obvious for $|x| \le R$; therefore we assume that $|x| > R$. In accord with Ito's formula applied to $e^{-\varphi_t - \varepsilon|x_t|}$ we obtain

$$e^{-\varepsilon R}\mathsf{M}e^{-\varphi_\gamma}\chi_{\gamma \le t} \le e^{-\varepsilon R}\mathsf{M}e^{-\varphi_{\tau_R}}\chi_{\tau_R \le t}$$

$$\le \mathsf{M}e^{-\varphi_{t\wedge\tau_R} - \varepsilon|x_{t\wedge\tau_R}|} = e^{-\varepsilon|x|} + \mathsf{M}\int_0^{t\wedge\tau_R} e^{-\varphi_s - \varepsilon|x_s|}\, I_s(x_s)\, ds,$$

where

$$I_s(x) = \varepsilon^2 \frac{(a_s x,x)}{|x|^2} - \varepsilon\frac{1}{|x|}\left[\operatorname{tr} a_s - \frac{(a_s x,x)}{|x|^2}\right] - \varepsilon b_s \frac{x}{|x|} - c_s$$

$$\le \varepsilon^2 \frac{(a_s x,x)}{|x|^2} + \varepsilon K_1 c_s - c_s \le \varepsilon^2 K_2 c_s + \varepsilon K_1 c_s - c_s = 0.$$

Hence $e^{-\varepsilon R}\mathsf{M}e^{-\varphi_\gamma}\chi_{\gamma \le t} \le e^{-\varepsilon|x|}$. Using Fatou's lemma, as $t \to \infty$, we arrive at the former inequality.

In order to prove the latter inequality, we note that under the assumption $r_t \le K_3 c_t$ we have on the set $\{t_1 \le y_\gamma + s\}$ that $t_1 - s \le K_3\varphi_\gamma$, from which it follows that

$$-\varphi_\gamma \le -K_3^{-1}(t_1 - s),$$

$$\mathsf{M}\chi_{t_1 \le y_\gamma + s,\, \gamma < \infty} e^{-\varphi_\gamma} \le e^{-K_3^{-1}(t_1 - s_\gamma)}.$$

Furthermore,

$$\mathsf{M}\chi_{t_1 \le y_\gamma + s,\, \gamma < \infty} e^{-\varphi_\gamma} \le \mathsf{M}\chi_{\gamma < \infty} e^{-\varphi_\gamma} \le e^{\varepsilon R - \varepsilon|x|}.$$

Having multiplied the extreme terms in the last two inequalities we establish the second assertion of the lemma, thus completing the proof of the lemma. □

3. Theorem. *There exists a finite function $N(d,K_1)$ increasing with respect K_1 and such that for all $f \in \mathscr{L}_{p+1}(E_{d+1})$*

$$\|v(f,\cdot,\cdot)\|_{p+1,E_{d+1}} \le K_3^{1/(p+1)}K_2^{d/(p+1)}N\left(d,\frac{K_1}{\sqrt{K_2}}\right)\|f\|_{p+1,E_{d+1}}.$$

PROOF. Suppose that we have proved the theorem under the condition that $K_2 = K_3 = 1$. In order to prove the theorem under the same assumption in the general case, we use arguments which replace implicitly the application of the self-scaling property of a Wiener process (see Exercise 3.5).

If $\alpha \in \mathfrak{A} = \mathfrak{A}(K_1,K_2,K_3)$, let

$$\alpha' = \left(\Omega^\alpha,\mathscr{F}^\alpha,\mathrm{P}^\alpha,d_1^\alpha,\mathrm{w}_t^\alpha,\mathscr{F}_t^\alpha,\frac{1}{\sqrt{K_2}}\sigma_t^\alpha,\frac{1}{\sqrt{K_2}}b_t^\alpha,c_t^\alpha,\frac{1}{K_3}r_t^\alpha\right).$$

It is seen that $\alpha' \in \mathfrak{A} = \mathfrak{A}(K_1/\sqrt{K_2},1,1)$. It is also seen that α' runs through the entire set $\mathfrak{A}(K_1/\sqrt{K_2},1,1)$ when α runs through the entire set $\mathfrak{A}(K_1,K_2,K_3)$.

Further, for $f \in \mathscr{L}_{p+1}(E_{d+1})$ let $f'(t,x) = f(K_3t,\sqrt{K_2}x)$. We have

$$\begin{aligned}
&v(K_1,K_2,K_3,f,s,x)\\
&= \sup_{\alpha\in\mathfrak{A}} \mathrm{M}^\alpha \int_0^\infty e^{-\varphi_t^\alpha}\psi_t^\alpha f(y_t^{\alpha,s},x_t^{\alpha,x})\,dt\\
&= K_3^{1/(p+1)}K_2^{d/(p+1)} \sup_{\alpha'\in\mathfrak{A}'} \mathrm{M}^{\alpha'} \int_0^\infty e^{-\varphi_t^{\alpha'}}\psi_t^{\alpha'} f(s + K_3y_t^{\alpha',0}, x + \sqrt{K_2}x_t^{\alpha',0})\,dt\\
&= K_3^{1/(p+1)}K_2^{d/(p+1)} \sup_{\alpha'\in\mathfrak{A}'} \mathrm{M}^{\alpha'} \int_0^\infty e^{-\varphi_t^{\alpha'}}\psi_t^{\alpha'} f'\left(\frac{s}{K_3} + y_t^{\alpha',0},\frac{x}{\sqrt{K_2}} + x_t^{\alpha',0}\right)dt\\
&= K_3^{1/(p+1)}K_2^{d/(p+1)}v\left(\frac{K_1}{\sqrt{K_2}},1,1,f',\frac{s}{K_3},\frac{x}{\sqrt{K_2}}\right).
\end{aligned}$$

Therefore, if we have proved our theorem for $K_2 = K_3 = 1$, then

$$\begin{aligned}
&\|v(K_1,K_2,K_3,f,\cdot,\cdot)\|_{p+1,E_{d+1}}^{p+1}\\
&= K_3K_2^d \int_{-\infty}^\infty \int \left|v\left(\frac{K_1}{\sqrt{K_2}},1,1,f',\frac{s}{K_3},\frac{x}{\sqrt{K_2}}\right)\right|^{p+1} dx\,ds\\
&= K_3^2K_2^{(3/2)d} \int_{-\infty}^\infty \int \left|v\left(\frac{K_1}{\sqrt{K_2}},1,1,f',s,x\right)\right|^{p+1} dx\,ds\\
&\le K_3^2K_2^{(3/2)d}N^{p+1}\left(d,\frac{K_1}{\sqrt{K_2}}\right)\|f'\|_{p+1,E_{d+1}}^{p+1}\\
&= K_3K_2^dN^{p+1}\left(d,\frac{K_1}{\sqrt{K_2}}\right)\|f\|_{p+1,E_{d+1}}^{p+1}.
\end{aligned}$$

Therefore, it suffices to prove this theorem only for $K_2 = K_3 = 1$. We use in our proof in this case the expression

$$I^\alpha(s,x) \equiv \mathsf{M}^\alpha \int_0^\infty e^{-\varphi_t^\alpha}\psi_t^\alpha f(y_t^{\alpha,s},x_t^{\alpha,x})\,dt$$

representable as the "sum" of terms each of which incorporates the change which occurs while the process $(y_t^{\alpha,s},x_t^{\alpha,s})$ moves across the region associated with the given term.

We assume without loss of generality that $f \geq 0$.

Let R be such that the volume of S_R is equal to unity. We denote by $w(t,x)$ the indicator of a set $\bar{C}_{1,R}$. Let $f_{(t_1,x_t)}(t,x) = w(t_1 - t,x_1 - x)f(t,x)$. It is seen that

$$f(t,x) = \int_{-\infty}^\infty \int f_{(t_1,x_1)}(t,x)\,dx_1\,dt_1,$$

$$I^\alpha(s,x) = \int_{-\infty}^\infty dt_1 \int dx_1\,\mathsf{M}^\alpha \int_0^\infty e^{-\varphi_t^\alpha}\psi_t^\alpha f_{(t_1,x_1)}(y_t^{\alpha,s},x_t^{\alpha,x})\,dt.$$

In order to estimate the last expectation for fixed t_1, x_1, we note that $f_{(t_1,x_1)}(t,x)$ can be nonzero only for $0 \leq t_1 - t \leq 1$, $|x_1 - x| \leq R$. Hence, if γ^α is the time of first entry of the process $(t_1 - y_t^{\alpha,s} x_1 - x_t^{\alpha,x})$ into the set $\bar{C}_{1,R}$, then

$$\mathsf{M}^\alpha \int_0^\infty e^{-\varphi_t^\alpha}\psi_t^\alpha f_{(t_1,x_1)}(y_t^{\alpha,s},x_t^{\alpha,x})\,dt = \mathsf{M}^\alpha \chi_{\gamma^\alpha<\infty} \int_{\gamma^\alpha}^\infty e^{-\varphi_t^\alpha}\psi_t^\alpha f_{(t_1,x_1)}(y_t^{\alpha,s},x_t^{\alpha,x})\,dt.$$

Furthermore, on the set $\{\gamma^\alpha < \infty\}$

$$0 \leq t_1 - y_{\gamma^\alpha}^{\alpha,s} \leq 1 \qquad \text{and} \qquad R \geq |x_1 - x_{\gamma^\alpha}^{\alpha,x}| = |x_{\gamma^\alpha}^{\alpha,x-x_1}|.$$

The last inequality in the preceding lemma implies the inequality $\gamma^\alpha \geq \tau_R^{\alpha,x-x_1}$. By Theorem 3.3 and the preceding lemma we obtain

$$\begin{aligned}
&\mathsf{M}^\alpha \int_0^\infty e^{-\varphi_t^\alpha}\psi_t^\alpha f_{(t_1,x_1)}(y_t^{\alpha,s},x_t^{\alpha,x})\,dt \\
&\quad = \mathsf{M}^\alpha \chi_{t_1-1\leq y_{\gamma^\alpha}^{\alpha,s},\,\gamma^\alpha<\infty} \times \mathsf{M}^\alpha\left\{\int_{\gamma^\alpha}^\infty e^{-\varphi_t^\alpha}\psi_t^\alpha f_{(t_1,x_1)}(y_t^{\alpha,s},x_t^{\alpha,x})\,dt \,\middle|\, \mathscr{F}_{\gamma^\alpha}^\alpha\right\} \\
&\quad \leq N_1 \|f_{(t_1 x_1)}\|_{p+1,E_{d+1}} \mathsf{M}^\alpha \chi_{t_1-1\leq y_{\gamma^\alpha}^{\alpha,s},\,\gamma^\alpha<\infty} e^{-\varphi_{\gamma^\alpha}^\alpha} \\
&\quad \leq N_1 \|f_{(t_1,x_1)}\|_{p+1,E_{d+1}} \exp\left\{\frac{\varepsilon}{2}R - \frac{\varepsilon}{2}|x - x_1| - \frac{1}{2}(t_1 - s - 1)\right\},
\end{aligned}$$

where $N_1 = N_1(d,K_1,1)$ is the constant given in Theorem 3.3. Also, we note that for $t_1 < s$ the first expression in the above computations is equal to zero since $t_1 - y_t^{\alpha,s} \leq t_1 - s < 0$ and $\gamma^\alpha = \infty$. Hence

$$I^\alpha(s,x) \leq N_1 \int_{-\infty}^\infty dt_1 \int dx_1 \|f_{(t_1,x_1)}\|_{p+1,E_{d+1}} \pi(s - t_1, x - x_1),$$

where $\pi(t,x) = \exp[(\varepsilon/2)R - (\varepsilon/2)|x| + \frac{1}{2}(t + 1)]$ for $t \leq 0$, $\pi(t,x) = 0$ for $t > 0$. Therefore, since $v = \sup_\alpha I^\alpha$,

$$v(s,x) \leq N_1 \int_{-\infty}^\infty \int \|f_{(t_1,x_1)}\|_{p+1,E_{d+1}} \pi(s - t_1, x - x_1)\,dx_1\,dt_1.$$

In the right side of the last expression there is a convolution (with respect to (t_1,x_1)) of the two functions $\|f_{(t_1,x_1)}\|_{p+1,E_{d+1}}$ and $\pi(t_1,x_1)$. It is a well-known fact that the norm of the convolution in $\mathscr{L}_p$ does not exceed the product of the norm of one function in $\mathscr{L}_p$ and the norm of the other function in $\mathscr{L}_1$. Using this fact, we conclude that

$$\begin{aligned}\|v\|_{p+1,E_{d+1}} &\leq N_1 \int_{-\infty}^{\infty}\int \pi(t,x)\,dx\,dt \big\| \|f_{(t_1,x_1)}\|_{p+1,E_{d+1}}\big\|_{p+1,E_{d+1}} \\ &= N(d,K_1)\Bigg[\int_{-\infty}^{\infty}\int dt_1\,dx_1 \\ &\quad \times \left(\int_{-\infty}^{\infty}\int w(t_1-t,x_1-x)f^{p+1}(t,x)\,dx\,dt\right)\Bigg]^{1/(p+1)} \\ &= N(d,K_1)\|f\|_{p+1,E_{d+1}}.\end{aligned}$$

To complete the proof of the theorem it remains only to show that the last constant $N(d,K_1)$ can be regarded as an increasing function of K_1.

Let

$$\tilde{N}(d,K_1) = \sup \|v(K_1,1,1,|f|,\cdot,\cdot)\|_{p+1,E_{d+1}}\|f\|_{p+1,E_{d+1}}^{-1},$$

where the upper bound is taken over all $f \in \mathscr{L}_{p+1}(E_{d+1})$ such that $\|f\|_{p+1,E_{d+1}} > 0$. According to what has been proved above, $\tilde{N}(d,K_1) < \infty$. In addition, the sets $\mathfrak{A}$ increase with respect to K_1. Hence v, $\tilde{N}(d,K_1)$ increase with respect to K_1. Finally, it is seen that $|v(f,s,x)| \leq v(|f|,s,x)$ and

$$\|v(K_1,1,1,f,\cdot,\cdot)\|_{p+1,E_{d+1}} \leq \tilde{N}(d,K_1)\|f\|_{p+1,E_{d+1}}.$$

The theorem has been proved. □

We extend the assertions of Theorem 1 and 3 to the case where the function $f(t,x)$ does not depend on t. However, we do not consider here the process r_t, as we did in the previous sections. Let

$$\begin{aligned}v(x) = v(g,x) &= v(K_1,K_2,g,x) \\ &= \sup_{\alpha\in\mathfrak{A}(K_1,K_2,0)} \mathsf{M}^\alpha \int_0^\infty e^{-\varphi_t^\alpha}(c_t^\alpha)^{(p-d)/p}(\det a_t^\alpha)^{1/p} g(x_t^{\alpha,x})\,dt.\end{aligned}$$

4. Theorem. (a) *Let $g \in \mathscr{L}_p(E_d)$; then $v(x)$ is a continuous function,*

$$|v(x)| \leq N(d,K_1,K_2)\|g\|_{p,E_d}.$$

(b) *There exists a finite function $N(d,K_1)$ increasing with respect to K_1 and such that for all $g \in \mathscr{L}_p(E_d)$*

$$\|v(g,\cdot)\|_{p,E_d} \leq K_2^{d/p} N\left(d,\frac{K_1}{\sqrt{K_2}}\right)\|g\|_{p,E_d}.$$

This theorem can be proved in the same way as Theorems 1 and 3.

We proceed now to consider the main results of the present section. Let numbers $K > 0$, $\delta > 0$ be fixed, and let each point $(t,x) \in E_{d+1}$ $(x \in E_d)$ be associated with some nonempty set $\mathfrak{B}(t,x)$ (respectively, $\mathfrak{B}(x)$) consisting of sets α of type (1). Let $\mathfrak{B}$ be the union of all sets $\mathfrak{B}(t,x)$, $\mathfrak{B}(x)$. We assume that a function $c_u^\alpha(\omega)$ is bounded on $\mathfrak{B} \times [0,\infty) \times \bigcup_\alpha \Omega^\alpha$ and that for all $\alpha \in \mathfrak{B}$, $u \in [0,\infty)$, $\omega \in \Omega^\alpha$, $y \in E_d$

$$|b_u^\alpha| \le K, \qquad \operatorname{tr} \sigma_u^\alpha[\sigma_u^\alpha]^* \le K, \qquad r_u^\alpha = 1, \\ |[\sigma_u^\alpha]^* y| \ge \delta |y|. \tag{2}$$

It is useful to note that (2) can be rewritten as

$$(a_t^\alpha y, y) = \sum_{i,j=1}^d (a_t^\alpha)^{ij} y^i y^j \ge \frac{1}{2}\delta^2 |y|^2,$$

since $(a_t^\alpha y,y) = \frac{1}{2}(\sigma_t^\alpha(\sigma_t^\alpha)^* y,y) = \frac{1}{2}|(\sigma_t^\alpha)^* y|^2$.

5. Theorem. (a) *Let $\lambda \ge \lambda_0 > 0$, let $Q \subset E_{d+1}$, let Q be an open set, and let $f \in \mathscr{L}_{p+1}(Q)$,*

$$\tau^\alpha = \tau^{\alpha,s,x} = \inf\{t \ge 0 : (t + s, x_t^{\alpha,x}) \notin Q\},$$

$$= \sup_{\alpha \in \mathfrak{B}(s,x)} \mathsf{M}^\alpha \int_0^{\tau^\alpha} e^{-\varphi_t^\alpha - \lambda t} f(t + s, x_t^{\alpha,x})\, dt.$$

Then

$$\lambda \| z^\lambda \|_{p+1,Q} \le N(d,K,\delta,\lambda_0) \| f \|_{p+1,Q}.$$

(b) *Let $\lambda \ge \lambda_0 > 0$, let $D \subset E_d$, let D be an open set, and let $g \in \mathscr{L}_p(D)$,*

$$\tau^\alpha = \tau^{\alpha,x} = \inf\{t \ge 0 : x_t^{\alpha,x} \notin D\},$$

$$z^\lambda(x) = \sup_{\alpha \in \mathfrak{B}(x)} \mathsf{M}^\alpha \int_0^{\tau^\alpha} e^{-\varphi_t^\alpha - \lambda t} g(x_t^{\alpha,x})\, dt.$$

Then

$$\lambda \| z^\lambda \|_{p,D} \le N(d,K,\delta,\lambda_0) \| g \|_{p,D}.$$

Proof. Since all eigenvalues of the matrix a_t^α are greater than $\frac{1}{2}\delta^2$, $\det a_t^\alpha \ge 2^\alpha \delta^{2d}$. From this, assuming that $\tilde{f} = |f| \chi_Q$, $\tilde{c}_t^\alpha = c_t^\alpha + \lambda$, $\tilde{\varphi}_t^\alpha = \varphi_t^\alpha + \lambda t$ and noting that $\tilde{c}_t^\alpha \ge \lambda$, we find

$$|z^\lambda(s,x)| \le N(\delta) \lambda^{(d-p)/(p+1)} \sup_{\alpha \in \mathfrak{B}(s,x)} \mathsf{M}^\alpha \int_0^\infty e^{-\tilde{\varphi}_t^\alpha} (\tilde{c}_t^\alpha)^{(p-d)/(p+1)} \\ \times (r_t^\alpha \det a_t^\alpha)^{1/(p+1)} \tilde{f}(y_t^{\alpha,s}, x_t^{\alpha,s})\, dt.$$

It is seen that

$$|b_t^\alpha| \le \frac{K}{\lambda} \tilde{c}_t^\alpha, \qquad \operatorname{tr} a_t^\alpha \le \frac{K}{\lambda} \tilde{c}_t^\alpha, \qquad \text{and} \qquad r_t^\alpha \le \frac{1}{\lambda} \tilde{c}_t^\alpha.$$

Therefore,

$$|z^\lambda(s,x)| \le v\left(\frac{K}{\lambda}, \frac{K}{\lambda}, \frac{1}{\lambda}, \tilde{f}, s, x\right) N(\delta) \lambda^{(d-p)/(p+1)},$$

which implies, by Theorem 3, that

$$]\|z^\lambda\|_{p+1,Q} \le N(\delta)\lambda^{(d-p)/(p+1)}\|v\|_{p+1,E_{d+1}}$$
$$\le N(\delta)\lambda^{-1}K^{d/(p+1)}N\left(d,\sqrt{\frac{K}{\lambda}}\right)\|\tilde{f}\|_{p+1,E_{d+1}},$$

thus proving assertion (a) of our theorem since

$$\|\tilde{f}\|_{p+1,E_{d+1}} = \|f\|_{p+1,Q}, \quad K^{d/(p+1)} \le 1 + K, \quad N\left(d,\sqrt{\frac{K}{\lambda}}\right) \le N\left(d,\sqrt{\frac{K}{\lambda_0}}\right).$$

Proceeding in the same way, we can prove assertion (b) with the aid of Theorem 4. The theorem is proved. □

6. Theorem. (a) *Suppose that Q is a region in E_{d+1}, $f_1(t,x)$ is a bounded Borel function, $f \in \mathscr{L}_{p+1}(Q)$, $\lambda > 0$,*

$$\tau^\alpha = \tau^{\alpha,s,x} = \inf\{t \ge 0 : (s+t, x_t^{\alpha,x}) \notin Q\},$$
$$z^\lambda(s,x) = z^\lambda(f,s,x)$$
$$= \sup_{\alpha\in\mathfrak{B}(s,x)} \mathsf{M}^\alpha\left[\int_0^{\tau^\alpha} e^{-\varphi_t^\alpha-\lambda t} f(s+t, x_t^{\alpha,x})\,dt + e^{-\varphi_{\tau^\alpha}^\alpha-\lambda\tau^\alpha} f_1(s+\tau^\alpha, x_{\tau^\alpha}^{\alpha,x})\right].$$

Then, there exists a sequence $\lambda_n \to \infty$ such that $\lambda_n z^{\lambda_n}(s,x) \to f(s,x)$ (Q-a.s.).

(b) *Suppose that D is a region in E_d, $g_1(x)$ is a bounded Borel function, $g \in \mathscr{L}_p(D)$, $\lambda > 0$,*

$$\tau^\alpha = \tau^{\alpha,x} = \inf\{t \ge 0 : x_t^{\alpha,x} \notin D\},$$
$$z^\lambda(x) = z^\lambda(g,x)$$
$$= \sup_{\alpha\in\mathfrak{B}(x)} \mathsf{M}^\alpha\left[\int_0^{\tau^\alpha} e^{-\varphi_t^\alpha-\lambda t} g(x_t^{\alpha,x})\,dt + e^{-\varphi_{\tau^\alpha}^\alpha-\lambda\tau^\alpha} g_1(x_{\tau^\alpha}^{\alpha,x})\right].$$

Then, there exists a sequence $\lambda_n \to \infty$ such that $\lambda_n z^{\lambda_n}(x) \to g(x)$($D$-a.s.).

7. Theorem. (a) *We introduce another element in Theorem* 6a. *Suppose that Q' is a bounded region $Q' \subset \bar{Q}' \subset Q$. Then $]\|\lambda z^\lambda - f\|_{p+1,Q'} \to 0$ as $\lambda \to \infty$. If $f_1 \equiv 0$, we can take $Q' = Q$.*

(b) *Suppose that in Theorem* 6b *D' is a bounded region, $D' \subset \bar{D}' \subset D$; then $]\|\lambda z^\lambda - g\|_{p,D'} \to 0$ as $\lambda \to \infty$. If $g_1 \equiv 0$, we can take $D' = D$.*

Proof of Theorems 6 and 7. It was noted in Section 2.1 that the property of convergence with respect to an exterior norm implies the existence of a subsequence convergent almost everywhere. Using this fact, we can easily see that only Theorem 7 is to be proved.

Proof of Theorem 7a. First, let $f_1 \equiv 0$. We take a sequence of functions $f^n \in C_0^\infty(Q)$ such that $\|f^n - f\|_{p+1,Q} \to 0$. It is seen that

$$|\lambda z^\lambda(f,s,x) - f(s,x)| \le \lambda|z^\lambda(f,s,x) - z^\lambda(f^n,s,x)|$$
$$+ |\lambda z^\lambda(f^n,s,x) - f^n(s,x)| + |f^n(s,x) - f(s,x)|,$$

from which, noting that $|z^\lambda(f,s,x) - z^\lambda(f^n,s,x)| \leq z^\lambda(|f - f^n|,s,x)$, *we obtain,* in accord with Theorem 5a,

$$\overline{\lim_{\lambda\to\infty}}\,]\!|\lambda z^\lambda(f,\cdot,\cdot) - f(\cdot,\cdot)\|_{p+1,Q}$$
$$\leq N(d,K,\delta,1)\|f - f^n\|_{p+1,Q} + \overline{\lim_{\lambda\to\infty}}\,]\!|\lambda z^\lambda(f^n,\cdot,\cdot) - f^n(\cdot,\cdot)\|_{p+1,Q} + \|f^n - f\|_{p+1,Q}$$

In the last inequality the left side does not depend on n; the first and third terms in the right side can be made arbitrarily small by choosing an appropriate n. In order to make sure that the left side of the last inequality is equal to zero, we need only to show that for each n

$$\overline{\lim_{\lambda\to\infty}}\,]\!|\lambda z^\lambda(f^n,\cdot,\cdot) - f^n(\cdot,\cdot)\|_{p+1,Q} = 0.$$

In short, it suffices to prove assertion (a) for $f_1 \equiv 0$ in the case $f \in C_0^\infty(Q)$.

In conjunction with Ito's formula applied to $f(s + t, x_t^{\alpha,x})l^{-\varphi_t^\alpha - \lambda t}$ for each $\alpha \in \mathfrak{B}(s,x)$, $t \geq 0$ we have

$$f(s,x) = \mathsf{M}^\alpha\left\{\int_0^{t\wedge\tau^\alpha} e^{-\varphi_r^\alpha - \lambda r}[\lambda f(s + r, x_r^{\alpha,x}) - L_r^\alpha f(s + r, x_r^{\alpha,x})]\, dr + f(s + t\wedge\tau^\alpha, x_{t\wedge\tau^\alpha}^{\alpha,x})e^{-\varphi_{t\wedge\tau^\alpha}^\alpha - \lambda t\wedge\tau^\alpha}\right\}, \tag{3}$$

where

$$L_r^\alpha f(t,x) \equiv \frac{\partial}{\partial t} f(t,x) + \sum_{i,j-1}^{d} (a_r^\alpha)^{ij} f_{x^ix^j}(t,x) + \sum_{i-1}^{d} (b_r^\alpha)^i f_{x^i}(t,x) - c_r^\alpha f(t,x).$$

Since a_r^α, b_r^α, c_r^α are bounded, $|L_r^\alpha f(t,x)|$ does not exceed the expression

$$N\left[\left|\frac{\partial}{\partial t} f(t,x)\right| + \sum_{i,j=1}^{d} |f_{x^ix^j}(t,x)| + \sum_{i=1}^{d} |f_{x^i}(t,x)| + |f(t,x)|\right]$$

Denoting the last expression by $h(t,x)$, we note that $h(t,x)$ is a bounded finite function; in particular, $h \in \mathscr{L}_{p+1}(Q)$.

Using the Lebesgue bounded convergence theorem, we pass to the limit in (3) as $t \to \infty$. Thus we have

$$f(s,x) = \lambda \mathsf{M}^\alpha \int_0^{\tau^\alpha} e^{-\varphi_t^\alpha - \lambda t} f(s + t, x_t^{\alpha,x})\, dt - \mathsf{M}^\alpha \int_0^{\tau^\alpha} e^{-\varphi_t^\alpha - \lambda t} L_t^\alpha f(s + t, x_t^{\alpha,x})\, dt,$$

which immediately yields

$$|\lambda z^\lambda(f,s,x) - f(s,x)| = \left|\sup_{\alpha\in\mathfrak{B}(s,x)} \mathsf{M}^\alpha \int_0^{\tau^\alpha} e^{-\varphi_t^\alpha - \lambda t} L_t^\alpha f(s + t, x_t^{\alpha,x})\, dt\right|$$
$$\leq \sup_{\alpha\in\mathfrak{B}(s,x)} \mathsf{M}^\alpha \int_0^{\tau^\alpha} e^{-\varphi_t^\alpha - \lambda t}|L_t^\alpha f(s + t, x_t^{\alpha,x})|\, dt \leq z^\lambda(h,s,x).$$

In short, we have

$$|\lambda z^\lambda(fs,x) - f(s,x)| \le z^\lambda(h,s,x),$$

which, according to Theorem 5a yields

$$\varlimsup_{\lambda\to\infty}]|\lambda z^\lambda(f,\cdot,\cdot) - f(\cdot,\cdot)\|_{p+1,Q} \le \lim_{\lambda\to\infty}]|z^\lambda(h,\cdot,\cdot)\|_{p+1,Q}$$

$$\le N\|h\|_{p+1,Q} \lim_{\lambda\to\infty}\frac{1}{\lambda} = 0,$$

thus proving Theorem 7a for $f_1 \equiv 0$. In the general case

$$|\lambda z^\lambda(f,s,x) - f(s,x)| \le \lambda \sup_{\alpha\in\mathfrak{B}(s,x)} \mathsf{M}^\alpha e^{-\varphi_\tau^\alpha - \lambda\tau^\alpha}|f_1(s+\tau^\alpha, x_\tau^{\alpha,x})|$$

$$+ \left|\lambda \sup_{\alpha\in\mathfrak{B}(s,x)} \mathsf{M}^\alpha \int_0^{\tau^\alpha} e^{-\varphi_t^\alpha - \lambda t} f(s+t, x_t^{\alpha,x})\,dt - f(s,x)\right|,$$

where the exterior norm of the second term tends to zero; due to the boundedness of f_1 the first term does not exceed the product of a constant and the expression

$$\pi^\lambda(s,x) \equiv \lambda \sup_{\alpha\in\mathfrak{B}(s,x)} \mathsf{M}^\alpha e^{-\lambda\tau^{\alpha,s,x}}.$$

Therefore, in order to complete proving Theorem 7a, it remains only to show that $]|\pi^\lambda\|_{p+1,Q'} \to 0$ as $\lambda\to\infty$ for any bounded region Q' lying together with the closure in Q. To this end, it suffices in turn to prove that $\pi^\lambda(s,x) \to 0$ uniformly on Q'. In addition, each region Q' can be covered with a finite number of cylinders of the type $C_{r,R}(s,y) = \{(t,x): |y-x| < R, |t-s| < r\}$, so that $C_{2r,2R}(s,y) \subset Q$. It is seen that we need only to prove that $\pi^\lambda(t,x) \to 0$ uniformly on any cylinder of this type.

We fix a cylinder $C_{r,R}(s,y)$ such that $C_{2r,2R}(s,y) \subset Q$. Let $\tau_R^\alpha(x) = \inf\{t \ge 0: |x - x_t^{\alpha,x}| \ge R\}$. Finally, we denote by $\mu(\lambda)$ the positive root of the equation $\lambda - \mu K - \mu^2 K = 0$. Also, we note that for $(t,x)\in C_{r,R}(s,y)$ we have $\tau^{\alpha,t,x} \ge r \wedge \tau_R^\alpha(x)$. Hence

$$\lambda\mathsf{M}^\alpha e^{-\lambda\tau^{\alpha,t,x}} \le \lambda\mathsf{M}^\alpha e^{-\lambda r\wedge\tau_R^\alpha(x)} \le \lambda e^{-\lambda r} + \lambda\mathsf{M}^\alpha e^{-\lambda\tau_R^\alpha(x)}.$$

Furthermore, by Lemma 3.2[3] the inequality

$$\mathsf{M}^\alpha e^{-\lambda\tau_R^\alpha(x)} \le \mathrm{ch}^{-1}\,\mu R$$

holds true. Therefore, the function $\pi^\lambda(t,x)$ does not exceed $\lambda e^{-\lambda r} + \lambda(\mathrm{ch}\,\mu(\lambda)R)^{-1}$ on $C_{r,R}(s,y)$. Simple computations show that the last constant tends to zero as $\lambda\to\infty$. Therefore, $\pi^\lambda(t,x)$ tends uniformly to zero on $C_{r,R}(s,y)$. This completes the proof of Theorem 7a.

Theorem 7b can be proved in a similar way, which we suggest the reader should do as an exercise. We have thus proved Theorems 6 and 7. □

[3] In Lemma 3.2, one should take $\tau = 0$, $c_t \equiv 1$.

8. Corollary. *Let* $f \in \mathscr{L}_{p+1}(Q)$, $f \geq 0$ (Q-a.s.) *and for all* $(s,x) \in Q$ *let*

$$\inf_{\alpha \in \mathfrak{B}(s,x)} \mathsf{M}^\alpha \int_0^{\tau^\alpha} e^{-\varphi_t^\alpha} f(s+t, x_t^{\alpha,s,x})\, dt = 0. \tag{4}$$

Then $f = 0$ (Q-a.s.).

In fact, by Theorem 2.4 the equality (4) still holds if we change f on the set of measure zero. It is then seen that for $\lambda \geq 0$

$$\inf_{\alpha \in \mathfrak{B}(s,x)} \mathsf{M}^\alpha \int_0^{\tau^\alpha} e^{-\varphi_t^\alpha - \lambda t} f(s+t, x_t^{\alpha,s,x})\, dt = 0.$$

Furthermore, for $f_1 \equiv 0$

$$z^\lambda(-f,s,x) = -\inf_{\alpha \in \mathfrak{B}(s,x)} \mathsf{M}^\alpha \int_0^{\tau^\alpha} e^{-\varphi_t^\alpha - \lambda t} f(s+t, x_t^{\alpha,s,x})\, dt.$$

Therefore, $z^\lambda = 0$ in Q and $-f = \lim_{n\to\infty} \lambda_n z^{\lambda_n} = 0$ (Q-a.s.).

5. Solutions of Stochastic Integral Equations and Estimates of the Moments

In this section, we list some generalizations of the kind we need of well-known results on existence and uniqueness of solutions of stochastic equations. Also, we present estimates of the moments of the above solutions. The moments of these solutions are estimated when the condition for the growth of coefficients to be linear is satisfied. The theorem on existence and uniqueness is proved for the case where the coefficients satisfy the Lipschitz condition (condition ($\mathscr{L}$)).

We fix two constants, $T > 0$, $K > 0$. Also, we adopt the following notation: $(\mathbf{w}_t, \mathscr{F}_t)$ is a d_1-dimensional Wiener process; x, y denote points of E_d; σ_t, $\sigma_t(x)$, $\tilde{\sigma}_t(x)$ are random matrices of dimension $d \times d_1$; $b_t(x)$, $\tilde{b}_t(x)$, ξ_t, $\tilde{\xi}_t$ are random d-dimensional vectors; r_t, h_t are nonnegative numbers. We assume all the processes to be given for $t \in [0,T]$, $x \in E_d$ and progressively measurable with respect to $\{\mathscr{F}_t\}$. If for all $t \in [0,T]$, ω, x, y

$$\|\sigma_t(x) - \sigma_t(y)\| \leq K|x-y|, \qquad |b_t(x) - b_t(y)| \leq K^2|x-y|,$$

we say that the condition ($\mathscr{L}$) is satisfied. If for all $t \in [0,T]$, ω, x

$$\|\sigma_t(x)\|^2 \leq 2r_t^2 + 2K^2|x|^2, \qquad |b_t(x)| \leq h_t + K^2|x|,$$

we say that the condition (R) is satisfied.

Note that we do not impose the condition ($\mathscr{L}$) and the condition (R) on $\tilde{\sigma}_t(x)$, $\tilde{b}_t(x)$. Furthermore, it is useful to have in mind that if the condition ($\mathscr{L}$) is satisfied, the condition (R) will be satisfied for $r_t = \|\sigma_t(0)\|$, $h_t = |b_t(0)|$

(with the same constant K) since, for example, $\|\sigma_t(x)\|^2 \le 2\|\sigma_t(0)\|^2 + 2\|\sigma_t(x) - \sigma_t(0)\|^2$.

As usual, by a solution of the stochastic equation

$$x_t = \xi_t + \int_0^t \sigma_s(x_s)\,dw_s + \int_0^t b_s(x_s)\,ds \tag{1}$$

we mean a progressively measuable (with respect to $\{\mathscr{F}_t\}$), process x_t for which the right side of (1) is defined[4] and, in addition, $x_t(\omega)$ coincides with the right side of (1) for some set Ω' of measure one for all $t \in [0,T]$, $\omega \in \Omega'$.

1. Lemma. *Let x_t be a solution of Eq. (1) for $\xi_t \equiv 0$. Then for $q \ge 1$*

$$\begin{aligned} d|x_t|^{2q} = {} & [2q|x_t|^{2q-2}x_t b_t(x_t) + q|x_t|^{2q-2}\|\sigma_t(x_t)\|^2 \\ & + 2q(q-1)|x_t|^{2q-4}|\sigma_t^*(x_t)x_t|^2]\,dt \\ & + 2q|x_t|^{2q-2}x_t\sigma_t(x_t)\,d\mathbf{w}_t \le q|x_t|^{2q-2}(2|x_t|\,|b_t(x_t)| \\ & + (2q-1)\|\sigma_t(x_t)\|^2)dt + 2q|x_t|^{2q-2}x_t\sigma_t(x_t)d\mathbf{w}_t. \end{aligned}$$

We prove this lemma by applying Ito's formula to the twice continuously differentiable function $|x|^{2q}$ and using the inequalities

$$x_t b_t \le |x_t|\,|b_t|, \qquad |\sigma_t^*(x_t)x_t|^2 \le \|\sigma_t(x_t)\|^2|x_t|^2.$$

2. Lemma. *Let the condition (R) be satisfied and let x_t be a solution of Eq. (1) for $\xi_t \equiv 0$. Then, for all $q \ge 1$, $\varepsilon > 0$, $t \in [0,T]$*

$$(\mathsf{M}|x_t|^{2q})^{1/q} \le \frac{1}{\varepsilon}\int_0^t e^{\lambda(t-s)}[\mathsf{M}h_s^{2q}]^{1/q}\,ds + 2(2q-1)\int_0^t e^{\lambda(t-s)}[\mathsf{M}r_s^{2q}]^{1/q}\,ds, \tag{2}$$

where $\lambda = 4qK^2 + \varepsilon \equiv \lambda_{q,\varepsilon}$. If the condition ($\mathscr{L}$) is satisfied, one can take in (2) $h_s = |b_s(0)|$, $r_s = \|\sigma_s(0)\|$.

PROOF. We fix $q \ge 1$, $\varepsilon > 0$, $t_0 \in [0,T]$. Also, denote by $\psi(t)$ the right side of (2). We prove (2) for $t = t_0$. We can obviously assume that $\psi(t_0) < \infty$. We make one more assumption which we will drop at the end of the proof. Assume that $x_t(\omega)$ is a bounded function of ω, t.

Using the preceding lemma and the condition (R), we obtain

$$\begin{aligned} d|x_t|^{2q} \le {} & [4q^2K^2|x_t|^{2q} + 2q|x_t|^{2q-1}h_t \\ & + 2q(2q-1)|x_t|^{2q-2}r_t^2]\,dt + 2q|x_t|^{2q-2}x_t\sigma_t(x_t)\,d\mathbf{w}_t. \end{aligned}$$

Next, we integrate the last inequality over t. In addition, we take the expectation from the both sides of this inequality. In this case, the expectation of the stochastic integral disappears because, due to the boundedness of

[4] Recall that the stochastic integral in (1) is defined and continuous in t for $t \le T$ if

$$\int_0^T \|\sigma_s(x_s)\|^2\,ds < \infty \quad \text{(a.s.)}.$$

$x_t(\omega)$, finiteness of $\psi(t_0)$, and, in addition, Hölder's inequality,

$$\begin{aligned}\mathsf{M}\int_0^{t_0}|x_t|^{4q-4}|\sigma_t^*(x_t)x_t|^2\,dt &\le N\mathsf{M}\int_0^{t_0}\|\sigma_t(x_t)\|^2\,dt\\ &\le N\mathsf{M}\int_0^{t_0}|x_t|^2\,dt + N\int_0^{t_0}e^{\lambda(t_0-t)}\mathsf{M}r_t^2\,dt\\ &\le N + N\int_0^{t_0}e^{\lambda(t_0-t)}[\mathsf{M}r_t^{2q}]^{1/q}\,dt < \infty.\end{aligned}$$

Furthermore, we use the following inequalities:

$$\begin{aligned}\mathsf{M}|x_t|^{2q-1}h_t &\le (\mathsf{M}|x_t|^{2q})^{1-(1/2q)}(\mathsf{M}h_t^{2q})^{1/2q}\\ &= (\mathsf{M}|x_t|^{2q})^{1/2}[(\mathsf{M}|x_t|^{2q})^{1-(1/q)}(\mathsf{M}h_t^{2q})^{1/q}]^{1/2}\\ &\le \frac{\beta}{2}\mathsf{M}|x_t|^{2q} + \frac{1}{2\beta}(\mathsf{M}|x_t|^{2q})^{1-(1/q)}(\mathsf{M}h_t^{2q})^{1/q},\end{aligned}$$

$$\mathsf{M}|x_t|^{2q-2}r_t^2 \le (\mathsf{M}|x_t|^{2q})^{1-(1/q)}(\mathsf{M}r_t^{2q})^{1/q}.$$

Also, let

$$m(t) = \mathsf{M}|x_t|^{2q}, \qquad g_t = \frac{q}{\varepsilon}(\mathsf{M}h_t^{2q})^{1/q} + 2q(2q-1)(\mathsf{M}r_t^{2q})^{1/q}.$$

In accord with what has been said above for $t \le t_0$,

$$m(t) \le \int_0^t [\lambda q m(s) + g_s m^{1-(1/q)}(s)]\,ds. \tag{3}$$

Further, we apply a well-known method of transforming such inequalities. Let $\delta > 0$. We introduce an operator F_δ on nonnegative functions of one variable, on $[0,t_0]$, by defining

$$F_\delta u(t) = \int_0^t [\lambda q u(s) + g_s u^{1-(1/q)}(s)]\,ds + \delta, \qquad t \in [0,t_0].$$

It is easily seen that F_δ is a monotone operator, i.e., if $0 \le u^1(t) \le u^2(t)$ for all t, then $0 \le F_\delta u^1(t) \le F_\delta u^2(t)$ for all t. Furthermore, if all the nonnegative functions u_t^n are bounded and if they converge for each t, then $\lim_{n\to\infty} F_\delta u_n(t) = F_\delta \lim_{n\to\infty} u_n(t)$. Finally, for the function $v(t) = Ne^{\lambda q t}$ for all sufficiently large N and $\delta \le 1$ we have $F_\delta v(t) \le v(t)$ if $t \in [0,t_0]$. In fact,

$$F_\delta v(t) \le N(e^{\lambda q t} - 1) + N^{1-(1/q)}e^{\lambda q t}\int_0^t g_s e^{-\lambda s}\,ds + \delta \le Ne^{\lambda q t}.$$

for

$$N^{-1/q}e^{\lambda q t_0}\int_0^{t_0} g_s e^{-\lambda s}\,ds + \delta N^{-1} \le 1.$$

It follows from (3) and the aforementioned properties, with N such that $m(t) \le v(t)$, that $m(t) \le F_\delta m(t) \le \cdots \le F_\delta^n m(t) \le v(t)$. Therefore, the limit $\lim_{n\to\infty} F_\delta^n m(t)$ exists. If we denote this limit by $v_\delta(t)$, then $m(t) \le v_\delta(t)$. Taking the limit in the equality $F_\delta^{n+1}m(t) = F_\delta(F_\delta^n m)(t)$, we conclude that $v_\delta = F_\delta v_\delta$. Therefore, for each $\delta \in (0,1)$ the function $m(t)$ does not exceed some nonnegative solution of the equation

$$v_\delta(t) = \int_0^t [\lambda q v_\delta(s) + g_s v_\delta^{1-(1/q)}(s)]\, ds + \delta.$$

We solve the last given equation, from which it follows that $v_\delta(t) \geq \delta$, $v_\delta(0) = \delta$, and

$$v_\delta'(t) = \lambda q v_\delta(t) + g v_\delta^{1-(1/q)}(t). \tag{4}$$

Equation (4), after we have multiplied it by $v_\delta^{(1/q)-1}$ (which is possible due to the inequality $v_\delta \geq \delta$) becomes a linear equation with respect to $v_\delta^{1/q}$. Having solved this equation, we find $v_\delta^{1/q}(t) = \delta^{1/q} + \psi(t)$.

Therefore, $m(t) \leq (\delta^{1/q} + \psi(t))^q$ for all $t \in [0,t_0]$, $\delta \in (0,1)$. We have proved the lemma for the bounded $x_t(\omega)$ as $\delta \to 0$.

In order to prove the lemma in the general case, we denote by τ_R the first exit time of x_t from S_R. Then $x_{t\wedge\tau_R}(\omega)$ is the bounded function of (ω,t), and, as is easily seen,

$$x_{t\wedge\tau_R} = \int_0^t \chi_{s<\tau_R}\sigma_s(x_{s\wedge\tau_R})\, d\mathbf{w}_s + \int_0^t \chi_{s<\tau_R} b_s(x_{s\wedge\tau_R})\, ds.$$

Therefore the process $x_{t\wedge\tau_R}$ satisfies the same equation as the process x_t does; however, $\sigma_s(x)$, $b_s(x)$ are to be replaced by $\chi_{s<\tau_R}\sigma_s(x)$, $\chi_{s<\tau_R}b_s(x)$, respectively. In accord with what has been proved above, $\mathsf{M}|x_{t\wedge\tau_R}|^{2q} \leq [\psi(t)]^q$. It remains only to allow $R \to \infty$, to use Fatou's lemma, and in addition, to take advantage of the fact that due to the continuity of x_t the time $\tau_R \to \infty$ as $R \to \infty$. We have thus proved our lemma. □

3. Corollary. *Let $\int_0^t \|\sigma_s\|^2\, ds < \infty$ with probability 1, and let τ be a Markov time with respect to $\{\mathscr{F}_s\}$. Then, for all $q \geq 1$*

$$\mathsf{M}\left|\int_0^{\tau\wedge t} \sigma_s\, d\mathbf{w}_s\right|^{2q} \leq 2^q(2q-1)^q\left\{\int_0^t [\mathsf{M}\|\sigma_s\|^{2q}\chi_{s<\tau}]^{1/q}\, ds\right\}^q$$
$$\leq 2^q(2q-1)^q t^{q-1}\mathsf{M}\int_0^{\tau\wedge t}\|\sigma_s\|^{2q}\, ds.$$

In fact, we have obtained the second inequality using Hölder's equality. The first inequality follows from the lemma, if we take $\sigma_s(x) = \sigma_s\chi_{s<\tau}$, $b_s(x) = 0$, write the assertion of the lemma with arbitrary K, ε, and, finally, assume that $K \downarrow 0$, $\varepsilon \downarrow 0$.

4. Exercise

In the proof of the lemma, show that the factor 2^q in Corollary 3 can be replaced by unity.

5. Corollary. *Let the condition $(\mathscr{L})$ be satisfied, let x_t be a solution of Eq. (1), and let $\tilde{x}_t$ be a solution of the equation*

$$\tilde{x}_t = \tilde{\xi}_t + \int_0^t \tilde{\sigma}_s(\tilde{x}_s)\, d\mathbf{w}_s + \int_0^t \tilde{b}_s(\tilde{x}_s)\, ds.$$

Then, for all $q \geq 1$, $t \in [0,T]$

$$\mathsf{M}|x_t - \tilde{x}_t|^{2q} \leq 4^q \mathsf{M}|\xi_t - \tilde{\xi}_t|^{2q} + N(q,K)t^{q-1}\mathsf{M}\int_0^t e^{\mu(t-s)}|\xi_s - \tilde{\xi}_s|^{2q}\,ds$$
$$+ N(q)t^{q-1}\mathsf{M}\int_0^t e^{\mu(t-s)}\{|b_s(\tilde{x}_s) - \tilde{b}_s(\tilde{x}_s)|^{2q}$$
$$+ \|\sigma_s(\tilde{x}_s) - \tilde{\sigma}_s(\tilde{x}_s)\|^{2q}\}\,ds,$$

where $\mu = 4q^2K^2 + q$.

PROOF. Let $y_t = (x_t - \tilde{x}_t) - (\xi_t - \tilde{\xi}_t)$. Then, as is easily seen,

$$y_t = \int_0^t [\sigma_s(y_s + \tilde{x}_s + \xi_s - \tilde{\xi}_s) - \tilde{\sigma}_s(\tilde{x}_s)]\,d\mathbf{w}_s + \int_0^t [b_s(y_s + \tilde{x}_s + \xi_s - \tilde{\xi}_s) - \tilde{b}_s(\tilde{x}_s)]\,ds,$$

in this case

$$[\sigma_s(x + \tilde{x}_s + \xi_s - \tilde{\xi}_s) - \tilde{\sigma}_s(\tilde{x}_s)], \qquad [b_s(x + \tilde{x}_s + \xi_s - \tilde{\xi}_s) - \tilde{b}_s(\tilde{x}_s)]$$

satisfy the condition ($\mathscr{L}$). From this, according to the lemma applied to the process y_t, we have

$$(\mathsf{M}|y_t|^{2q})^{1/q} \leq \int_0^t e^{(1/q)\mu(t-s)}[\mathsf{M}|b_s(\tilde{x}_s + \xi_s - \tilde{\xi}_s) - \tilde{b}_s(\tilde{x}_s)|^{2q}]^{1/q}\,ds$$
$$+ 2(2q-1)\int_0^t e^{(1/q)\mu(t-s)}[\mathsf{M}\|\sigma_s(\tilde{x}_s + \xi_s - \tilde{\xi}_s) - \tilde{\sigma}_s(\tilde{x}_s)\|^{2q}]^{1/q}\,ds.$$

We raise both sides of the last inequality to the q^{th} power. We use Hölder's inequality as well as the fact that

$$|b_s(\tilde{x}_s + \xi_s - \tilde{\xi}_s) - \tilde{b}_s(\tilde{x}_s)| \leq |b_s(\tilde{x}_s + \xi_s - \tilde{\xi}_s) - b_s(\tilde{x}_s)| + |b_s(\tilde{x}_s) - \tilde{b}_s(\tilde{x}_s)|$$
$$\leq K^2|\xi_s - \tilde{\xi}_s| + |b_s(\tilde{x}_s) - \tilde{b}_s(\tilde{x}_s)|,$$
$$(a+b)^q \leq 2^{q-1}(a^q + b^q),$$

which yields

$$\mathsf{M}|y_t|^{2q} \leq 2^{q-1}t^{q-1}\mathsf{M}\int_0^t e^{\mu(t-s)}[2^{2q-1}K^{4q}|\xi_s - \tilde{\xi}_s|^{2q}$$
$$+ 2^{2q-1}|b_s(\tilde{x}_s) - \tilde{b}_s(\tilde{x}_s)|^{2q} + 2^q(2q-1)^q 2^{2q-1}K^{2q}|\xi_s - \tilde{\xi}_s|^{2q}$$
$$+ 2^q(2q-1)^q 2^{2q-1}\|\sigma_s(\tilde{x}_s) - \tilde{\sigma}_s(\tilde{x}_s)\|^{2q}]\,ds.$$

It remains to note that $|x_t - \tilde{x}_t| \leq |y_t| + |\xi_t - \tilde{\xi}_t|$, $|x_t - \tilde{x}_t|^{2q} \leq 2^{2q-1}|y_t|^{2q} + 2^{2q-1}|\xi_t - \tilde{\xi}_t|^{2q}$, thus proving Corollary 5. □

6. Corollary. *Let the condition* (*R*) *be satisfied, and let* x_t *be a solution of* (1). *Then there exists a constant* $N = N(q,K)$ *such that for all* $q \geq 1$, $t \in [0,T]$

$$\mathsf{M}|x_t|^{2q} \leq N\mathsf{M}|\xi_t|^{2q} + Nt^{q-1}\mathsf{M}\int_0^t [|\xi_s|^{2q} + h_s^{2q} + r_s^{2q}]e^{N(t-s)}\,ds.$$

In fact, the process $y_t \equiv x_t - \xi_t$ satisfies the equation

$$dy_t = \sigma(y_t + \xi_t)\,d\mathbf{w}_t + b(y_t + \xi_t)\,dt, \qquad y_0 = 0,$$

the coefficients of this equation satisfying the condition (R), however with different h_t, r_t, K. For example,

$$\|\sigma_t(x + \xi_t)\|^2 \leq 2r_t^2 + 2K^2\|x + \xi_t\|^2 \leq 2r_t^2 + 4K^2|\xi_t|^2 + 4K^2|x|.$$

Therefore, using this lemma we can estimate $\mathsf{M}|y_t|^{2q}$. Having done this, we need to use the fact that $|x_t|^{2q} \leq 2^{2q-1}|y_t| + 2^{2q-1}|\xi_t|$.

In our previous assertions we assumed that a solution of Eq. (1) existed and we also wrote the inequalities which may sometimes take the form $\infty \leq \infty$. Further, it is convenient to prove one of the versions of the classical Ito theorem on the existence of a solution of a stochastic equation. Since the proofs of these theorems are well known, we shall dwell here only on the most essential points.

7. Theorem. *Let the condition* ($\mathscr{L}$) *be satisfied and let*

$$\mathsf{M} \int_0^T [|\xi_t|^2 + |b_t(0)|^2 + \|\sigma_t(0)\|^2]\, dt < \infty.$$

Then for $t \leq T$ *Eq.* (1) *has a solution such that* $\mathsf{M} \int_0^T |x_t|^2\, dt < \infty$. *If* x_t, y_t *are two solutions of* (1), *then* $\mathsf{P}\{\sup_{t \in [0,T]}|x_t - y_t| > 0\} = 0$.

PROOF. Due to Corollary 5, $\mathsf{M}|x_t - y_t|^2 = 0$ for each t. Furthermore, the process $x_t - y_t$ can be represented as the sum of stochastic integrals and ordinary integrals. Hence the process $x_t - y_t$ is continuous almost surely. The equality $x_t = y_t$ (a.s.) for each t implies that $x_t = y_t$ for all t (a.s.), thus proving the last assertion of the theorem.

For proving the first assertion of the theorem we apply, as is usually done in similar cases, the method of successive approximation. We define the operator I using the formula

$$Ix_t = \int_0^t \sigma_s(x_s)\, d\mathbf{w}_s + \int_0^t b_s(x_s)\, ds. \tag{5}$$

This operator is defined on those processes x_t for which the right side of (5) makes sense, and, furthermore, this operator maps these processes into processes Ix_t whose values can be found with the aid of the formula (5).

Denote by V a space of progressively measurable processes x_t with values in E_d such that

$$\|x_t\| = \left(\mathsf{M} \int_0^T |x_t|^2\, dt\right)^{1/2} < \infty.$$

It can easily be shown that the operator I maps V into V. In addition, it can easily be deduced from the condition ($\mathscr{L}$) that

$$\mathsf{M}|Ix_t - Iy_t|^2 \leq \alpha \mathsf{M} \int_0^t |x_s - y_s|^2\, ds, \tag{6}$$

where $\alpha = 2K^2(1 + TK^2)$.

Let $x_t^{(0)} \equiv 0$, $x_t^{(n+1)} = \xi_t + Ix_t^{(n)}$ ($n = 0,1,2,\ldots$). It follows from (6) that

$$\mathsf{M}|x_t^{(n+1)} - x_t^{(n)}|^2 \leq \alpha \mathsf{M} \int_0^t |x_s^{(n)} - x_s^{(n-1)}|^2\, ds.$$

Iterating the last inequality, we find

$$\|x_t^{(n+1)} - x_t^{(n)}\|^2 \leq \frac{T^n\alpha^n}{n!}\|x_t^{(1)}\|^2. \tag{7}$$

Since a series of the numbers $(T\alpha)^{n/2}(n!)^{-n/2}$ converges, it follows from (7) that a series of functions $x_t^{(n+1)} - x_t^{(n)}$ converges in V. In other words, the functions $x_t^{(n+1)}$ converge in V, and furthermore, there exists a process $\tilde{x}_t \in V$ such that $\|x_t^{(n)} - \tilde{x}_t\| \to 0$ as $n \to \infty$.

Further, integrating (6), we obtain

$$\|Ix_t - Iy_t\| \leq \alpha T\|x_t - y_t\|. \tag{8}$$

In particular, the operator I is continuous in V. Passing to the limit in the equality $\|x_t^{(n+1)} - (\xi_t + Ix_t^{(n)})\| = 0$, we conclude $\|\tilde{x}_t - (\xi_t + I\tilde{x}_t)\| = 0$, from which and also from (8) it follows that $I\tilde{x}_t = I(\xi + I\tilde{x})_t$ for almost all t, ω. However, the both sides of this equality are continuous with respect to t for almost all ω. Hence they coincide for all t at once almost surely. Finally, taking $x_t = \xi_t + I\tilde{x}_t$, we have $x_t = \xi_t + I(\xi + I\tilde{x})_t = \xi_t + Ix_t$ for all t almost surely. Therefore, x_t is a solution of the primary equation, (1), thus completing the proof of the theorem. □

8. Exercise

Noting that $\sigma_s(x) = [\sigma_s(x) - \sigma_s(0)] + \sigma_s(0)$, prove that the assertions of the theorem still hold if $\mathsf{M}\int_0^T |\eta_t|^2\,dt < \infty$, where

$$\eta_t = \xi_t + \int_0^t \sigma_s(0)\,d\mathbf{w}_s + \int_0^t b_s(0)\,ds.$$

We continue estimating the moments of solutions of a stochastic equation.

9. Theorem. *Suppose the condition* ($\mathscr{L}$) *is satisfied,* x_t *is a solution of Eq.* (1), *and* $\tilde{x}_t$ *is a solution of the equation*

$$\tilde{x}_t = \tilde{\xi}_t + \int_0^t \tilde{\sigma}_s(\tilde{x}_s)\,d\mathbf{w}_s + \int_0^t \tilde{b}_s(\tilde{x}_s)\,ds.$$

Then, if the process $\xi_t - \tilde{\xi}_t$ *is separable, the process* $x_t - \tilde{x}_t$ *is also separable, and for all* $q \geq 1$, $t \in [0,T]$

$$\mathsf{M}\sup_{s\leq t}|x_s - \tilde{x}_s|^{2q} \leq Ne^{Nt}\mathsf{M}\sup_{s\leq t}|\xi_s - \tilde{\xi}_s|^{2q}$$
$$+ Nt^{q-1}e^{Nt}\mathsf{M}\int_0^t [|b_s(\tilde{x}_s) - \tilde{b}_s(\tilde{x}_s)|^{2q} + \|\sigma_s(\tilde{x}_s) - \tilde{\sigma}_s(\tilde{x}_s)\|^{2q}]\,ds,$$

where $N = N(q,K)$.

Proof. It is seen that $x_t - \tilde{x}_t$ is the sum of $\xi_t - \tilde{\xi}_t$, stochastic integrals and Lebesque integrals. Both types of integrals are continuous with respect to t. Hence, the separability property of $\xi_t - \tilde{\xi}_t$ implies that $x_t - \tilde{x}_t$ is separable, and in particular, the quantity $\sup_{s\leq t}|x_s - \tilde{x}_s|$ is measurable with respect to ω.

As was done in proving Corollary 5, the assertion of the theorem in the general case can easily be deduced from that in the case where $\xi_t = \tilde{\xi}_t = \tilde{x}_t = 0$, $\tilde{\sigma}_s(x) = 0$, $\tilde{b}_s(x) = 0$. It is required to prove in the latter case that

$$\mathsf{M} \sup_{s \le t} |x_s|^{2q} \le N t^{q-1} e^{Nt} \mathsf{M} \int_0^t [|b_s(0)|^{2q} + \|\sigma_s(0)\|^{2q}]\, ds. \tag{9}$$

Reasoning in the same way as in proving Lemma 2, we convince ourselves that it is possible to consider only the case with bounded functions $x_t(\omega)$ and to assume, in addition, that the right side of (9) is finite.

First, we prove that the process

$$\eta_t = |x_t| e^{K^2 t} + \int_0^t e^{K^2 s} |b_s(0)|\, ds$$

is a submartingale. We fix $\varepsilon > 0$ and introduce an auxiliary function of the real variable r using the formula $\varphi(r) = \sqrt{r^2 + \varepsilon^2}$. Note that $\varphi(|x|)$ is a smooth function on E_d. In conjunction with Ito's formula,

$$\begin{aligned} d[\varphi(|x_t|)e^{K^2 t}] = e^{K^2 t} \Bigg\{ & K^2 \varphi(|x_t|) + \varphi'(|x_t|) \frac{b_t(x_t)x_t}{|x_t|} \\ & + \varphi''(|x_t|) \frac{1}{2} \frac{|\sigma_t^*(x_t)x_t|^2}{|x_t|^2} + \varphi'(|x_t|) \frac{1}{2} \frac{1}{|x_t|} \times \\ & \left[\|\sigma_t(x_t)\|^2 - \frac{|\sigma_t^*(x_t)x_t|^2}{|x_t|^2} \right] \Bigg\} dt + e^{K^2 t} \varphi'(|x_t|) \frac{x_t \sigma_t(x_t)}{|x_t|}\, d\mathbf{w}_t. \end{aligned}$$

Let us integrate the last expression over t from s_1 up to $s_2 \ge s_1$, and also, let us take the conditional expectation under the condition $\mathscr{F}_{s_1}$. In this case, the expectation of the stochastic integral disappears (see Proof of Lemma 2). In addition, we use the fact that since

$$b_t(x_t)x_t \ge -|b_t(x_t)|\,|x_t| \ge -K^2|x_t|^2 - |b_t(0)|\,|x_t|, \qquad 0 \le \varphi'(r) \le 1,\ |r| \le \varphi(r),$$

then

$$K^2 \varphi(|x_t|) + \varphi'(|x_t|) \frac{b_t(x_t)x_t}{|x_t|} \ge -|b_t(0)|.$$

Furthermore, $\varphi'' \ge 0$, $|x_t|^2 \|\sigma_t(x_t)\|^2 \ge |\sigma_t^*(x_t)x_t|^2$. Therefore,

$$\mathsf{M}\{\varphi(|x_{s_2}|)e^{K^2 s_2} | \mathscr{F}_{s_1}\} - \varphi(|x_{s_1}|)e^{K^2 s_1} \ge \mathsf{M}\left\{ \int_{s_1}^{s_2} e^{K^2 t} |b_t(0)|\, dt \,\Big|\, \mathscr{F}_{s_1} \right\}.$$

from which, letting ε go to zero, we obtain, using the theorem on bounded convergence, $\mathsf{M}\{\eta_{s_2} | \mathscr{F}_{s_1}\} \ge \eta_{s_1}$. Therefore η_t is a submartingale.

From well-known inequalities for submartingales (see Appendix 2) as well as Hölder's inequality we have

$$\begin{aligned} \mathsf{M} \sup_{s \le t} |x_s|^{2q} &\le \mathsf{M} \sup_{s \le t} \eta_s^{2q} \le 4\mathsf{M}\eta_t^{2q} \\ &\le 4 \cdot 2^{2q-1} e^{2qK^2 t} \mathsf{M}|x_t|^{2q} + 4 \cdot 2^{2q-1} e^{2qK^2 t} t^{2q-1} \mathsf{M} \int_0^t |b_s(0)|^{2q}\, ds. \end{aligned}$$

It remains only to use Lemma 2 or Corollary 6 for estimating $\mathsf{M}|x_t|^{2q}$, and, furthermore, to note that $t^a e^{bt} \le N(a,b)e^{2bt}$ for $a > 0$, $b > 0$, $t > 0$. The theorem is proved. □

10. Corollary. *Let the condition (R) be satisfied, and let x_t be a solution of Eq. (1). Then there exists a constant $N(q,K)$ such that for all $q \ge 1$, $t \in [0,T]$*

$$\mathsf{M} \sup_{s \le t} |x_s - \xi_s|^{2q} \le N t^{q-1} e^{Nt} \mathsf{M} \int_0^t [|\xi_s|^{2q} + h_s^{2q} + r_s^{2q}]\, ds.$$

If ξ_t is a separable process, then

$$\mathsf{M} \sup_{s \le t} |x_s|^{2q} \le N \mathsf{M} \sup_{s \le t} |\xi_t|^{2q} + N t^{q-1} e^{Nt} \mathsf{M} \int_0^t [|\xi_s|^{2q} + h_s^{2q} + r_s^{2q}]\, ds.$$

First, we note that the second inequality follows readily from the first expression. In order to prove the first inequality, we introduce the process $y_t = x_t - \xi_t$. It is seen that $dy_t = \sigma_t(y_t + \xi_t)\, d\mathbf{w}_t + b_t(y_t + \xi_t)\, dt$, $y_0 = 0$. In estimating y_t it suffices, as was done in proving Lemma 2, to consider only the case where $y_t(\omega)$ is a bounded function. Similarly to what we did in proving our theorem above, we use here the inequality $b_t(y_t + \xi_t)y_t \ge -K^2|y_t|^2 - (K^2|\xi_t| + h_t)|y_t|$, thus obtaining that the process

$$\eta_t = |y_t| e^{K^2 t} + \int_0^t e^{K^2 s}(K^2|\xi_s| + h_s)\, ds$$

is a submartingale.

From the above, using the inequalities for submartingales as well as Hölder's inequality, we find

$$\begin{aligned} \mathsf{M} \sup_{s \le t} |y_t|^{2q} &\le \mathsf{M} \sup_{s \le t} \eta_s^{2q} \le 4\mathsf{M}\eta_t^{2q} \\ &\le N e^{Nt} \mathsf{M}|y_t|^{2q} + N e^{Nt} t^{2q-1} \mathsf{M} \int_0^t (|\xi_s|^{2q} + h_s^{2q})\, ds. \end{aligned}$$

For estimating $\mathsf{M}|y_t|^{2q}$, it remains to apply Lemma 2, noting that $\sigma_t(x + \xi_t)$, $b_t(x + \xi_t)$ satisfy the condition (R) in which we replace r_t^2, h_t, K by $r_t^2 + 2K^2|\xi_t|^2$, $h_t + K^2|\xi_t|$, $2K$, respectively.

11. Corollary. *Let $\int_0^t \|\sigma_s\|^2\, ds < \infty$ (a.s.). Then for all $q \ge 1$*

$$\mathsf{M} \sup_{s \le t} \left| \int_0^s \sigma\, d\mathbf{w}_s \right|^{2q} \le 2^{q+2}(2q-1)^q t^{q-1} \mathsf{M} \int_0^t \|\sigma_s\|^{2q}\, ds.$$

This corollary as well as Corollary 10 can be proved by arguments similar to those used to prove the theorem. Taking $\sigma_s(x) = \sigma_s$, $b_s(x) = 0$, we have the process $x_t = \int_0^t \sigma_s\, d\mathbf{w}_s$. The proof of the theorem for $K = 0$ shows that $|x_t|$ is a submartingale. Hence $\mathsf{M}\, \sup_{s \ge t} |x_s|^{2q} \le 4\mathsf{M}|x_t|^{2q}$, which can be estimated with the aid of Corollary 3.

12. Corollary. *Let there exist a constant K_1 such that $\|\sigma_t(x)\| + |b_t(x)| \leq K_1(1 + |x|)$ for all t, ω, x. Let x_t be a solution of Eq. (1) for $\xi_t \equiv x_0$, where x_0 is a fixed point on E_d. There exists a constant $N(q,K_1)$ such that for all $q \geq 0$, $t \in [0,T]$*

$$\mathsf{M} \sup_{s \leq t} |x_s - x_0|^q \leq N t^{q/2} e^{Nt} (1 + |x_0|)^q,$$

$$\mathsf{M} \sup_{s \leq t} |x_s|^q \leq N e^{Nt} (1 + |x_0|)^q.$$

In fact, for $q \geq 2$ these estimates are particular cases of the estimates given in Corollary 10. To prove these inequalities for $q \in [0,2]$ we need only to take $\eta_1 = \sup_{s \leq t} |x_s - x_0|(1 + |x_0|)^{-1}$, $\eta_2 = \sup_{s \leq t} |x_s|(1 + |x_0|)^{-1}$ and, furthermore, use the fact that in conjunction with Hölder's inequality, $\mathsf{M}|\eta_t|^q \leq (\mathsf{M}|\eta_t|^2)^{q/2}$.

13. Remark. The sequential approximations x_t^n defined in proving Theorem 7 have the property that

$$\lim_{n \to \infty} \mathsf{M} \sup_{t \leq T} |x_t^n - x_t|^2 = 0,$$

where x_t is a solution of Eq. (1). Indeed,

$$x_t^{n+1} = \xi_t + I x_t^n, \qquad x_t = \xi_t + I \tilde{x}_t, \qquad x_t^{n+1} - x_t = I x_t^n - I \tilde{x}_t,$$

from which, using Corollary 11 and the Cauchy inequality, we obtain

$$\begin{aligned} \mathsf{M} \sup_{t \leq T} |x_t^{n+1} - x_t|^2 &\leq 2\mathsf{M} \sup_{t \leq T} \left| \int_0^t [\sigma_s(x_s^n) - \sigma_s(\tilde{x}_s)] \, d\mathbf{w}_s \right|^2 \\ &\quad + 2\mathsf{M} \sup_{t \leq T} \left| \int_0^t [b_s(x_s^n) - b_s(\tilde{x}_s)] \, ds \right|^2 \\ &\leq N\mathsf{M} \int_0^T \|\sigma_s(x_s^n) - \sigma_s(\tilde{x}_s)\|^2 \, ds \\ &\quad + N\mathsf{M} \int_0^T |b_s(x_s^n) - b_s(\tilde{x}_s)|^2 \, ds \leq N\mathsf{M} \int_0^T |x_s^n - \tilde{x}_s|^2 \, ds. \end{aligned}$$

As was seen in the proof of Theorem 7, the expression given above tends to zero as $n \to \infty$.

6. Existence of a Solution of a Stochastic Equation with Measurable Coefficients

In this section, using the estimates obtained in Sections 2.2–2.5 we prove that in a wide class of cases there exists a probability space and a Wiener process on this space such that a stochastic equation having measurable coefficients as well as this Wiener process is solvable. In other words, ac-

acording to conventional terminology, we construct here "weak" solutions of a stochastic equation. The main difference between "weak" solutions and usual ("strong") solutions consists in the fact that the latter can be constructed on any a priori given probability space on the basis of any given Wiener process.

Let $\sigma(t,x)$ be a matrix of dimension $d \times d$, and let $b(t,x)$ be a d-dimensional vector. We assume that $\sigma(t,x)$, $b(t,x)$ are given for $t \geq 0$, $x \in E_d$, and, in addition, are bounded and Borel measurable with respect to (t,x). Also, let the matrix $\sigma(t,x)$ be positive definite, and, moreover, let

$$(\sigma(t,x)\lambda,\lambda) \geq \delta|\lambda|^2$$

for some constant $\delta > 0$ for all (t,x), $\lambda \in E_d$.

1. Theorem. *Let $x \in E_d$. There exists a probability space, a Wiener process $(\mathbf{w}_t,\mathscr{F}_t)$ on this space, and a continuous process x_t which is progressively measurable with respect to $\{\mathscr{F}_t\}$, such that almost surely for all $t \geq 0$*

$$x_t = x + \int_0^t \sigma(s,x_s)\,d\mathbf{w}_s + \int_0^t b(s,x_s)\,ds.$$

For proving our theorem we need two assertions due to A. V. Skorokhod.

2. Lemma.[5] *Suppose that d_1-dimensional random processes ξ_t^n ($t \geq 0$, $n = 0, 1,2,\ldots$) are defined on some probability space. Assume that for each $T \geq 0$, $\varepsilon > 0$*

$$\lim_{c\to\infty} \sup_n \sup_{t\leq T} \mathsf{P}\{|\xi_t^n| > c\} = 0,$$

$$\lim_{h\downarrow 0} \sup_n \sup_{\substack{t_1,t_2\leq T\\|t_1-t_2|\leq h}} \mathsf{P}\{|\xi_{t_1}^n - \xi_{t_2}^n| > \varepsilon\} = 0.$$

Then, one can choose a sequence of numbers n', a probability space, and random processes $\tilde{\xi}_t$, $\tilde{\xi}_t^{n'}$ defined on this probability space such that all finite-dimensional distributions of $\tilde{\xi}_t^{n'}$ coincide with the pertinent finite-dimensional distributions of $\xi_t^{n'}$ and $\mathsf{P}\{|\tilde{\xi}_t^{n'} - \tilde{\xi}_t| > \varepsilon\} \to 0$ as $n' \to \infty$ for all $\varepsilon > 0$, $t \geq 0$.

3. Lemma.[6] *Suppose the assumptions of Lemma 2 are satisfied. Also, suppose that d_1-dimensional Wiener processes $(\mathbf{w}_t^n,\mathscr{F}_t^n)$ are defined on the aforegoing probability space. Assume that the functions $\xi_t^n(\omega)$ are bounded on $[0,\infty) \times \Omega$ uniformly in n and that the stochastic integrals $I_t^n = \int_0^t \xi_s^n\, d\mathbf{w}_s^n$ are defined. Finally, let $\xi_s^n \to \xi_s^0$, $\mathbf{w}_s^n \to \mathbf{w}_s^0$ in probability as $n \to \infty$ for each $s \geq 0$. Then $I_t^n \to I_t^0$ as $n \to \infty$ in probability for each $t \geq 0$.*

4. Proof of Theorem 1. We smooth out σ, b using the convolution. Let $\sigma_n(t,x) = \sigma^{(\varepsilon_n)}(t,x)$, $b_n(t,x) = b^{(\varepsilon_n)}(t,x)$[7] (see Section 2.1), where $\varepsilon_n \to 0$ as $n \to \infty$,

[5] See [70, Chapter 1, §6].

[6] See [70, Chapter 2, §6].

[7] In computing the convolution we assume that $\sigma^{ij}(t,x) = \delta\delta^{ij}$, $b^i(t,x) = 0$ for $t \leq 0$.

$\varepsilon_n \neq 0$. It is clear that σ_n, b_n are bounded, $\sigma_n \to \sigma$, $b_n \to b$ (a.s.) as $n \to \infty$,

$$(\sigma_n\lambda,\lambda) = (\sigma\lambda,\lambda)^{(\varepsilon_n)} \geq \delta|\lambda|^2$$

for all $\lambda \in E_d$, $n \geq 1$. Let $\sigma_0 = \sigma$, $b_0 = b$.

We take some d-dimensional Wiener process $(\mathbf{w}_t,\mathscr{F}_t)$. Furthermore, we consider for $n = 1, 2, \ldots$ solutions of the following stochastic equations: $dx_t^n = \sigma_n(t,x_t^n)\,d\mathbf{w}_t + b_n(t,x_t^n)\,dt$, $t \geq 0$, $x_0^n = x$. Note that the derivatives σ^n, b^n are bounded for each n. Hence the functions of σ^n, b^n satisfy the Lipschitz condition and the solutions of the aforegoing equations in fact exist.

According to Corollary 5.12, for each T

$$\sup_n \mathsf{M} \sup_{t\leq T} |x_t^n| < \infty.$$

Using Chebyshev's inequality we then obtain

$$\lim_{c\to\infty} \sup_n \sup_{t\leq T} \mathsf{P}\{|x_t^n| > c\} = 0.$$

Further, for $t_2 > t_1$

$$x_{t_2}^n - x_{t_1}^n = \int_{t_1}^{t_2} \sigma_n(s,x_s^n)\,d\mathbf{w}_s + \int_{t_1}^{t_2} b_n(s,x_s^n)\,ds,$$

from which, according to Corollary 5.3[8] for $t_2 - t_1 < 1$ we have

$$\begin{aligned}\mathsf{M}|x_{t_2}^n - x_{t_1}^n|^4 &\leq N\left\{\int_{t_1}^{t_2} [\mathsf{M}\|\sigma_n(s,x_s^n)\|^4]^{1/2}\,ds\right\}^2 + N\mathsf{M}\left|\int_{t_1}^{t_2} b_n(s,x_s^n)\,ds\right|^4 \\ &\leq N(t_2 - t_1)^2 + N(t_2 - t_1)^4 \leq N(t_2 - t_1)^2,\end{aligned}$$

where the constants N depend only on the upper bounds $\|\sigma\|$, $|b|$, and do not depend on n. In conjunction with Chebyshev's inequality

$$\lim_{h\downarrow 0} \sup_n \sup_{|t_1 - t_2| < h} \mathsf{P}\{|x_{t_2}^n - x_{t_1}^n| > \varepsilon\} = 0. \tag{1}$$

Using Lemma 2 we conclude that there exists a sequence of numbers n', a probability space, and random processes $(\tilde{x}_t^{n'};\tilde{\mathbf{w}}_t^{n'})$ on this probability space such that the finite-dimensional distributions of $(\tilde{x}_t^{n'},\tilde{\mathbf{w}}_t^{n'})$ coincide with the corresponding finite-dimensional distributions of the processes $(x_t^{n'};\mathbf{w}_t)$, and for all $t \geq 0$ the limit, say $(\tilde{x}_t^0;\tilde{\mathbf{w}}_t^0)$, exists in probability of the sequence $(\tilde{x}_t^{n'};\tilde{\mathbf{w}}_t^{n'})$ as $n' \to \infty$. For brevity of notation we assume that the sequence $\{n'\}$ coincides with $\{1,2,3,\ldots\}$.

The processes $(\tilde{x}_t^n;\tilde{\mathbf{w}}_t^n)$ can be regarded as separable processes for all $n \geq 0$. Since $\mathsf{M}|\tilde{x}_{t_2}^n - \tilde{x}_{t_1}^n|^4 = \mathsf{M}|x_{t_2}^n - x_{t_1}^n|^4 \leq N|t_2 - t_1|^2$ for $n > 0$, $|t_2 - t_1| \leq 1$ (by Fatou's lemma), the relationship between the extreme terms of this inequality holds for $n = 0$ as well. Then, by Kolmogorov's theorem $\tilde{x}_t^n$ is a continuous process for all $n \geq 0$. $\tilde{\mathbf{w}}_t^n$, being separable Wiener processes, are continuous as well.

[8] In Corollary 5.3 we need take $\tau = \infty$, $t = t_2$, $\sigma_s = \sigma_n(s.\,x_s^n)\chi_{s>t_1}$.

Further, we fix some $T > 0$. The processes $(x_t^n; \mathbf{w}_t)$ are measurable with respect to $\mathscr{F}_T$ for $t \le T$; the increments $\mathbf{w}_s$ after an instant of time T do not depend on $\mathscr{F}_T$. Therefore, the processes $(x_t^n; \mathbf{w}_t)$ $(t \le T)$ do not depend on the increments $\mathbf{w}_s$ after the instant of time T. Due to the coincidence of finite-dimensional distributions, the processes $(\tilde{x}_t^n; \tilde{\mathbf{w}}_t^n)$ $(t \le T)$ do not depend on the increments $\tilde{\mathbf{w}}_s^n$ after the time T for $n \ge 1$. This property obviously holds true for a limiting process as well, i.e., it holds for $n = 0$. This readily implies that for $n \ge 0$ the processes $\tilde{\mathbf{w}}_t^n$ are Wiener processes with respect to σ-algebras of $\mathscr{F}_t^{(n)}$, defined as the completion of $\sigma\{\tilde{x}_s^n, \tilde{\mathbf{w}}_s^n : s \le t\}$. Furthermore, for $n \ge 0$ and each $s \le t$ the variable $\tilde{x}_s^n$ is $\mathscr{F}_t^{(n)}$-measurable. Since $\tilde{x}_s^n$ is continuous with respect to s, $\tilde{x}_s^n$ is a progressively measurable process with respect to $\{\mathscr{F}_t^{(n)}\}$. These arguments show that the stochastic integrals given below make sense.

Let $\kappa_m(a) = 2^{-m}[2^m a]$, where $[a]$ is the largest integer $\le a$. Since $\sigma_n(t, \tilde{x}_t^n)$ for $n \ge 1$ are bounded functions of (ω, t), continuous with respect to t, and since $\kappa_m(t) \to t$ as $m \to \infty$, then

$$\lim_{m \to \infty} \mathsf{M} \int_0^T \|\sigma_n(t, \tilde{x}_t^n) - \sigma_n(\kappa_m(t), \tilde{x}_{\kappa_m(t)}^n)\|^2 \, dt \to 0$$

for $n \ge 1$ for each $T \ge 0$. Hence for each $t \ge 0$

$$\int_0^t \sigma_n(s, \tilde{x}_s^n) \, d\tilde{\mathbf{w}}_s^n = \operatorname*{l.i.m.}_{m \to \infty} \int_0^t \sigma_n(\kappa_m(s), x_{\kappa_m(s)}^n) \, d\tilde{\mathbf{w}}_s^n$$

$$= \operatorname*{l.i.m.}_{m \to \infty} \sum_{k2^{-m} \le t} \sigma_n(k2^{-m}, \tilde{x}_{k2^{-m}}^n)(\tilde{\mathbf{w}}_{(k+1)2^{-m}}^n - \tilde{\mathbf{w}}_{k2^{-m}}^n).$$

Writing similar relations for $\int_0^t \sigma_n(s, x_s^n) \, d\mathbf{w}_s$, $\int_0^t b_n(s, \tilde{x}_s^n) \, ds$, $\int_0^t b_n(s, x_s^n) \, ds$, and using the fact that the familiar finite-dimensional distributions coincide, we can easily prove that for all $n \ge 1$, $t \ge 0$

$$\mathsf{M} \left| \tilde{x}_t^n - x - \int_0^t \sigma_n(s, \tilde{x}_s^n) \, d\tilde{\mathbf{w}}_s^n - \int_0^t b_n(s, \tilde{x}_s^n) \, ds \right|^2 = 0.$$

In other words,

$$\tilde{x}_t^n = x + \int_0^t \sigma_n(s, \tilde{x}_s^n) \, d\tilde{\mathbf{w}}_s^n + \int_0^t b_n(s, \tilde{x}_s^n) \, ds \tag{2}$$

for each $t \ge 0$ almost surely. We have thus completed the first stage of proving Theorem 1. If we had so far the processes x_t^n, the convergence property of which we knew nothing about, we have now the convergent processes $\tilde{x}_t^n$. However, in contrast to x_t^n, the processes $\tilde{x}_t^n$ satisfy an equation containing a Wiener process which changes as n changes.

We take the limit in (2) as $n \to \infty$. For each $n_0 \ge 1$, we have

$$\int_0^t \sigma_n(s, \tilde{x}_s^n) \, d\tilde{\mathbf{w}}_s^n = \int_0^s \sigma_{n_0}(s, \tilde{x}_s^n) \, d\tilde{\mathbf{w}}_s^n + \int_0^t [\sigma_n - \sigma_{n_0}](s, \tilde{x}_s^n) \, d\tilde{\mathbf{w}}_s^n, \tag{3}$$

where $\sigma_{n_0}(s, x_s)$ satisfies the Lipschitz condition with respect to (s, x). Hence $\|\sigma_{n_0}(t_2, \tilde{x}_{t_2}^n) - \sigma_{n_0}(t_1, \tilde{x}_{t_1}^n)\| \le N(|t_2 - t_1| + |\tilde{x}_{t_2}^n - \tilde{x}_{t_1}^n|)$.

In addition, by virtue of (1)

$$\lim_{h\downarrow 0}\sup_{n}\sup_{|t_2-t_1|<h}\mathsf{P}\{\|\sigma_{n_0}(t_2,\tilde{x}^n_{t_2})-\sigma_{n_0}(t_1,\tilde{x}^n_{t_1})\|>\varepsilon\}=0.$$

From this it follows, according to Lemma 3, that the first term in (3) tends in probability to $\int_0^t \sigma_{n_0}(s,\tilde{x}^0_s)\,d\tilde{\mathrm{w}}^0_s$. Therefore, applying Chebyshev's inequality, we obtain

$$\lim_{n\to\infty}\mathsf{P}\left\{\left|\int_0^t \sigma_n(s,\tilde{x}^n_s)\,d\tilde{\mathrm{w}}^n_s-\int_0^t \sigma_0(s,\tilde{x}^0_s)\,d\tilde{\mathrm{w}}^0_s\right|>\varepsilon\right\}$$

$$\leq \overline{\lim_{n\to\infty}}\,\mathsf{P}\left\{\left|\int_0^t \sigma_{n_0}(s,\tilde{x}^n_s)\,d\tilde{\mathrm{w}}^n_s-\int_0^t \sigma_{n_0}(s,\tilde{x}^0_s)\,d\tilde{\mathrm{w}}^0_s\right|>\frac{\varepsilon}{3}\right\}$$

$$+\overline{\lim_{n\to\infty}}\,\mathsf{P}\left\{\left|\int_0^t [\sigma_n(s,\tilde{x}^n_s)-\sigma_{n_0}(s,\tilde{x}^n_s)]\,d\tilde{\mathrm{w}}^n_s\right|>\frac{\varepsilon}{3}\right\}$$

$$+\mathsf{P}\left\{\left|\int_0^t [\sigma_{n_0}(s,\tilde{x}^0_s)-\sigma_0(s,\tilde{x}^0_s)]\,d\tilde{\mathrm{w}}^0_s\right|>\frac{\varepsilon}{3}\right\}$$

$$\leq \frac{9}{\varepsilon^2}\left[\overline{\lim_{n\to\infty}}\,\mathsf{M}\int_0^t \|\sigma_n-\sigma_{n_0}\|^2(s,\tilde{x}^n_s)\,ds+\mathsf{M}\int_0^t \|\sigma_{n_0}-\sigma_0\|^2(s,\tilde{x}^0_s)\,ds\right].$$

We estimate the last expression. It is seen that

$$\mathsf{M}\int_0^t |f(s,\tilde{x}^n_s)|\,ds\leq e^t\mathsf{M}\int_0^t e^{-s}|f(s,\tilde{x}^n_s)|\,ds\leq e^t\mathsf{M}\int_0^\infty e^{-s}|f(s,\tilde{x}^n_s)|\,ds.$$

Therefore, by Theorem 3.4[9]

$$\mathsf{M}\int_0^T |f(s,\tilde{x}^n_s)|\,ds\leq N\|f\|_{d+1,H_\infty}$$

for $n\geq 1$, where N does not depend on n. For $n=0$ the last inequality as well holds, which fact we can can easily prove for continuous f taking the limit as $n\to\infty$ and using Fatou's lemma. Furthermore, we can prove it for all Borel f applying the results obtained in [54, Chapter 1, §2]. Let $w(t,x)$ be a continuous function equal to zero for $t^2+|x|^2\geq 1$ and such that $w(0,0)=1$, $0\leq w(t,x)\leq 1$. Then, for $R>0$

$$\overline{\lim_{n\to\infty}}\,\mathsf{M}\int_0^t \|\sigma_n-\sigma_{n_0}\|^2(s,\tilde{x}^n_s)\,ds\leq N\mathsf{M}\int_0^t\left[1-w\left(\frac{s}{R},\frac{\tilde{x}^0_s}{R}\right)\right]ds$$

$$+\overline{\lim_{n\to\infty}}\,\mathsf{M}\int_0^t w\left(\frac{s}{R},\frac{\tilde{x}^n_s}{R}\right)$$

$$\times\|\sigma_n-\sigma_{n_0}\|^2(s,\tilde{x}^n_s)\,ds$$

$$\leq N\mathsf{M}\int_0^t\left[1-w\left(\frac{s}{R},\frac{\tilde{x}^0_s}{R}\right)\right]ds$$

$$+N\|\|\sigma_0-\sigma_{n_0}\|^2\|_{d+1,C_{R,R}}.$$

[9] It should be noted that $\frac{1}{2}(\sigma_n\sigma_n^*\lambda)=\frac{1}{2}|\sigma_n^*\lambda|^2\geq\frac{1}{2}\delta^2|\lambda|^2$ since $\delta|\lambda|^2\leq(\sigma_n\lambda,\lambda)=(\lambda,\sigma_n^*\lambda)\leq|\lambda|\,|\sigma_n^*\lambda|$.

Estimating $\mathsf{M}\int_0^t \|\sigma_{n_0} - \sigma_0\|^2(r,\tilde{x}_r^0)\,dr$ in similar fashion, we find

$$\overline{\lim_{n\to\infty}}\, \mathsf{P}\left\{\left|\int_0^t \sigma_n(s,\tilde{x}_s^n)\,d\tilde{\mathbf{w}}_s^n - \int_0^t \sigma_0(s,\tilde{x}_s^0)\,d\tilde{\mathbf{w}}_s^0\right| > \varepsilon\right\}$$

$$\leq \varepsilon^{-2} N\left\{\mathsf{M}\int_0^t \left[1 - w\left(\frac{s}{R},\frac{\tilde{x}_s^0}{R}\right)\right] ds + \|\|\sigma_0 - \sigma_{n_0}\|^2\|_{d+1,\,C_{R,R}}\right\}$$

for each $n_0 > 0$, $R > 0$. Finally, we note that the last expression tends to zero if we assume first that $n_0 \to \infty$, and next, that $R \to \infty$. Therefore,

$$\int_0^t \sigma_n(s,\tilde{x}_s^n)\,d\tilde{\mathbf{w}}_s^n \to \int_0^t \sigma_0(s,\tilde{x}_s^0)\,d\tilde{\mathbf{w}}_s^0$$

in probability. We have a similar situation for the second integral in (2). Therefore it follows from (2) that

$$\tilde{x}_t^0 = x + \int_0^t \sigma_0(s,\tilde{x}_s^0)\,d\tilde{\mathbf{w}}_s^0 + \int_0^t b_0(s,\tilde{x}_s^0)\,ds$$

for each $t \geq 0$ almost surely. It remains only to note that each side of the last equality is continuous with respect to t; hence both sides coincide on a set of complete probability. We have thus proved Theorem 1. □

7. Some Properties of a Random Process Depending on a Parameter

In investigating the smoothness property of a payoff function in optimal control problems it is convenient to use theorems on differentiability in the mean of random variables over some parameter. It turns out frequently that the random variable in question, say $J(p)$, depends on a parameter p in a complicated manner. For example, $J(p)$ can be given as a functional of trajectories of some process x_t^p which depends on p. In this section we prove the assertions about differentiability in the mean of such or other functionals of the process.

Three constants T, K, $m > 0$ will be fixed throughout the entire section.

1. Definition. Let a real random process $x_t(\omega)$ be defined for $t \in [0,T]$. We write $x_t \in \mathscr{L}$ if the process $x_t(\omega)$ is measurable with respect to (ω,t) and for all $q \geq 1$

$$\mathsf{M}\int_0^T |x_t|^q\,dt < \infty.$$

We write $x_t \in \mathscr{L}B$ if x_t is a separable process and for all $q \geq 1$

$$\mathsf{M}\sup_{t\leq T}|x_t|^q < \infty.$$

The convergence property in the sets $\mathscr{L}$, $\mathscr{L}B$ can be defined in a natural way.

2. Definition. Let $x_t^0, x_t^1, \ldots, x_t^n, \ldots \in \mathscr{L}(\mathscr{L}B)$. We say that the $\mathscr{L}$-limit ($\mathscr{L}B$-limit) of the process x_t^n equals x_t^0, and we write $\mathscr{L}\text{-}\lim_{n\to\infty} x_t^n = x_t^0$ ($\mathscr{L}B\text{-}\lim_{n\to\infty} x_t^n = x_t^0$) if for all $q \geq 1$

$$\lim_{n\to\infty} \mathsf{M} \int_0^T |x_t^n - x_t^0|^q \, dt = 0 \qquad \left(\lim_{n\to\infty} \mathsf{M} \sup_{t\leq T} |x_t^n - x_t^0|^q = 0\right).$$

Having introduced the notions of the $\mathscr{L}$-limit ($\mathscr{L}B$-limit), it is clear what is meant by $\mathscr{L}$-continuity ($\mathscr{L}B$-continuity) of the process x_t^p with respect to the parameter p at a point p_0.

3. Definition. Suppose that $p_0 \in E_d$, unit vector $l \in E_d$, $y_t \in \mathscr{L}(\mathscr{L}B)$. Further, suppose that for each p from some neighborhood of the point p_0 a process $x_t^p \in \mathscr{L}(\mathscr{L}B)$ is given. We say that y_t is an $\mathscr{L}$-derivative ($\mathscr{L}B$-derivative) of x_t^p at the point p_0 along the l direction, and also, we write

$$y_t = \mathscr{L}\text{-}\frac{\partial}{\partial l} x_t^p\Big|_{p=p_0} \qquad \left(y_t = \mathscr{L}B\text{-}\frac{\partial}{\partial l} x_t^p\Big|_{p=p_0}\right),$$

if

$$y_t = \mathscr{L}\text{-}\lim_{r\to 0} \frac{1}{r}(x_t^{p_0+rl} - x_t^{p_0}) \qquad y_t = \mathscr{L}B\text{-}\lim_{r\to 0} \frac{1}{r}(x_t^{p_0+rl} - x_t^{p_0}) \ .$$

We say that the process x_t^p is once $\mathscr{L}$-differentiable ($\mathscr{L}B$-differentiable) at the point p_0 if this process x_t^p has $\mathscr{L}$-derivatives ($\mathscr{L}B$-derivatives) at the point p_0 along all l directions. The process x_t^p is said to be i times ($i \geq 2$) $\mathscr{L}$-differentiable ($\mathscr{L}$B-differentiable) at the point p_0 if this process x_t^p is once $\mathscr{L}$-differentiable ($\mathscr{L}B$-differentiable) in some neighborhood[10] of the point p_0 and, in addition, each (first) $\mathscr{L}$-derivative ($\mathscr{L}B$-derivative) of this process x_t^p is $i-1$ times $\mathscr{L}$-differentiable ($\mathscr{L}B$-differentiable) at the point p_0.

Definitions 1–3 have been given for numerical processes x_t only. They can be extended to vector processes and matrix processes x_t in the obvious way.

Further, as is commonly done in conventional analysis, we write $y_t^p = \mathscr{L}\text{-}(\partial/\partial l)x_t^p$ if $y_t^{p_0} = \mathscr{L}\text{-}(\partial/\partial l)x_t^p|_{p=p_0}$ for all p_0 considered, $\mathscr{L}\text{-}(\partial/\partial l_1\,\partial l_2)x_t^p \equiv \mathscr{L}\text{-}(\partial/\partial l_1)[\mathscr{L}\text{-}(\partial/\partial l_2)x_t^p]$, etc. We say that x_t^p is i times $\mathscr{L}$-continuously $\mathscr{L}$-differentiable if all $\mathscr{L}$-derivatives of x_t^p up to order i inclusive are $\mathscr{L}$-continuous. We shall not dwell in future on the explanation of such obvious facts.

We shall apply Definitions 1–3 to random variables as well as random processes, the former being regarded as time independent processes.

In order to grow familiar with the given definitions, we note a few simple properties these definitions imply. It is obvious that the notion of $\mathscr{L}$-continuity is equivalent to that of $\mathscr{L}B$-continuity for random variables. Furthermore, $|\mathsf{M}x^p - \mathsf{M}x^{p_0}| \leq \mathsf{M}|x^p - x^{p_0}|$. Hence the expectation of an $\mathscr{L}$-

[10] That is, at each point of this neighborhood.

continuous random variable is continuous. Since

$$\left|\frac{1}{r}(\mathsf{M}x^{p_0+rl} - \mathsf{M}x^{p_0}) - \mathsf{M}y\right| \le \mathsf{M}\left|\frac{1}{r}(x^{p_0+rl} - x^{p_0}) - y\right|,$$

the derivative of $\mathsf{M}x^p$ along the l direction at a point p_0 is equal to the expectation of the $\mathscr{L}$-derivative of x^p if the latter exists. Therefore, the sign of the first derivative is interchangeable with the sign of the expectation. Combining the properties listed in an appropriate way, we deduce that $(\partial/\partial l)\mathsf{M}x^p$ exists and it is continuous at the point p_0 if the variable x^p is $\mathscr{L}$-continuously $\mathscr{L}$-differentiable at the point p_0 along the l direction. A similar situation is observed for derivatives of higher orders.

Since for $\tau \le T$

$$\mathsf{M}|x_\tau^p - x_\tau^{p_0}|^q \le \mathsf{M}\sup_{t\le T}|x_t^p - x_t^{p_0}|^q,$$

x_τ^p is an $\mathscr{L}$-continuous variable if $\tau(\omega) \le T$ for all ω, x_t^p is an $\mathscr{L}B$-continuous process, and x_τ^p is a measurable function of ω. A similar inequality shows that for the same τ

$$\mathscr{L}\text{-}\frac{\partial}{\partial l}x_\tau^p = \left(\mathscr{L}B\text{-}\frac{\partial}{\partial l}x_t^p\right)\bigg|_{t=\tau} \tag{1}$$

if x_t^p has an $\mathscr{L}B$-derivative along the l direction, and if x_τ^p and the right side of (1) are measurable functions of ω. These arguments allow us to derive the properties of $\mathscr{L}$-continuity and $\mathscr{L}$-differentiability of the random variable x_τ^p from the properties of $\mathscr{L}$B-continuity and $\mathscr{L}$B-differentiability of the process x_t^p. Furthermore, (1) shows that the order of the substitution of t for τ and the order of the computation of derivatives are interchangeable.

Suppose that the process x_t^p is continuous with respect to t and is $\mathscr{L}$B-continuous with respect to p at a point p_0. Also, suppose that $\tau(p)$ are random functions with values in $[0,T]$, continuous in probability at the point p_0. We assert that in this case $x_{\tau(p)}^{p_0}$, $x_{\tau(p)}^p$ are $\mathscr{L}$-continuous at the point p_0. In fact, the difference $|x_{\tau(p)}^{p_0} - x_{\tau(p_0)}^{p_0}|^q \to 0$ in probability as $p \to p_0$ and in addition, this difference is bounded by the summable quantity $2^{q-1}\sup_t|x_t^{p_0}|^q$. Therefore, the expectation of the difference indicated tends to zero, i.e., the variable $x_{\tau(p)}^{p_0}$ is $\mathscr{L}$-continuous. The $\mathscr{L}$-continuity of the second variable follows from the $\mathscr{L}$-continuity of the first variable and from the inequalities

$$\begin{aligned}\mathsf{M}|x_{\tau(p)}^p - x_{\tau(p_0)}^{p_0}|^q &\le 2^{q-1}\mathsf{M}|x_{\tau(p)}^p - x_{\tau(p)}^{p_0}|^q + 2^{q-1}\mathsf{M}|x_{\tau(p)}^{p_0} - x_{\tau(p_0)}^{p_0}|^q\\ &\le 2^{q-1}\mathsf{M}\sup_{t\in[0,T]}|x_t^p - x_t^{p_0}|^q + 2^{q-1}\mathsf{M}|x_{\tau(p)}^{p_0} - x_{\tau(p_0)}^{p_0}|^q.\end{aligned}$$

In conjunction with Hölder's inequality

$$\mathsf{M}\sup_{t\le T}\left|\int_0^t x_s^p\,ds - \int_0^t x_s^{p_0}\,ds\right|^q \le \mathsf{M}\left|\int_0^T |x_s^p - x_s^{p_0}|\,ds\right|^q \le T^{q-1}\mathsf{M}\int_0^T |x_s^p - x_s^{p_0}|^q\,ds.$$

Therefore, $\int_0^t x_s^p\,ds$ is an $\mathscr{L}B$-continuous process if the process x_t^p is $\mathscr{L}$-continuous. We prove in a similar way that this integral has an $\mathscr{L}B$-derivative along the l direction, which coincides with the integral of the $\mathscr{L}$-derivative of x_t^p along the l direction if the latter derivative exists. In other words, the derivative can be brought under the integral sign.

Combining the assertions given above in an appropriate way, we can obtain many necessary facts. They are, however, too simple to require formal proof.

It is useful to have in mind that if $\{\mathscr{F}_t\}$ is a family of σ-algebras in Ω and if the process x_t^p is k times $\mathscr{L}$-differentiable at the point p_0 and, in addition, progressively measurable with respect to $\{\mathscr{F}_t\}$, all the derivatives of the process x_t^p can be chosen to be progressively measurable with respect to $\{\mathscr{F}_t\}$. Keeping in mind that induction is possible in this situation, we prove the foregoing assertion only for $k = 1$. Let $y_t^p = \mathscr{L}\text{-}(\partial/\partial l)x_t^p$. Having fixed p, we find a sequence $r_n \to 0$ such that $(1/r_n)(x_t^{p+r_n l} - x_t^p) \to y_t^p$ almost everywhere $d\mathsf{P} \times dt$. Further, we take $\tilde{y}_t^p = \lim_{n\to\infty} (1/r_n)(x_t^{p+r_n l} x_t^p)$ for those ω, t for which this limit exists and $\tilde{y}_t^p = 0$ on the remaining set. It is seen that the process $\tilde{y}_t^p$ is progressively measurable. Also, it is seen that

$$\tilde{y}_t^p = \mathscr{L}\text{-}\frac{\partial}{\partial l} x_t^p$$

since $\tilde{y}_t^p = y_t^p$ ($d\mathsf{P} \times dt$-a.s.).

We shall take this remark into account each time we calculate $\mathscr{L}$-derivatives of a stochastic integral.

We have mentioned above that differentiation is interchangeable with the (standard) integration. Applying Corollary 5.11, we immediately obtain that if $(\mathbf{w}_t, \mathscr{F}_t)$ is a d_1-dimensional Wiener process, σ_t^p is a matrix of dimension $d_2 \times d_1$, which is progressively measurable with respect to $\{\mathscr{F}_t\}$ and is $\mathscr{L}$-continuous at a point p_0, the integral $\int_0^t \sigma_t^p\,d\mathbf{w}_t$ is $\mathscr{L}B$-continuous at the point p_0. If σ_t^p is $\mathscr{L}$-differentiable along the l direction at the point p_0, then for $p = p_0$

$$\mathscr{L}B\text{-}\frac{\partial}{\partial l}\int_0^t \sigma_s^p\,d\mathbf{w}_s = \int_0^t \left(\mathscr{L}\text{-}\frac{\partial}{\partial l}\sigma_s^p\right) d\mathbf{w}_s.$$

A similar assertion is valid in an obvious way for derivatives of higher orders.

4. Exercise

Prove that if the function x_t^p is continuous (continuously differentiable) with respect to p in the usual sense for all (t,ω) and, in addition, the function $\mathsf{M}\int_0^T |x_t^p|^q\,dt$ ($\mathsf{M}\int_0^T |(\partial/\partial x)x_t^p|^q\,dt$ for each l, $|l| = 1$) is bounded in some region for each $q \geq 1$, the process x_t^p is $\mathscr{L}$-continuous ($\mathscr{L}$-differentiable and $\mathscr{L}\text{-}(\partial/\partial l)x_t^p = (\partial/\partial l)x_t^p$) in this region.

Further, we turn to investigating the continuity and differentiability properties of a composite function. To do this, we need three lemmas.

5. Lemma. *Suppose that for* $n = 1, 2, \ldots, t \in [0, T]$, $x \in E_d$, *the* d_1*-dimensional processes* x_t^n *measurable with respect to* (ω,t) *are defined, and, in addition, the variables* $h_t^n(x)$ *measurable with respect to* (ω,t,x) *are given. Assume that* $x_t^n \to 0$ *as* $n \to \infty$ *with respect to the measure* $d\mathsf{P} \times dt$, *and that the variable* $h_t^n(x)$ *is continuous in* x *for all* n, ω, t. *Furthermore, we assume that one of the following two conditions is satisfied*:

a. *for almost all* (ω,t)

$$\lim_{\delta\to 0} \varlimsup_{n\to\infty} w_t^n(\delta) = 0,$$

where $w_t^n(\delta) = \sup_{|x| \le \delta} |h_t^n(x)|$;

b. *for each* $\varepsilon > 0$

$$\lim_{\delta\to 0} \varlimsup_{n\to\infty} \int_0^T \mathsf{P}\{w_t^n(\delta) > \varepsilon\}\, dt = 0.$$

Then $|h_t^n(x_t^n)| \le w_t^n(|x_t^n|) \to 0$ *as* $n \to \infty$ *in measure* $d\mathsf{P} \times dt$.

Proof. We note that since $h_t^n(x)$ is continuous in $x, w_t^n(\delta)$ will be measurable with respect to (ω,t). Further, condition (b) follows from condition (a) since (a) implies that $w_t^n(\delta) \to 0$ as $n \to \infty$, $\delta \to 0$ almost everywhere; (b) implies the same although with respect to $d\mathsf{P} \times dt$.

Finally, for each $\varepsilon > 0$, $\delta > 0$

$$\varlimsup_{n\to\infty} \int_0^T \mathsf{P}\{w_t^n(|x_t^n|) > \varepsilon\}\, dt \le \varlimsup_{n\to\infty} \int_0^T \mathsf{P}\{|x_t^n| > \delta\}\, dt + \varlimsup_{n\to\infty} \int_0^T \mathsf{P}\{w_t^n(\delta) > \varepsilon\}\, dt,$$

where the first summand equals zero by assumption. Thus, letting $\delta \to 0$ and using (b), we have proved the lemma. □

6. Lemma. *Let* x_t^n *be* d_1*-dimensional processes measurable with respect to* (ω,t) $(n = 0,1,2, \ldots, t \in [0,T])$, *such that* $\mathscr{L}\text{-}\lim_{n\to\infty} x_t^n = x_t^0$. *Let* $f_t(x)$ *be random variables defined for* $t \in [0,T]$, $x \in E_d$, *measurable with respect to* (ω,t), *continuous in* x *for all* (ω,t), *and such that* $|f_t(x)| \le K(1 + |x|)^m$ *for all* ω, t, x. *Then* $\mathscr{L}\text{-}\lim_{n\to\infty} f_t(x_t^n) = f(x_t^0)$.

Proof. First we note that under the condition $|f_t(x)| \le K(1 + |x|)^m$ the processes $f_t(x_t^n) \in \mathscr{L}$ for all $n \ge 0$. Next, we write $f_t(x_t^n) - f_t(x_t^0)$ as $h_t(y_t^n)$, where $h_t(x) = f_t(x + x_t^0) - f_t(x_t^0)$, $y_t^n = x_t^n - x_t^0$. Since $\mathsf{M}\int_0^T |y_t^n|\, dt \to 0$, $y_t^n \to 0$ in measure $d\mathsf{P} \times dt$, from which we have, using Lemma 5 applied to y_t^n and $h_t(x)$, that $h_t(y_t^n) \to 0$ in measure $d\mathsf{P} \times dt$.

Since the function $|a|/(|a| + 1)$ is bounded and

$$g_t^n = |h_t(y_t^n)|[|h_t(y_t^n)| + 1]^{-1} \to 0$$

in measure $d\mathsf{P} \times dt$, then for each $q \geq 1$

$$\lim_{n\to\infty} \mathsf{M} \int_0^T |g_t^n|^{2q}\, dt = 0. \tag{2}$$

Moreover, in view of the estimate $|f_t(x)| \leq K(1 + |x|)^m$ and the fact that

$$\mathsf{M} \int_0^T |x_t^n - x_t^0|^{2qm}\, dt \to 0, \qquad \mathsf{M} \int_0^T |x_t^0|^{2qm}\, dt < \infty,$$

we have

$$\sup_n \mathsf{M} \int_0^T |x_t^n|^{2qm}\, dt < \infty,$$

$$\sup_n \mathsf{M} \int_0^T (1 + |h_t(y_t^n)|)^{2q}\, dt \leq \sup_n \mathsf{M} \int_0^T [1 + K(1 + |x_t^n|)^m + K(1 + |x_t^0|)^m]^{2q}\, dt < \infty. \tag{3}$$

Using the Cauchy inequality, we derive from (2) and (3) that

$$\lim_{n\to\infty} \mathsf{M} \int_0^T |h_t(y_t^n)|^q\, dt \leq \lim_{n\to\infty} \left(\mathsf{M} \int_0^T |g_t^n|^{2q}\, dt\right)^{1/2} \left(\mathsf{M} \int_0^T (1 + |h_t(y_t^n)|)^{2q}\, dt\right)^{1/2} = 0$$

for each $q \geq 1$. The lemma is proved. □

We note a simple corollary of Lemma 6.

7. Corollary. *If for $n = 0, 1, 2, \ldots$ the one-dimensional processes x_t^n, y_t^n are defined and $\mathscr{L}\text{-}\lim_{n\to\infty} x_t^n = x_t^0$, $\mathscr{L}\text{-}\lim_{n\to\infty} y_t^n = y_t^0$, then*

$$\mathscr{L}\text{-}\lim_{n\to\infty} x_t^n y_t^n = x_t^0 y_t^0$$

Indeed, the two-dimensional process (x_t^n, y_t^n) has the $\mathscr{L}$-limit equal to (x_t^0, y_t^0). Furthermore, the function $f(x,y) \equiv xy$ satisfies the growth condition $|f(x,y)| \leq (1 + \sqrt{x^2 + y^2})^2$. Hence $\mathscr{L}\text{-}\lim_{n\to\infty} f(x_t^n, y_t^n) = f(x_t^0, y_t^0)$.

8. Lemma. *Suppose that the assumptions of Lemma 6 are satisfied. Also, suppose that for $n = 1, 2, \ldots, u \in [0,1]$ the d_1-dimensional random variables $x_t^n(u)$ are defined which are continuous in u, measurable with respect to (ω,t) and such that $|x_t^n(u) - x_t^0| \leq |x_t^n - x_t^0|$. Then*

$$\mathscr{L}\text{-}\lim_{n\to\infty} \int_0^1 f_t(x_t^n(u))\, du = f_t(x_t^0). \tag{4}$$

Proof. In accord with Hölder's inequality, for $q \geq 1$

$$\left|\int_0^1 f_t(x_t^n(u))\, du - f_t(x_t^0)\right|^q = \left|\int_0^1 [f_t(x_t^n(u)) - f_t(x_t^0)]\, du\right|^q \leq \int_0^1 \left|f_t(x_t^n(u)) - f_t(x_t^0)\right|^q du.$$

It follows from the inequalities $|x_t^n(u) - x_t^0| \le |x_t^n - x_t^0|$, $|x_t^n(u)| \le |x_t^0| + |x_t^n - x_t^0|$ that $x_t^n(u) \in \mathscr{L}$ and $\mathscr{L}\text{-}\lim_{n\to\infty} x_t^n(u) = x_t^0$ for each $u \in [0,1]$. Therefore, by Lemma 6

$$I_n(u) \equiv \mathsf{M} \int_0^T |f_t(x_t^n(u)) - f_t(x_t^0)|^q \, dt \to 0.$$

Finally, by the inequalities

$$|f_t(x_t^n(u))| \le K(1 + |x_t^n(u)|)^m \le K(1 + |x_t^n - x_t^0| + |x_t^0|)^m$$

we have that the limiting expression in (4) belongs to $\mathscr{L}$, and the totality of variables $I_n(u)$ is bounded. By the Lebesgue theorem, as $n \to \infty$

$$\mathsf{M} \int_0^T \left| \int_0^1 f_t(x_t^n(u)) \, du - f_t(x_t^0) \right|^q dt \le \int_0^1 I_n(u) \, du \to 0,$$

thus proving the lemma. □

Further, we prove a theorem on continuity and differentiability of a composite function.

9. **Theorem.** *Suppose for $x \in E_{d_1}$ and p in a neighborhood of a point $p_0 \in E_d$ the random processes $x_t^p = x_t^p(\omega)$, $f_t(x) = f_t(\omega,x)$ with values in E_d and E_1, respectively, are given for $t \in [0,T]$ and measurable with respect to (t,ω).*

(a) *For all t, ω let the function $f_t(x)$ be continuous in x, let $|f_t(x)| \le K(1 + |x|)^m$, and let the process x_t^p be $\mathscr{L}$-continuous at p_0. Then the process $f_t(x_t^p)$ is also $\mathscr{L}$-continuous at p_0.*

(b) *Suppose that for all t,ω the function $f_t(x)$ is i times continuously differentiable over x. Furthermore, suppose that for all t, ω the absolute values of the function $f_t(x)$ as well as those of its derivatives up to order i inclusively do not exceed $K(1 + |x|)^m$. Then, if the process x_t^p is i times ($\mathscr{L}$-continuously) $\mathscr{L}$-differentiable at the point p_0, the process $f_t(x_t^p)$ is i times ($\mathscr{L}$-continuously) $\mathscr{L}$-differentiable at the point p_0 as well. In addition, for the unit vector $l \in E_d$*

$$\mathscr{L}\text{-}\frac{\partial}{\partial l} f_t(x_t^p) = f_{t(y_t^p)}(x_t^p)|y_t^p|, \tag{5}$$

$$\mathscr{L}\text{-}\frac{\partial^2}{\partial l^2} f_t(x_t^p) = f_{t(z_t^p)}(x_t^p)|z_t^p| + f_{t(y_t^p)(y_t^p)}(x_t^p)|y_t^p|^2, \tag{6}$$

where

$$z_t^p = \mathscr{L}\text{-}\frac{\partial^2}{\partial l^2} x_t^p, \qquad y_t^p = \mathscr{L}\text{-}\frac{\partial}{\partial l} x_t^p,$$

for those i, p for which the existence of the left sides of (5) and (6) has been established.

Proof. For proving (a) it suffices to take any sequence of points $p_n \to p_0$, to put $x_t^{(n)} = x_t^{p_n}$ and, finally, to make use of Lemma 6.

We shall prove (b) for $i = 1$. First we note that $f_t(x,y) \equiv f_{t(y)}(x)|y|$ is a continuous function of (x,y) and

$$|f_t(x,y)| = |f_{t(y)}(x)|\,|y| \leq K(1 + |x|)^m|y| \leq N(1 + \sqrt{|x|^2 + |y|^2})^{m+1}.$$

Further, we take the unit vector $l \in E_d$, a sequence of numbers $r_n \to 0$, and we put

$$x_t^{(n)}(u) = ux_t^{p_0 + r_n l} + (1 - u)x_t^{p_0}, \qquad y_t^{(n)} = \frac{1}{r_n}(x_t^{p_0 + r_n l} - x_t^{p_0}).$$

Using the Newton-Leibniz rule we have

$$\frac{1}{r_n}[f_t(x_t^{p_0 + r_n l}) - f_t(x_t^{p_0})] = \frac{1}{r_n}\int_0^1 \frac{\partial}{\partial u} f_t(x_t^{(n)}(u))\,du$$
$$= \int_0^1 f_t(x_t^{(n)}(u), y_t^{(n)})\,du,$$

where $|x_t^{(n)}(u) - x_t^{p_0}|^2 + |y_t^{(n)} - y_t^{p_0}|^2 \leq |x_t^{p_0 + r_n l} - x_t^{p_0}|^2 + |y_t^{(n)} - y_t^{p_0}|^2$ and where, by Lemma 8 applied to $x_t^{(n)}(u)$ and $y_t^{(n)}(u) \equiv y_t^{(n)}$,

$$\mathscr{L}\text{-}\lim_{n\to\infty} \int_0^1 f_t(x_t^{(n)}(u), y_t^{(n)})\,du = f_t(x_t^{p_0}, y_t^{p_0}).$$

Therefore,

$$\mathscr{L}\text{-}\lim_{r\to 0} \frac{1}{r}[f_t(x_t^{p_0 + rl}) - f_t(x_t^{p_0})] = f_t(x_t^{p_0}, y_t^{p_0}).$$

Finally, by (a), $f_t(x^{p_0}, y_t^{p_0})$ is $\mathscr{L}$-continuous with respect to p_0 if $x_t^{p_0}$ is $\mathscr{L}$-continuously $\mathscr{L}$-differentiable with respect to p_0. This proves assertion one in (b) for $i = 1$. At the same time we have proved Eq. (5), which we find convenient to write as follows:

$$\mathscr{L}\text{-}\frac{\partial}{\partial l} f_t(x_t^p) = f_t(x_t^p, y_t^p).$$

For proving (b) for all i we apply the method of induction. Assume that the first assertion in (b) is proved for $i \leq j$ and for any processes $f_t(x)$, x_t^p satisfying condition (b). Let the pair $f_t(x)$, x_t^p satisfy the conditions of (b) for $i = j + 1$. We take a derivative $\mathscr{L}\text{-}(\partial/\partial l)f_t(x_t^p)$ and prove that this derivative is j times $\mathscr{L}$-differentiable at a point p_0. Let us write this derivative as $f_t(x_t^p, y_t^p)$. We note that the process (x_t^p, y_t^p) is j times $\mathscr{L}$-differentiable at the point p_0 by assumption, the function $f_t(x,y)$ is continuously differentiable j times with respect to the variables (x,y). Also, we note that the absolute values of the derivatives of the above function up to order j inclusively do not exceed $N(1 + \sqrt{|x|^2 + |y|^2})^{m+1}$. Therefore, by the induction assumption, $f_t(x_t^p, y_t^p)$ is j times $\mathscr{L}$-differentiable at the point p_0. Since l is a vector, $f_t(x_t^p)$ is, by definition, $j + 1$ times $\mathscr{L}$-differentiable at the point p_0.

In a similar way we can prove $\mathscr{L}$-continuity of $\mathscr{L}$-derivatives of $f_t(x_t^p)$ at the point p_0 if $\mathscr{L}$-derivatives of x_t^p are $\mathscr{L}$-continuous at the point p_0. Finally, in conjunction with (5)

$$\mathscr{L}\text{-}\frac{\partial}{\partial l} f_t(x_t^p, y_t^p) = f_{t(y_t^p, z_t^p)}(x_t^p, y_t^p)\sqrt{|y_t^p|^2 + |z_t^p|^2},$$

which, after simple transformations, yields (6). The theorem is proved. □

10. Remark. The theorem proved above can easily be used for proving the $\mathscr{L}$-continuity and $\mathscr{L}$-differentiability of various expressions which contain random processes. For example, arguing in the same way as in Corollary 7, we can prove that if x_t^p, y_t^p are real i times $\mathscr{L}$-differentiable processes, the product $x_t^p y_t^p$ is i times $\mathscr{L}$-differentiable as well. If the real nonnegative process x_t^p is i times $\mathscr{L}$-differentiable, the process $e^{-x_t^p}$ is i times $\mathscr{L}$-differentiable as well. In fact, notwithstanding that the function e^{-x} grows more rapidly than any polynomial as $x \to -\infty$, we consider the nonnegative process x_t^p, and moreover, we can take any smooth function $f(x)$ equal to zero for $x \leq -1$ and equal to e^{-x} for $x \geq 0$. In this situation the hypotheses of the theorem concerning $f(x)$ will be satisfied and $e^{-x_t^p} = f(x_t^p)$. Combining the foregoing arguments with the known properties of integrals of $\mathscr{L}$-continuous and $\mathscr{L}$-differentiable functions, we arrive at the following assertion.

11. Lemma. *Let the processes $x_t^p, f_t^1(x), f_t^2(x)$ satisfy the conditions of Theorem* 9a (*Theorem* 9b), *and, in addition, let $f_t^1(x) \geq 0$; then the process*

$$f_t^2(x_t^p)\exp\left\{-\int_0^t f_s^1(x_s^p)\,ds\right\}$$

is $\mathscr{L}$-continuous at the point p_0 (i times ($\mathscr{L}$-continuously) $\mathscr{L}$-differentiable at the point p_0).

Fixing $\tau \in [0,T]$, and regarding $\int_0^\tau f_s^1(x_s^p)\,ds$ as a time independent process, we conclude that the following lemma is valid.

12. Lemma. *Let the processes x_t^p, $f_t^1(x)$, $f_t^2(x)$ satisfy the hypotheses of Theorem* 9a (*Theorem* 9b), *and, in addition, let $f_t^1(x) \geq 0$. Let the random variable $\tau(\omega) \in [0,T]$ and let the random processes y_t^p, $f_\tau^2(x)$ be such that the processes $\tilde{x}_t \equiv y_\tau^p$, $\tilde{f}_t(x) \equiv f_\tau^2(x)$ satisfy the hypotheses of Theorem* 9a (*Theorem* 9b). *Then the random variable*

$$f_\tau^2(y_\tau^p)\exp\left\{-\int_0^\tau f_s^1(x_s^p)\,ds\right\}$$

is $\mathscr{L}$-continuous at the point p_0 (i times ($\mathscr{L}$-continuously) $\mathscr{L}$-differentiable at the point p_0).

13. Remark. Equation (5) shows that in computing an $\mathscr{L}$-derivative of a composite function the usual formulas familiar in analysis can be applied.

14. Exercise

Derive a formula for the derivative of a product, using (5). (*Hint*: Take a function $f(x,y) \equiv xy$.)

We have investigated the properties of the functions $f_t(x_t^n)$ in the case where $f_t(x)$ does not depend on n. We prove a few assertions for the case where $f_t(x)$ depends on the parameter n in an explicit manner.

15. Lemma. *Let $\xi(\omega)$ be a d_1-dimensional random vector. Further, let $h(x) = h(\omega,x)$, $w(R,\varepsilon) = w(\omega,R,\varepsilon)$ be measurable variables which are defined for $x \in E_{d_1}$, $R \geq 0$, $\varepsilon \geq 0$, $\omega \in \Omega$. Assume that $w(R,\varepsilon)$ increases with respect to R and to ε, $|h(x) - h(y)| \leq w(|x| \vee |y|, |x - y|)$ for all ω, x, y and $|h(x)| \leq K(1 + |x|)^m$ for all ω, x. Then, for all $R \geq 0$, $\varepsilon \in (0,1)$*

$$\mathsf{M}|h(\xi)| \leq K\mathsf{M}(1 + |\xi|)^m \chi_{|\xi| > R-1} + \mathsf{M}w(R,\varepsilon) + N(d_1)\varepsilon^{-d_1}\mathsf{M}\int_{|y| \leq R} |h(y)|\, dy.$$

Proof. We fix $R \geq 0$, $\varepsilon \in (0,1)$, and also we take a d_1-dimensional vector η such that it does not depend on ξ, w and is uniformly distributed in the sphere $\{x \in E_{d_1} : |x| < \varepsilon\}$.

It is seen that

$$\mathsf{M}|h(\xi)| \leq \mathsf{M}|h(\xi)|\chi_{|\xi| > R-1} + \mathsf{M}|h(\xi) - h(\xi + \eta)|\chi_{|\xi| \leq R-1} + \mathsf{M}\,|h(\xi + \eta)|\chi_{\xi \leq R-1}.$$

The assertion of our lemma follows from the above expression as well as the assumptions of the lemma since $|\eta| < \varepsilon < 1$, for $|\xi| \leq R - 1$

$$|\xi + \eta| < R, \qquad |h(\xi) - h(\xi + \eta)| \leq w(R,\varepsilon),$$

and

$$\begin{aligned} &\mathsf{M}|h(\xi + \eta)|\chi_{|\xi| \leq R-1} \\ &= N(d_1)\varepsilon^{-d_1}\mathsf{M}\int \chi_{|x| < \varepsilon, |\xi| \leq R-1}|h(\xi + x)|\, dx \\ &= N(d_1)\varepsilon^{-d_1}\mathsf{M}\int \chi_{|y-\xi| < \varepsilon, |\xi| \leq R-1}|h(y)|\, dy \\ &\leq N(d_1)\varepsilon^{-d_1}\mathsf{M}\int_{|y| \leq R} |h(y)|\, dy. \end{aligned}$$

□

16. Lemma. *Suppose that for $x \in E_{d_1}$, $t \in [0,T]$, $n = 1, 2, 3, \ldots$, $R > 0$, $\varepsilon > 0$ d_1-dimensional processes x_t^n are defined which are measurable with respect to (ω,t). Furthermore, suppose that the variables $h_t^n(x)$ and $w_t^n(R,\varepsilon)$ increasing with respect to R and ε are defined, these variables being measurable with respect to (ω,t,x) and (ω,t), respectively. Assume that $w_t^n(|x| \vee |y|, |x - y|) \geq |h_t^n(x) - h_t^n(y)|$ for all ω, t, x, y,*

$$\lim_{R \to \infty} \overline{\lim_{n \to \infty}} \int_0^T \mathsf{P}\{|x_t^n| > R\}\, dt = 0, \tag{7}$$

and for each $R > 0, \delta > 0$

$$\lim_{\varepsilon \downarrow 0} \overline{\lim_{n \to \infty}} \int_0^T \mathsf{P}\{w_t^n(R,\varepsilon) > \delta\}\, dt = 0. \tag{8}$$

Finally, let $h_t^n(x) \to 0$ as $n \to \infty$ in measure $d\mathsf{P} \times dt$ for each $x \in E_d$.
Then, $h_t^n(x_t^n) \to 0$ as $n \to \infty$ in measure $d\mathsf{P} \times dt$.

We shall prove this lemma later. We derive from Lemma 16 (in the same way as we derived Lemma 6 from Lemma 5) the following theorem.

17. Theorem. *Let the hypotheses of Lemma 16 be satisfied. Furthermore, let $|h_t^n(x)| \le K(1 + |x|)^m$ for all n, ω, t, x and for all $q \ge 1$*

$$\sup_n \mathsf{M} \int_0^T |x_t^n|^q\, dt < \infty. \tag{9}$$

Then

$$\mathscr{L}\text{-}\lim_{n \to \infty} h_t^n(x_t^n) = 0.$$

18. Remark. By Chebyshev's inequality, (7) follows from (9). Using Chebyshev's inequality it can easily be proved that the condition (8) is satisfied if $w_t^n(R,\varepsilon)$ is nonrandom and

$$\lim_{\varepsilon \downarrow 0} \overline{\lim_{n \to \infty}} \int_0^T w_t^n(R,\varepsilon)\, dt = 0.$$

For $w_t^n(R,\varepsilon)$ it is convenient to take $K\varepsilon$ if $|h_t^n(x) - h_t^n(y)| \le K|x - y|$.

Proof of Lemma 16. Since the convergence of $h_t^n(x_t^n)$ to zero in measure is equivalent to the same convergence of $(2/\pi)\arctan h_t^n(x_t^n)$, and furthermore, since the latter variable is bounded and

$$|\arctan h_t^n(x) - \arctan h_t^n(y)| \le |h_t^n(x) - h_t^n(y)| \le w_t^n(|x| \vee |y|, |x - y|),$$

we can consider without loss of generality that $|h_t^n| \le 1$.

It is clear that in this case $2 \wedge w_t^n$ can be taken instead of w_t^n so that w_t^n will be assumed to be bounded as well.

According to Lemma 15 (we take in Lemma 15 $K = 1$, $m = 0$) for any $R > 0$, $\varepsilon \in (0,1)$

$$\begin{aligned} \int_0^T \mathsf{M}|h_t^n(x_t^n)|\, dt \le \int_0^T & \mathsf{P}\{|x_t^n| > R - 1\}\, dt \\ & + \int_0^T \mathsf{M} w_t^n(R,\varepsilon)\, dt + N(d_1)\varepsilon^{-d_1} \\ & \times \int_0^T \mathsf{M} \int_{|y| \le R} |h_t^n(y)|\, dy\, dt. \end{aligned} \tag{10}$$

We make use of the fact that the convergence in measure is equivalent to the convergence in the mean for uniformly bounded sequences. Thus we have that the sequence

$$\int_0^T \mathsf{M}|h_t^n(y)|\,dt \to 0$$

as $n \to \infty$ for any $y \in E_{d_1}$. Furthermore, each term of this sequence does not exceed T. This implies that the last expression in (10) tends to zero as $n \to \infty$ for any $\varepsilon > 0$, $R > 0$. Letting $n \to \infty$ in (10), next, $\varepsilon \downarrow 0$, $R \to \infty$, and in addition, using (7), (8) and the fact mentioned above that the convergence in the mean and the convergence in measure are related, we complete the proof of Lemma 16. □

8. The Dependence of Solutions of a Stochastic Equation on a Parameter

Let E be a Euclidean space, let a region $D \subset E$ (D denotes a region of parameter variation), and let T, K, m be fixed nonnegative constants; $(\mathbf{w}_t, \mathscr{F}_t)$ is a d_1-dimensional Wiener process. Furthermore, for $t \in [0, T]$, $x \in E_d$, $p \in D$, $n = 0, 1, 2, \ldots$ we are given: $\sigma_t(x)$, $\sigma_t^n(x)$, $\sigma_t(p,x)$ are random matrices of dimension $d \times d_1$ and $b_t(x)$, $b_t^n(x)$, $b_t(p,x)$, ξ_t^n, $\xi_t(p)$ are random d-dimensional vectors which are progressively measurable with respect to $\{\mathscr{F}_t\}$. Assume that for all t, ω, x, y

$$\|\sigma_t(x) - \sigma_t(y)\| + |b_t(x) - b_t(y)| \le K|x - y|. \tag{1}$$

Also, assume that $\sigma_t^n(x)$, $b_t^n(x)$ satisfy (1) for each $n \ge 0$ and, in addition, $\sigma_t(p,x)$, $b_t(p,x)$ satisfy (1) for each $p \in D$.

Suppose that all the processes in question belong to $\mathscr{L}$ for all values of x, n, p. Recall that the space $\mathscr{L}$ was introduced in Section 7. We shall frequently use further on other concepts and results given in Section 7.

We define the processes x_t^x, x_t^n, x_t^p as the solutions of the following equations:

$$x_t^x = x + \int_0^t \sigma_s(x_s^x)\,d\mathbf{w}_s + \int_0^t b_s(x_s^x)\,ds;$$

$$x_t^n = \xi_t^n + \int_0^t \sigma_s^n(x_s^n)\,d\mathbf{w}_s + \int_0^t b_s^n(x_s^n)\,ds;$$

$$x_t^p = \xi_t(p) + \int_0^t \sigma_s(p, x_s^p)\,d\mathbf{w}_s + \int_0^t b_s(p, x_s^p)\,ds.$$

Note that by Theorem 5.7 the above equations have solutions. We also note that by Corollary 5.6 these solutions belong to $\mathscr{L}$. If ξ_t^n, $\xi_t(p) \in \mathscr{L}B$ for all n, p, according to Corollary 5.10 x_t^x, x_t^n, $x_t^p \in \mathscr{L}B$ as well for all n, p, x.

1. Theorem. *Let $\sigma_t^n(x) \to \sigma_t^0(x)$, $b_t^n(x) \to b_t^0(x)$ in $\mathscr{L}$ as $n \to \infty$ for each $x \in E_d$ and let $\xi_t^n \to \xi_t^0$ in $\mathscr{L}$ as $n \to \infty$. Then $x_t^n \to x_t^0$ in $\mathscr{L}$ as $n \to \infty$.*

If $\xi_t^n \to \xi_t^0$ in $\mathscr{L}B$ as well as $n \to \infty$, then $x_t^n \to x_t^0$ in $\mathscr{L}B$ as $n \to \infty$.

PROOF. Let $\hat{\sigma}_t^n(x) = \sigma_t^n(x) - \sigma_t^n(0)$. It is seen that $\hat{\sigma}_t^n(x)$ satisfies the Lipschitz condition (1) and the growth condition $\|\hat{\sigma}_t^n(x)\| \leq K|x|$. Furthermore, $\hat{\sigma}_t^n(x) \to \hat{\sigma}_t^0(x)$ in $\mathscr{L}$ for all x. From this, using Theorem 7.17 and Remark 7.18, we conclude that $\hat{\sigma}_t^n(x_t^0) \to \hat{\sigma}_t^0(x_t^0)$ in $\mathscr{L}$. Adding the last relation with $\sigma_t^n(0) \to \sigma_t^0(0)$ in $\mathscr{L}$, we obtain: $\sigma_t^n(x_t^0) \to \sigma_t^0(x_t^0)$ in $\mathscr{L}$. Similarly, $b_t^n(x_t^0) \to b_t^0(x_t^0)$.

Applying Corollary 5.5 and Theorem 5.9 for $(\tilde{x}_t,\tilde{\sigma}_t,\tilde{b}_t) = (x_t^0,\sigma_t^0,b_t^0)$, $(x_t,\sigma_t,b_t) = (x_t^n,\sigma_t^n,b_t^n)$, we immediately arrive at the assertions of the theorem. □

2. Corollary. *If the process $\xi_t(p)$ is $\mathscr{L}$-continuous ($\mathscr{L}B$-continuous), and, in addition, for each $x \in E_d$ the processes $\sigma_t(p,x)$, $b_t(p,x)$ are $\mathscr{L}$-continuous in p at the point $p_0 \in D$, the process x_t^p is $\mathscr{L}$-continuous ($\mathscr{L}B$-continuous) at the point p_0.*

3. Lemma. *Suppose that for each $t \in [0,T]$, $p \in D$, ω the functions $\sigma_t(p,x)$, $b_t(p,x)$ are linear with respect to x. Let the process $\xi_t(p)$ and for each $x \in E_d$ the processes $\sigma_t(p,x)$ and $b_t(p,x)$ be i times ($\mathscr{L}$-continuously) $\mathscr{L}$-differentiable at the point $p_0 \in D$. Then, the process x_t^p is i times ($\mathscr{L}$-continuously) $\mathscr{L}$-differentiable at p_0. If, in addition, $\xi_t(p)$ is i times ($\mathscr{L}B$-continuously) $\mathscr{L}B$-differentiable at the point p_0, the process x_t^p will be the same as the process $\xi_t(p)$.*

PROOF. Due to the linearity of $\sigma_t(p,x)$, $b_t(p,x)$

$$x_t^p = \xi_t(p) + \sum_{j=1}^d \int_0^t (x_s^p)^j \sigma_s(p,e_j)\,d\mathbf{w}_s + \sum_{j=1}^d \int_0^t (x_s^p)^j b_s(p,e_j)\,ds,$$

where $(x_s^p)^j$ is the jth coordinate of the vector x_s^p in the basis $\{e_j\}$. This implies that the last assertion of the lemma is a corollary of the first assertion as well as the results, which were proved in Section 7, related to the $\mathscr{L}B$-differentiability of integrals and the $\mathscr{L}$-differentiability of products of $\mathscr{L}$-differentiable processes.

We prove the first assertion. To this end, we make use of the induction with respect to i and, in addition, assume that $i = 1$. We take a unit vector $l \in E$ and, in accord with what was said in Section 7, let the processes

$$\mathscr{L}\text{-}\frac{\partial}{\partial l}\,\xi_t(p_0), \qquad \mathscr{L}\text{-}\frac{\partial}{\partial l}\,\sigma_s(p_0,x), \qquad \mathscr{L}\text{-}\frac{\partial}{\partial l}\,b_s(p_0,x)$$

be progressively measurable for $x \in E_d$.

By Corollary 2, we conclude that the process x_t^p is $\mathscr{L}$-continuous at the point p_0. It is not hard to see that for $p = p_0$ the process

$$\eta_t(p) \equiv \mathscr{L}\text{-}\frac{\partial}{\partial l}\,\xi_t(p) + \sum_{j=1}^d \int_0^t (x_s^p)^j \mathscr{L}\text{-}\frac{\partial}{\partial l}\,\sigma_s(p,e_j)\,d\mathbf{w}_s$$
$$+ \sum_{j=1}^d \int_0^t (x_s^p)^j \mathscr{L}\text{-}\frac{\partial}{\partial l}\,b_s(p,e_j)\,ds$$

exists, is progressively measurable (and $\mathscr{L}$-continuous with respect to p, if

$$\mathscr{L}\text{-}\frac{\partial}{\partial l}\xi_t(p), \qquad \mathscr{L}\text{-}\frac{\partial}{\partial l}\sigma_s(p,x), \qquad \mathscr{L}\text{-}\frac{\partial}{\partial l}b_s(p,x)$$

are $\mathscr{L}$-continuous with respect to p). Furthermore, $\eta_t(p_0) \in \mathscr{L}$.

According to Theorem 5.7, the solution of the equation

$$y_t^p = \eta_t(p) + \int_0^t \sigma_s(p,y_s^p)\,d\mathbf{w}_s + \int_0^t b_s(p,y_s^p)\,ds \tag{2}$$

exists and is unique for $p = p_0$.

Let us show that $y_t^p = \mathscr{L}\text{-}(\partial/\partial l)x_t^p$ for $p = p_0$. To this end, we take a sequence $r_n \to 0$ and assume that $y_t^p(n) = r_n^{-1}(x_t^{p+r_nl} - x_t^p)$. It can easily be seen that

$$y_t^p(n) = \eta_t(p,n) + \int_0^t \sigma_s(p + r_nl, y_s^p(n))\,d\mathbf{w}_s + \int_0^t b_s(p + r_nl, y_s^p(n))\,ds, \tag{3}$$

where

$$\eta_t(p,n) = r_n^{-1}[\xi_t(p + r_nl) - \xi_t(p)] + \int_0^t r_n^{-1}[\sigma_s(p + r_nl, x_s^p) - \sigma_s(p,x_s^p)]\,d\mathbf{w}_s$$
$$+ \int_0^t r_n^{-1}[b_s(p + r_nl, x_s^p) - b_s(p,x_s^p)]\,ds.$$

We are given the expression

$$\mathscr{L}\text{-}\lim_{n\to\infty} r_n^{-1}[\xi_t(p_0 + r_nl) - \xi_t(p_0)] = \mathscr{L}\text{-}\frac{\partial}{\partial l}\xi_t(p_0).$$

In addition, since the $\mathscr{L}$-limit of the product (sum) equals the product (sum) of $\mathscr{L}$-limits, we have that in $\mathscr{L}$

$$r_n^{-1}[\sigma_s(p_0 + r_nl, x_s^{p_0}) - \sigma_s(p_0,x_s^{p_0})] = \sum_{j=1}^d (x_s^{p_0})^j r_n^{-1}[\sigma_s(p_0 + r_nl, e_j) - \sigma_s(p_0,e_j)]$$
$$\to \sum_{j=1}^d (x_s^{p_0})^j \mathscr{L}\text{-}\frac{\partial}{\partial l}\sigma_s(p_0,e_j).$$

Similarly, in $\mathscr{L}$

$$r_n^{-1}[b_s(p_0 + r_nl, x_s^{p_0}) - b_s(p_0,x_s^{p_0})] \to \sum_{j=1}^d (x_s^{p_0})^j \mathscr{L}\text{-}\frac{\partial}{\partial l}b_s(p_0,e_j).$$

Thus, $\eta_t(p_0,n) \to \eta_t(p_0)$ in $\mathscr{L}$ as $n \to \infty$. Comparing (2) with (3), we have from Theorem 3 that $y_t^{p_0}(n) \to y_t^{p_0}$ in $\mathscr{L}$. Hence

$$y_t^p = \mathscr{L}\text{-}\frac{\partial}{\partial l}x_t^p \tag{4}$$

for $p = p_0$, proving thereby that x_t^p is $\mathscr{L}$-differentiable.

It is clear that (4) is satisfied at any point p at which there exist $\mathscr{L}$-derivatives $\xi_t(p)$, $\sigma_t(p,x)$, $b_t(p,x)$. Further, if the foregoing derivatives are continuous at the point p_0, they are defined in some neighborhood in which (4) is satisfied. In this case, as we noted above, $\eta_t(p)$ is $\mathscr{L}$-continuous at the point p_0. Also, by Corollary 2, it follows from Eq. (2) that the process y_t^p is $\mathscr{L}$-continuous at the point p_0. This fact implies that the process x_t^p is $\mathscr{L}$-continuously $\mathscr{L}$-differentiable at p_0.

Suppose that our lemma is proved for $i = i_0$ and that the assumptions of the lemma are satisfied for $i = i_0 + 1$. We shall complete proving our lemma if we show that each first $\mathscr{L}$-derivative of x_t^p is i_0 times ($\mathscr{L}$-continuously) $\mathscr{L}$-differentiable at the point p_0. We consider, for instance, $\mathscr{L}$-$(\partial/\partial l)x_t^p$. This process exists and satisfies Eq. (2) for p close to p_0.

Since the assumptions of Lemma 3 are satisfied for $i = i_0$ (even for $i = i_0 + 1$), by the induction assumption, the process x_t^p is i_0 times ($\mathscr{L}$-continuously) $\mathscr{L}$-differentiable at p_0. From this, it follows that the process $\eta_t(p)$ is i_0 times ($\mathscr{L}$-continuously) $\mathscr{L}$-differentiable at p_0. Applying the induction assumption to (2), we convince ourselves that the process y_t^p is i_0 times ($\mathscr{L}$-continuously) $\mathscr{L}$-differentiable at the point p_0. The lemma is proved. □

4. Theorem. *Suppose that the process $\xi_t(p)$ is i times ($\mathscr{L}$-continuously) $\mathscr{L}$-differentiable at a point $p_0 \in D$, and that the functions $\sigma_s(p,x)$, $b_s(p,x)$ for each s, ω are i times continuously (with respect to p, x) differentiable with respect to p, x for $p \in D$, $x \in E_d$. Furthermore, assume that all derivatives of the foregoing functions, up to order i inclusive, do not exceed $K(1 + |x|)^m$ with respect to the norm for any $p \in D$, s, ω, x. Then the process x_t^p is i times ($\mathscr{L}$-continuously) $\mathscr{L}$-differentiable at the point p_0. If, in addition, the process $\xi_t(p)$ is i times ($\mathscr{L}B$-continuously) $\mathscr{L}B$-differentiable at the point p_0, the process x_t^p will be the same as the process $\xi_t(p)$.*

PROOF. Because the notion of the $\mathscr{L}$-derivative is local, it suffices to prove the theorem in any subregion D' of a region D, which together with its closure lies in D. We construct an infinitely differentiable function $w(p)$ in such a way that $w(p) = 1$ for $p \in D'$, $w(p) = 0$ for $p \notin D$. Let $\bar{\xi}_t(p) = \xi_t(p)w(p)$, $\bar{\sigma}_s(p,x) = \sigma_s(p,x)w(p)$, $\bar{b}_s(p,x) = b_s(p,x)w(p)$. Then ξ_t, $\bar{b}_s$, $\bar{\sigma}_s$ satisfy the conditions of the theorem for $D = E$. Further, since the assertions of the theorem hold for $\bar{\xi}$, $\bar{\sigma}$, $\bar{b}$ in E, they hold as well for ξ, b, σ in the region D'. This reasoning shows that in proving our theorem, we can assume that the assumptions of the theorem are satisfied for $D = E$.

In this case we use the induction over i. First, let $i = 1$. Further, we take a unit vector $l \in E$ and a sequence of numbers $r_n \to 0$. Let

$$y_t^p(n) = r_n^{-1}(x_t^{p+r_n l} - x_t^p),$$
$$x_t^p(n,u) = u x_t^{p+r_n l} + (1-u)x_t^p.$$

Using the Newton–Leibniz formula, we easily obtain

$$y_t^p(n) = \eta_t(p,n) + \int_0^t \tilde{\sigma}_s^n(p,y_s^p(n))\,d\mathrm{w}_s + \int_0^t \tilde{b}_s^n(p,y_s^p(n))\,ds, \tag{5}$$

where

$$\eta(p,n) = r_n^{-1}[\xi_t(p + r_n l) - \xi_t(p)] + \int_0^t \left[\sum_j l^j \int_0^1 \sigma_{s,p^j}(p + ur_n l, x_s^p(n,u))\, du\right] d\mathbf{w}_s + \int_0^t \left[\sum_j l^j \int_0^1 b_{s,p^j}(p + ur_n l, x_s^p(n,u))\, du\right] ds,$$

$$\tilde{\sigma}_s^n(p,x) = \sum_{j=1}^d x^j \int_0^1 \sigma_{s,x^j}(p + ur_n l,\ x_s^p(n,u))\, du,$$

$$\tilde{b}_s^n(p,x) = \sum_{j=1}^d x^j \int_0^1 b_{s,x^j}(p + ur_n l, x_s^p(n,u))\, du,$$

We look upon the pair $(p + ur_n l, x_t^p(n,u))$ as a process in $E \times E_d$ with a time parameter t. It is seen that

$$|(p + ur_n l, x_t^p(n,u)) - (p,x_t^p)| \leq |(p + r_n l, x_t^{p+r_n l}) - (p,x_t^p)|.$$

Furthermore, by Corollary 2 and the $\mathscr{L}$-continuity of $\mathscr{L}$-differentiable functions, $x^{p_0+r_n l} \to x_t^{p_0}$ in $\mathscr{L}$. In order to apply Lemma 7.8, we note that, for instance, $|b_{s,x^j}(p,x)| \leq K(1 + \sqrt{|x|^2 + |p|^2})^m$ for all ω, s, p, x. By this lemma, for $p = p_0$

$$\tilde{\sigma}_s^n(p,x) \to \tilde{\sigma}_s(p,x), \qquad \tilde{b}_s^n(p,x) \to \tilde{b}_s(p,x), \qquad \eta_t(p,n) \to \eta_t(p)$$

in the sense of convergence in the space $\mathscr{L}$ in which

$$\tilde{\sigma}_s(p,x) \equiv \sum_{j=1}^d x^j \sigma_{s,x^j}(p,x_s^p),$$

$$\tilde{b}_s(p,x) \equiv \sum_{j=1}^d x^j b_{s,x^j}(p,x_s^p).$$

$$\eta_t(p) \equiv \mathscr{L}\text{-}\frac{\partial}{\partial l}\xi_t(p) + \int_0^t \sum_j l^j \sigma_{s,p^j}(p,x_s^p)\, dw_s + \int_0^t \sum_j l^j b_{s,p^j}(p,x_s^p)\, ds$$

Note that $\tilde{\sigma}_s$, $\tilde{b}_s$, η_t, $\tilde{\sigma}_s^n$, $\tilde{b}_s^n$, $\eta_t(p,n)$ are progressively measurable for those p, x, for which they exist. In fact, one can take the derivative $\mathscr{L}$-$(\partial/\partial l)\xi_t(p)$ to be progressively measurable. Also, for example, $\sigma_{s,x^j}(p,x)$ is progressively measurable (ordinary derivative with respect to the parameter of a progressively measurable process) and continuous with respect to p, x. Hence the process $\sigma_{s,x^j}(p + ur_n l, x_s^p(n,u))$ is progressively measurable and continuous with respect to u, which, in turn, implies the progressive measurability of the Riemann integral

$$\int_0^1 \sigma_{s,x^j}(p + ur_n l, x_s^p(n,u))\, du$$

and the progressive measurability of the process $\tilde{\sigma}_s^n(p,x)$.

Further, since $\sigma_s(p,x)$, $b_s(p,x)$ satisfy the Lipschitz condition (1) with respect to x, $\sigma_{s,x^j}(p,x)$, $b_{s,x^j}(p,x)$ are bounded variables. This implies that the functions $\tilde{\sigma}_s(p,x)$ and $\tilde{b}_s(p,s)$, linear with respect to x, satisfy the Lipschitz condition (1). By Theorem 5.7, for $p = p_0$ there exists a solution of the equation

$$y_t^p = \eta_t(p) + \int_0^t \tilde{\sigma}_s(p,y_s^p)\,d\mathbf{w}_s + \int_0^t \tilde{b}_s(p,y_s^p)\,ds. \tag{6}$$

By Theorem 1, comparing (5) with (6), we conclude that

$$\mathscr{L}\text{-}\lim_{n\to\infty} r_n^{-1}(x_t^{p+r_n l} - x_t^p) = \mathscr{L}\text{-}\lim_{n\to\infty} y_t^p(n) = y_t^p$$

for $p = p_0$. This shows that $y_t^p = \mathscr{L}\text{-}(\partial/\partial l)x_t^p$ for $p = p_0$, and therefore the process x_t^p is $\mathscr{L}$-differentiable at the point p_0. It is also seen that $y_t^p = \mathscr{L}\text{-}(\partial/\partial l)x_t^p$ at each point p at which $\mathscr{L}\text{-}(\partial/\partial l)\xi_t(p)$ exists.

Next, let $\xi_t(p)$ be $\mathscr{L}$-continuously $\mathscr{L}$-differentiable at the point p_0. Then $\mathscr{L}\text{-}(\partial/\partial l)\xi_t(p)$ exists in some neighborhood of the point p_0, y_t^p being the $\mathscr{L}$-derivative of x_t^p along the l direction in this neighborhood. In addition, the process (p,x_t^p) is $\mathscr{L}$-continuous at the point p_0, and the functions $\sigma_{s,p^j}(p,x)$, $b_{s,p^j}(p,x)$ are continuous with respect to (p,x) and do not exceed $K(1 + |x|)^m$ with respect to the norm. Therefore, by Theorem 7.9, the processes $\sigma_{s,p^j}(p,x_s^p)$, $b_{s,p^j}(p,x_s^p)$ are $\mathscr{L}$-continuous, and consequently, the process $\eta_t(p)$ is $\mathscr{L}$-continuous at p_0. Similarly, the fact that the functions $\sigma_{s,x^j}(p,x)$, $b_{s,x^j}(p,x)$ are bounded and continuous with respect to p, x, implies that the processes $\tilde{\sigma}_s(p,x)$ and $\tilde{b}_s(p,x)$ are $\mathscr{L}$-continuous at p_0 for each x. To conclude our reasoning, we observe that, by Corollary 2, the process y_t^p interpreted as the solution of Eq. (6) is $\mathscr{L}$-continuous at the point p_0.

Thus, we have proved the first assertion of the theorem for $i = 1$. Further, suppose that this theorem has been proved for $i = i_0$, and, in addition, the assertions of the theorem are satisfied for $i = i_0 + 1$. Consider the derivative $\mathscr{L}\text{-}(\partial/\partial l)x_t^p$. As was shown above, we may assume that this process is y_t^p and that it satisfies Eq. (6). By the induction assumption, x_t^p is i_0 times $\mathscr{L}$-differentiable at p_0. Therefore the pair (p,x_t^p) is i_0 times $\mathscr{L}$-differentiable as well. By Theorem 7.9, the processes $\sigma_{s,p^j}(p,x_s^p)$, $b_{s,p^j}(p,x_s^p)$, $\sigma_{s,x^j}(p,x_s^p)$, $b_{s,x^j}(p,x_s^p)$ are i_0 times $\mathscr{L}$-differentiable at the point p_0. Hence, in Eq. (6) the processes $\eta_t(p)$, $\tilde{\sigma}_s(p,x)$, $\tilde{b}_s(p,x)$ are i_0 times $\mathscr{L}$-differentiable with respect to p. Since $\tilde{\sigma}_s(p,x)$, $\tilde{b}_s(p,x)$ are linear functions of x, according to the preceding lemma the process y_t^p is i_0 times $\mathscr{L}$-differentiable at the point p_0. We have thus proved that the derivative $(\partial/\partial l)x_t^p$ is i_0 times $\mathscr{L}$-differentiable at the point p_0. Since l is an arbitrary unit vector from E, this implies, by definition, that x_t^p is $i_0 + 1$ times $\mathscr{L}$-differentiable at the point p_0.

In addition, if $\xi_t(p)$ is $i_0 + 1$ times $\mathscr{L}$-continuously $\mathscr{L}$-differentiable at the point p_0, we can prove that x_t^p is $i_0 + 1$ times $\mathscr{L}$-continuously $\mathscr{L}$-differentiable at the point p_0 if we put the word "$\mathscr{L}$-continuously" in the appropriate places in the above arguments. This completes the proof of the first assertion of Theorem 4.

For proving the second assertion of the theorem, we need only to prove, due to the equality

$$x_t^p = \xi_t(p) + \int_0^t \sigma_s(p,x_s^p)\, d\mathbf{w}_s + \int_0^t b_s(p,x_s^p)\, ds,$$

that the processes $\sigma_s(p,x_s^p)$, $b_s(p,x_s^p)$ are i times ($\mathscr{L}$-continuously) $\mathscr{L}$-differentiable at the point p_0. It is obvious that a process which is identically equal to $(p,0)$ is i times $\mathscr{L}$-continuously $\mathscr{L}$-differentiable. It is also seen that, since the function $\sigma_s(p,0)$ is i times continuously differentiable with respect to p and, in addition, the derivatives of this function are bounded, the process $\sigma_s(p,0)$ is i times $\mathscr{L}$-continuously $\mathscr{L}$-differentiable in accord with Theorem 7.9. Furthermore, the process (p,x_s^p) is i times ($\mathscr{L}$-continuously) $\mathscr{L}$-differentiable at the point p_0, the function $\sigma_s(p,x) - \sigma_s(p,0)$, with respect to the norm, does not exceed $K|x|$, and, in addition, the derivatives of this function satisfy the necessary restrictions on the growth. By Theorem 7.9, the process $\sigma_s(p,x_s^p) - \sigma_s(p,0)$ is i times ($\mathscr{L}$-continuously) $\mathscr{L}$-differentiable at the point p_0; the same holds for the process $\sigma_s(p,x_s^p) = \sigma_s(p,0) + [\sigma_s(p,x_s^p) - \sigma_s(p,0)]$. The process in $b_s(p,x_s^p)$ can be considered in a similar way. The theorem is proved. □

5. Remark. For $i \geq 1$ we have proved that for any unit vector $l \in E$ the solution of Eq. (6) is the $\mathscr{L}$-derivative of x_t^p along the l direction:

$$y_t^p = \mathscr{L}\text{-}\frac{\partial}{\partial l}\, \xi_t(p) + \int_0^t \sigma_{s,(l)}(p,x_s^p)\, d\mathbf{w}_s$$
$$+ \int_0^t b_{s,(l)}(p,x_s^p)\, ds + \int_0^t \sigma_{s,(y_s^p)}(p,x_s^p)|y_s^p|\, d\mathbf{w}_s$$
$$+ \int_0^t b_{s,(y_s^p)}(p,x_s^p)|y_s^p|\, ds.$$

We have seen that the last equation is linear with respect to y_s^p; also, we applied Lemma 3 to this equation for $i \geq 2$. In Lemma 3 we derived Eq. (2), according to which the solution of the equation which follows is an $\mathscr{L}$-derivative of y_t^p along the l direction, that is, a second $\mathscr{L}$-derivative of x_t^p along the l direction. This equation is the following:

$$z_t^p = \eta_t(p) + \int_0^t \sigma_{s,(z_s^p)}(p,x_s^p)|z_s^p|\, d\mathbf{w}_s + \int_0^t b_{s,(z_s^p)}(p,x_s^p)|z_s^p|\, ds,$$

where, according to the rules of $\mathscr{L}$-differentiation of a composite function (see (2)),

$$\eta_t(p) = \mathscr{L}\text{-}\frac{\partial}{\partial l}\left[\mathscr{L}\text{-}\frac{\partial}{\partial l}\, \xi_t(p) + \int_0^t \sigma_{s,(l)}(p,x_s^p)\, d\mathbf{w}_s + \int_0^t b_{s,(l)}(p,x_s^p)\, ds\right]$$
$$+ \sum_{j=1}^d \int_0^t (y_s^p)^j[\sigma_{s,x^j(l)}(p,x_s^p) + \sigma_{s,x^j(y_s^p)}(p,x_s^p)|y_s^p|]\, d\mathbf{w}_s$$
$$+ \sum_{j=1}^d \int_0^t (y_s^p)^j[b_{s,x^j(l)}(p,x_s^p) + b_{s,x^j(y_s^p)}(p,x_s^p)|y_s^p|]\, ds.$$

Note that the above equations as well as equations for the highest $\mathscr{L}$-derivatives of x_t^p can be obtained proceeding from the fact that x_t^p is $\mathscr{L}$-differentiable the desired number of times, if we differentiate the equality

$$x_t^p = \xi_t(p) + \int_0^t \sigma_s(p,x_s^p)\, d\mathbf{w}_s + \int_0^t b_s(p,x_s^p)\, ds,$$

interchange the order of the derivatives with those of the integrals, and, in addition, make use of the formula for an $\mathscr{L}$-derivative of a composite function.

The following assertion is a simple consequence of Theorem 4 and Corollary 2 in the case where $D = E_d$, $\xi_t(p) \equiv p$, $\sigma_t(p,x) = \sigma_t(x)$, $b_t(p,x) = b_t(x)$.

6. Theorem. *The process x_t^x is $\mathscr{L}B$-continuous. If $\sigma_s(x)$, $b_s(x)$ are i times continuously differentiable with respect to x for all ω, s, and if, in addition, each derivative of these functions up to order i inclusively does not exceed $K(1 + |x|)^m$ with respect to the norm for any s, x, ω, the process x_t^x is i times $\mathscr{L}B$-continuously $\mathscr{L}B$-differentiable.*

In concluding this section, we give two theorems on estimation of moments of derivatives of a solution of a stochastic equation. Since, as we saw in Remark 5, it was possible to write equations for such derivatives, it is reasonable to apply Corollaries 5.6 and 5.10–5.12 for estimating the moments of these derivatives. The reader can easily prove the theorems which follow.

7. Theorem. *Let there be a constant K_1 such that for all s, x, p, ω*

$$|b_s(p,x)| + \|\sigma_s(p,x)\| \le K_1(1 + |x|).$$

Suppose that the process $\xi_t(p)$ is $\mathscr{L}B$-differentiable at a point $p_0 \in D$. Further, suppose that $\mathscr{L}B$-derivatives of the process $\xi_t(p)$ have modifications which are progressively measurable and separable at the same time. Let the functions $\sigma_s(p,x)$, $b_s(p,x)$ for each s, ω be continuously differentiable with respect to p, x for $p \in D$, $x \in E_d$. In addition, let the matrix norms of the derivatives of the function $\sigma_s(p,x)$ and the norms of the derivatives of the function $b_s(p,x)$ be smaller than $K(1 + |x|)^m$ $(m \ge 1)$ along all directions for all $p \in D$, s, ω, x. Then for any unit vector $l \in E$, $q \ge 1$, $t \in [0,T]$

$$\mathsf{M} \sup_{s \le t} \left| \mathscr{L}B\text{-}\frac{\partial}{\partial l} x_s^{p_0} \right|^{2q} \le N e^{Nt} \left(1 + \mathsf{M} \sup_{s \le t} \left| \mathscr{L}B\text{-}\frac{\partial}{\partial l} \xi_s(p_0) \right|^{2q} + \mathsf{M} \int_0^t |\xi_s(p_0)|^{2qm}\, ds \right),$$

where $N = N(q,K,m,K_1)$.

8. Theorem. (a) *Let the functions $\sigma_s(x)$, $b_s(x)$ be continuously differentiable with respect to x for each s, ω. Then for any unit vector $l \in E_d$, $q \ge 1$, $t \in [0,T]$, $x \in E_d$*

$$\mathsf{M} \sup_{s \le t} \left| \mathscr{L}B\text{-}\frac{\partial}{\partial l} x_s^x \right|^q \le N e^{Nt},$$

where $N = (q,K)$.

(b) *Let the functions $\sigma_s(x)$, $b_s(x)$ be twice continuously differentiable for each s, ω. Further, for each x, s, ω and unit vectors $l \in E_d$ let*

$$\|\sigma_{s(l)(l)}(x)\| + |b_{s(l)(l)}(x)| \leq K(1 + |x|)^m.$$

Also, suppose that $\|\sigma_s(x)\| + |b_s(x)| \leq K_1(1 + |x|)$ for all x, s, ω for some constant K_1. Then for any $q \geq 1$, $t \in [0,T]$, $x \in E_d$ and the unit vector $l \in E_d$

$$\mathsf{M} \sup_{s \leq t} \left| \mathscr{L}B\text{-}\frac{\partial^2}{\partial l^2} x_s^x \right|^q \leq N(1 + |x|)^{qm} e^{Nt},$$

where $N = N(q,K,m,K_1)$.

9. The Markov Property of Solutions of Stochastic Equations

The Markov property of solutions of a stochastic equation with non random coefficients is well known (see [9, 11, 24]). In this section, we shall prove a similar property for random coefficients of the equation (Theorem 4), and moreover, deduce some consequences from this property.

We fix two constants T, $K > 0$. In this section we repeatedly assume about $(\mathbf{w}_t,\mathscr{F}_t)$, ξ_t, $\sigma_t(x)$, $b_t(x)$, with indices and tildes or without them, the following: $(\mathbf{w}_t,\mathscr{F}_t)$ is a d_1-dimensional Wiener process, $\sigma_t(x)$ is a random matrix of dimension $d \times d_1$, $b_t(x)$, ξ_t are random d-dimensional vectors; $\sigma_t(x)$, $b_t(x)$, ξ_t are defined for $t \in [0,T]$, $x \in E_d$, progressively measurable with respect to $\{\mathscr{F}_t\}$, and

$$\mathsf{M} \int_0^T [|\xi_t|^2 + \|\sigma_t(x)\|^2 + |b_t(x)|^2]\, dt < \infty,$$

$$\|\sigma_t(x) - \sigma_t(y)\| + |b_t(x) - b_t(y)| \leq K|x - y|$$

for all possible values of the indices and arguments.

We can now specify the objective of this section. It consists in deriving formulas for a conditional expectation under the condition $\mathscr{F}_0$ of functionals of solutions of the stochastic equation

$$x_t = \xi_t + \int_0^t \sigma_s(x_s)\, d\mathbf{w}_s + \int_0^t b_s(x_s)\, ds. \tag{1}$$

Note that if the assumptions made above are satisfied, in accord with Theorem 5.7 the solution of Eq. (1) on an interval $[0,T]$ exists and is unique.

1. Lemma. *Suppose that for all integers $i, j > 0$, $t_1, \ldots, t_i \in [0,T]$, $z_1, \ldots, z_j \in E_d$ the vector*

$$\{\mathbf{w}_{t_p}, \xi_{t_p}, \sigma_{t_p}(z_q), b_{t_p}(z_q) : p = 1, \ldots, i,\ q = 1, \ldots, j\}$$

does not depend on $\mathscr{F}_0$. Then the process x_t, which is a solution of Eq. (1), does not depend on $\mathscr{F}_0$ either.

PROOF. As we did in proving Theorem 5.7, we introduce here an operator I using the formula

$$Iy_t = \int_0^t \sigma_s(y_s)\,d\mathbf{w}_s + \int_0^t b_s(y_s)\,ds.$$

In proving Theorem 5.7 we said that the operator I is defined on a set of progressively measurable functions in $\mathscr{L}_2([0,T] \times \Omega)$ and also that this operator maps this set into itself.

Let a function $y_t(\omega)$ from the set indicated (for example, $y_t \equiv 0$) be such that the totality of random variables

$$\{\mathbf{w}_t,\xi_t,y_t,\sigma_t(x),b_t(x):t \in [0,T], x \in E_d\} \tag{2}$$

does not depend on $\mathscr{F}_0$. We prove that in this case the totality of random variables

$$\{\mathbf{w}_t, \xi_t + Iy_t, \sigma_t(x), b_t(x):t \in [0,T], x \in E_d\} \tag{3}$$

does not depend on $\mathscr{F}_0$ either.

We denote by Σ the completion of a σ-algebra of subsets Ω, which is generated by the totality of random variables (2). By assumption, Σ does not depend on $\mathscr{F}_0$. It is seen that for proving that (3) is independent of $\mathscr{F}_0$, it suffices to prove that random variables Iy_t are Σ-measurable for $t \in [0,T]$.

For real a let $x_n(a) = 2^{-n}[2^n a]$, where $[a]$ is the greatest integer less than or equal to a. If $y \in E_d$, we assume that $x_n(y) = (x_n(y^1), \dots, x_n(y^d))$, and, in addition, that Γ_n is a set of values of the function $x_n(y)$, $y \in E_d$. Due to the continuity of $\sigma_t(x)$ with respect to x we have

$$\sigma_t(y_t) = \lim_{n\to\infty} \sigma_t(\kappa_n(y_t)) = \lim_{n\to\infty} \sum_{y \in \Gamma_n} \sigma_t(y)\chi_{\kappa_n}(y_t) = y. \tag{4}$$

Therefore, the variable $\sigma_t(y_t)$ is Σ-measurable. The Σ-measurability of $b_t(y_t)$ can be proved in a similar way. Further (see Appendix 1), for almost all $s \in [0,1]$ for some sequence of integers n' in probability

$$\lim_{n'\to\infty}\left[\int_0^t \sigma_{\kappa_{n'}(r+s)-s}(y_{\kappa_{n'}(r+s)-s})\,d\mathbf{w}_r + \int_0^t b_{\kappa_{n'}(r+s)-s}(y_{\kappa_{n'}(r+s)-s})\,dr\right] = Iy_t. \tag{5}$$

Since the function $\kappa_n(r + s) - s$ assumes only a finite number of values on an interval $[0,t]$, the integrals in a limiting expression are integrals of step functions. The former integrals are to be written as finite sums which consist of the product of values of $\sigma_r(y_r)$ and an increment $\mathbf{w}_r$ and the product of values of $b_r(y_r)$ and increments r. The foregoing sums are Σ-measurable. Hence the limiting expressions are Σ-measurable, which implies the Σ-measurability of Iy_t.

As we did in proving Theorem 5.7, we define here the sequence x_t^n using the recurrence formula

$$x_t^0 \equiv 0, \qquad x_t^{n+1} = \xi_t + Ix_t^n, \qquad n \geq 0.$$

By induction, it follows from what has been proved above that the processes x_t^n do not depend on $\mathscr{F}_0$ for $n \geq 0$, $t \in [0,T]$. According to Remark 5.13, for $t \in [0,T]$

$$\underset{n \to \infty}{\text{l.i.m.}}\; x_t^n = x_t.$$

Therefore, the process x_t does not depend on $\mathscr{F}_0$. The lemma is proved. □

In the next lemma we consider $(\tilde{\mathbf{w}}_t, \tilde{\mathscr{F}}_t)$, $\tilde{\xi}_t$, $\tilde{\sigma}_t(x)$, $\tilde{b}_t(x)$ as well as $(\mathbf{w}_t, \mathscr{F}_t)$, ξ_t, $\sigma_t(x)$, $b_t(x)$. As we agreed above, we assume here that these elements satisfy the same conditions. Let $\tilde{x}_t$ be a solution of the equation

$$\tilde{x}_t = \tilde{\xi}_t + \int_0^t \tilde{\sigma}_r(\tilde{x}_r)\, d\tilde{\mathbf{w}}_r + \int_0^t \tilde{b}_r(\tilde{x}_r)\, dr.$$

2. Lemma. *Suppose that for all integers $i, j > 0$ and $t_1, \ldots, t_i \in [0,T]$, $z_1, \ldots, z_j \in E_d$ the following vectors are identically distributed:*

$$\{\mathbf{w}_{t_p}, \xi_{t_p}, \sigma_{t_p}(z_q), b_{t_p}(z_q) : p = 1, \ldots, i, q = 1, \ldots, j\},$$
$$\{\tilde{\mathbf{w}}_{t_p}, \tilde{\xi}_{t_p}, \tilde{\sigma}_{t_p}(z_q), \tilde{b}_{t_p}(z_q) : p = 1, \ldots, i, q = 1, \ldots, j\}.$$

Then the finite-dimensional distribution of the process x_t is equivalent to that of the process $\tilde{x}_t$.

Proof. We make use again of the operator I from the previous proof. Let

$$\tilde{I}\tilde{y}_t = \int_0^t \tilde{\sigma}_s(\tilde{y}_s)\, d\tilde{\mathbf{w}}_s + \int_0^t \tilde{b}_s(\tilde{y}_s)\, ds$$

and let the processes y_t, $\tilde{y}_t$ be progressively measurable with respect to $\{\mathscr{F}_t\}$, $\{\tilde{\mathscr{F}}_t\}$, respectively:

$$\mathsf{M} \int_0^t |y_t|^2\, dt < \infty, \qquad \mathsf{M} \int_0^t |\tilde{y}_t|^2\, dt < \infty.$$

Further, for any $i, j > 0$, $t_1, \ldots, t_i \in [0,T]$, $z_1, \ldots, z_j \in E_d$ let the vectors

$$\begin{gathered}\{\mathbf{w}_{t_p}, \xi_{t_p}, y_{t_p}, \sigma_{t_p}(z_q), b_{t_p}(z_q) : p = 1, \ldots, i, q = 1, \ldots, j\},\\ \{\tilde{\mathbf{w}}_{t_p}, \tilde{\xi}_{t_p}, \tilde{y}_{t_p}, \tilde{\sigma}_{t_p}(z_q), \tilde{b}_{t_p}(z_q) : p = 1, \ldots, i, q = 1, \ldots, j\}\end{gathered} \tag{6}$$

have identical distributions. Note that if two random vectors have identical distributions, any (Borel) function of one vector has the same distribution as the other has. From this it follows, in accord with Eq. (4), that for any $i, j > 0$, $t_1, \ldots, t_i \in [0,T]$, $z_1, \ldots, z_j \in E_d$ the vectors

$$\begin{gathered}\{\mathbf{w}_{t_p}, \xi_{t_p}, y_{t_p}, \sigma_{t_p}(y_{t_p}), \sigma_{t_p}(z_q), b_{t_p}(y_{t_p}), b_{t_p}(z_q) : p = 1, \ldots, i, q = 1, \ldots, j\},\\ \{\tilde{\mathbf{w}}_{t_p}, \tilde{\xi}_{t_p}, \tilde{y}_{t_p}, \tilde{\sigma}_{t_p}(\tilde{y}_{t_p}), \tilde{\sigma}_{t_p}(z_q), \tilde{b}_{t_p}(\tilde{y}_{t_p}), \tilde{b}_{t_p}(z_q) : p = 1, \ldots, i, q = 1, \ldots, j\}\end{gathered} \tag{7}$$

have the same distributions. It is useful to draw the reader's attention to the fact that in order to prove the proposition made above, we need to use vectors of type (6) at the values of z_q different from those which appear in (7).

We choose $s \in [0,1]$ so that Eq. (5) holds for $t = t_1, \dots, t_i$, and, in addition, that similar representations hold for $\tilde{I}$. Having done this, we can see that the vectors

$$\begin{gathered}\{\mathbf{w}_{t_p}, \xi_{t_p}, \xi_{t_p} + Iy_{t_p}, \sigma_{t_p}(z_q), b_{t_p}(z_q) : p = 1, \dots, i, q = 1, \dots, j\},\\ \{\tilde{\mathbf{w}}_{t_p}, \tilde{\xi}_{t_p}, \tilde{\xi}_{t_p} + \tilde{I}\tilde{y}_{t_p}, \tilde{\sigma}_{t_p}(z_q), \tilde{b}_{t_p}(z_q) : p = 1, \dots, i, q = 1, \dots, j\}\end{gathered} \tag{8}$$

are representable as the limits in probability of identical functions of vectors of type (7). Therefore, the vectors (8) have identical distributions for any $i, j > 0$, $t_1, \dots, t_i \in [0,T]$, $z_1, \dots, z_j \in E_d$.

Next, we compare the vectors (6) and (8). Also, we find sequences of the processes

$$x_t^0 \equiv 0, \qquad \tilde{x}_t^0 \equiv 0, \qquad x_t^{n+1} = \xi_t + Ix_t^n, \qquad \tilde{x}_t^{n+1} = \tilde{\xi}_t + \tilde{I}\tilde{x}_t^n.$$

Passing from vectors of type (6) to vectors of type (8), we prove by induction that the finite-dimensional distribution of x_t^n is equivalent to that of $\tilde{x}_t^n$. Therefore, the finite-dimensional distributions of the limits of these processes in the mean square coincide, i.e., x_t and $\tilde{x}_t$. The lemma is proved. □

3. Corollary. *If ξ_t, $\sigma_t(x)$, $b_t(x)$ are nonrandom and if, in addition, they are equal to $\tilde{\xi}_t$, $\tilde{\sigma}_t(x)$, $\tilde{b}_t(x)$, respectively, for all $t \in [0,T]$, $x \in E_d$, the processes x_t, $\tilde{x}_t$ have identical finite-dimensional distributions. Furthermore, the process x_t does not depend on $\mathscr{F}_0$, and the process $\tilde{x}_t$ does not depend on $\tilde{\mathscr{F}}_0$.*

This corollary follows from Lemmas 1 and 2 and the fact that all Wiener processes have identical finite-dimensional distributions and that, for example, $\mathbf{w}_t = \mathbf{w}_t - \mathbf{w}_0$ does not depend on $\mathscr{F}_0$.

The formula mentioned at the beginning of the section can be found in the next theorem. In order not to complicate the formulation of the theorem, we list the conditions under which we shall prove the theorem.

Let Z be a separable metric space with metric ρ and let $(\mathbf{w}_t^z, \mathscr{F}_t^z) \equiv (\mathbf{w}_t, \mathscr{F}_t)$, $\sigma_t^z(x)$, $b_t^z(x)$ be defined for $z \in Z$. We assume (in addition to the assumption mentioned at the beginning of the section) that the functions $\sigma_t^z(x,\omega)$, $b_t^z(x,\omega)$ are continuous with respect to z for all t, ω, x and

$$\mathsf{M} \int_0^T \left[\sup_z \|\sigma_t^z(x)\|^2 + \sup_z |b_t^z(x)|^2\right] dt < \infty$$

for all x.

4. Theorem. *Suppose that the assumptions made before proving the theorem are satisfied. Let the totality of variables*

$$\{\mathbf{w}_t, \sigma_t^z(x), b_t^z(x) : t \in [0,T], x \in E_d\}$$

be independent of $\mathscr{F}_0$ for all $z \in Z$. Further, let ξ be an $\mathscr{F}_0$-measurable random variable with values in E_d and a finite second moment, let ζ be an $\mathscr{F}_0$-measurable *random function with values in Z. Finally, let y_t be a solution of the equation*

$$y_t = \xi + \int_0^t \sigma_r^{\zeta}(y_r)\,d\mathbf{w}_r + \int_0^t b_r^{\zeta}(y_r)\,dr. \tag{9}$$

We denote by $x_t^{z,x}$ a solution of the equation

$$x_t^{z,x} = x + \int_0^t \sigma_r^{z}(x_r^{z,x})\,d\mathbf{w}_r + \int_0^t b_r^{z}(x_r^{z,x})\,dr. \tag{10}$$

Let $F(z,x_{[0,T]})$ be a nonnegative measurable function on $Z \times C([0,T],E_d)$. Then

$$\mathsf{M}\{F(\zeta,y_{[0,T]})|\mathscr{F}_0\} = \Phi(\zeta,\xi) \qquad \text{(a.s.)}, \tag{11}$$

where

$$\Phi(z,x) \equiv \mathsf{M}F(z,x_{[0,T]}^{z,x}).$$

Proof. First we note that due to the conditions imposed, Eq. (9) and Eq. (10) are solvable and, in addition, are continuous with respect to t. Further, it suffices to prove Eq. (11) for functions of the form $F(z,x_{t_1}, \ldots ,x_{t_n})$, where $t_1, \ldots , t_n \in [0,T]$ and $F(z,x_1, \ldots ,x_n)$ is a bounded continuous function of $(z,x_1, \ldots ,x_n)$. In fact, in this case Eq. (11) extends in a standard manner to all nonnegative functions $F(z,x_{[0,T]})$, which are measurable with respect to a product of a σ-algebra of Borel sets in z and the smallest σ-algebra which contains cylinder sets of the space $C([0,T],E_d)$. It is a well-known fact that the latter σ-algebra is equivalent to the σ-algebra of Borel sets of the metric space $C([0,T],E_d)$.

In future, we shall consider functions F only of the type indicated. Let $A = \{z^{(i)}; i \geq 1\}$ be a countable everywhere dense subset in Z. For $z \in Z$ we denote by $\bar{\kappa}_n(z)$ the first member of the sequence $\{z^{(i)}\}$ for which $\rho(z,z^{(i)} \leq 2^{-n}$. It is easily seen that $\bar{\kappa}_n(z)$ is the measurable function of z and that $\rho(z,\bar{\kappa}_n(z)) \leq 2^{-n}$ for all $z \in Z$. In addition, we define the function $\kappa_n(x)$ as in the proof of Lemma 1.

By Lemma 1, almost surely

$$\begin{aligned}
&\mathsf{M}\,\{F(\bar{\kappa}_n(\zeta), x_{[0,T]}^{\bar{\kappa}_n(\zeta),\kappa_n(\xi)}|\mathscr{F}_0\} \\
&\quad = \sum_{z\in A}\sum_{x\in\Gamma_n} \chi_{\bar{\kappa}_n(\zeta)=z,\kappa_n(\xi)=x}\mathsf{M}\,\{F(z,x_{[0,T]}^{z,x})|\mathscr{F}_0\} \\
&\quad = \sum_{z\in A}\sum_{x\in\Gamma_n} \chi_{\bar{\kappa}_n(\zeta)=z,\kappa_n(\xi)=x}\Phi(z,x) = \Phi(\bar{\kappa}_n(\zeta),_{\kappa_n}(\xi))
\end{aligned} \tag{12}$$

where we take the limit as $n \to \infty$. We agreed to consider only bounded continuous functions $F(z,x_{[0,T]})$ (moreover, of special type). Hence, the left side of (12) yields the left side of (11), if we show that for some subsequence $\{n'\}$

$$\mathsf{P}\left\{\lim_{n'\to\infty}\sup_{t\leq T}|x_t^{\bar{\kappa}_{n'}(\zeta),\kappa_{n'}(\xi)} - y_t| = 0\right\} = 1. \tag{13}$$

In this case the right side of (12) yields the right side of (11) if we prove that $\Phi(z,x)$ is a continuous function of (z,x).

Since the variables $\bar{\kappa}_n(\zeta)$, $\kappa_n(\xi)$ are $\mathscr{F}_0$-measurable, we can bring an indicator of the set $\{\bar{\kappa}_n(\zeta) = z, \kappa_n(\xi) = x\}$ under the sign of a stochastic integral. Multiplying (10) by the indicator of the above set, bringing this indicator under the integral signs, replacing the values z, x by values $\bar{\kappa}_n(\zeta)$, $\kappa_n(\xi)$, which are equal to z, x on the set considered, and, finally, bringing the indicator out, we have that on each set $\{\bar{\kappa}_n(\zeta) = z, \kappa_n(\xi) = x\}$ the process $x^{\kappa_n(\zeta),\kappa_n(\xi)}$ satisfies the equation

$$x_t = \kappa_n(\xi) + \int_0^t \sigma_r^{\bar{\kappa}_n(\zeta)}(x_r)\,d\mathbf{w}_r + \int_0^t b_r^{\bar{\kappa}_n(\zeta)}(x_r)\,dr. \tag{14}$$

The combination of the sets $\{\bar{\kappa}_n(\zeta) = z, \kappa_n(\xi) = x\}$ with respect to $z \in A$, $x \in \Gamma_n$ produces all Ω. Hence $x^{\kappa_n(\zeta)\kappa_n(\xi)}$ satisfies Eq. (14) on Ω. Comparing (9) with (14), we have in accord with Theorem 5.9 that

$$\begin{aligned} \mathsf{M} \sup_{t \le T} |x_t^{\bar{\kappa}_n(\zeta),\kappa_n(\xi)} - y_t|^2 \le N\mathsf{M}|\xi - \kappa_n(\xi)|^2 \\ + N\mathsf{M} \int_0^T \big[|b_t^{\bar{\kappa}_n(\zeta)}(y_t) - b_t^{\zeta}(y_t)|^2 \\ + \|\sigma^{\bar{\kappa}_n(\zeta)}(y_t) - \sigma_t^{\zeta}(y_t)\|^2\big]\,dt. \end{aligned}$$

Here $|\xi - \kappa_n(\xi)| \to 0$ uniformly on Ω, $b_t^{\bar{\kappa}_n(\zeta)}(y_t) \to b_t^{\zeta}(y_t)$ for each t, due to continuity of $b_t^z(x)$ with respect to z. Furthermore, $|b_t^{\bar{\kappa}_n(\zeta)}(y_t)|^2 + |b_t^{\zeta}(y_t)|^2$ does not exceed $4 \sup_z |b_t^z(0)|^2 + 4K^2|y_t|^2$.

The last expression is summable over $d\mathsf{P} \times dt$. Investigating $\sigma_t^z(x)$ in a similar way, we conclude using the Lebesgue theorem that

$$\mathsf{M} \sup_{t \le T} |x_t^{\bar{\kappa}_n(\zeta),\kappa_n(\xi)} - y_t|^2 \to 0.$$

This implies (13). For proving the continuity of $\Phi(z,x)$ with respect to (z,x) it suffices to prove that for any sequence $(z_n,x_n) \to (z,x)$ there is a subsequence $(z_{n'},x_{n'})$ for which $\Phi(z_{n'},x_{n'}) \to \Phi(z,x)$. From a form of $\Phi(z,x)$ we easily find that it is enough to have

$$\mathsf{P}\left\{\lim_{n' \to \infty} \sup_{t \le T} |x_t^{z_{n'},x_{n'}} - x_t^{z,x}| = 0\right\} = 1.$$

The existence of such a subsequence $\{n'\}$ for any sequence (z_n,x_n) converging to (z,x) follows from the considerations which are very similar to the preceding considerations concerning Eq. (13). The theorem is proved. □

5. Remark. The function $\mathsf{M}F(z,x_{[0,T]}^{z,x})$ is measurable with respect to (z,x).

Indeed, the set of functions $F(z,x_{[0,T]})$ for which $\Phi(z,x)$ is measurable contains all continuous and bounded functions F. For these functions F, $\Phi(z,x)$ is continuous even with respect to (z,x). From this we derive in a usual way that the set mentioned contains all nonnegative Borel functions $F(z,x_{[0,T]})$.

6. Exercise

Prove that the assumptions of Theorem 4 about the finiteness of

$$\mathsf{M}\int_0^T\left[\sup_z\|\sigma_t^z(x)\|^2+\sup_z|b_t^z(x)|^2\right]dt$$

can be weakened, and that it is possible to require instead uniform integrability of the values $\|\sigma_t^z(0)\|^2$, $|b_t^z(0)|^2$ over $d\mathsf{P}\times dt$ for z which run through each bounded subset Z.

Further, we consider the problem of computing a conditional expectation under the condition $\mathscr{F}_s$, where $s\in[0,T]$. We shall reduce this problem to that of computing a conditional expectation under the condition $\tilde{\mathscr{F}}_0$ using a time shift. If the function $F(x_{[0,T-s]})$ is defined on $C([0,T-s],E_d)$ and $x_{[0,T-s]}\in C([0,T-s],E_d)$, we denote by $F(x_{[s,T]})$ a value of F on the function $\theta_s x$ which is given by the formula $(\theta_s x)_t=x_{t+s}$ for $t\in[0,T-s]$. Sometimes $F(x_{[s,T]})$ is written as $\theta_s F(x_{[0,T-s]})$. Similar notation can be used for the functions $F(x_{[0,\infty)})$.

7. Theorem. *Let the assumptions of Theorem 4 be satisfied. Further, let $s\in[0,T]$, and let $\zeta=\zeta(\omega)$, $\xi=\xi(\omega)$ be $\mathscr{F}_s$-measurable variables with values in Z and E_d, respectively. Finally, let $\sigma_{s+t}^z(x)$ and $b_{s+t}^z(x)$ be independent of ω for all $t\geq 0$.*

Suppose the process y_t satisfies the equation

$$y_t=\xi+\int_s^t\sigma_r^\zeta(y_r)\,d\mathbf{w}_r+\int_s^t b_r^\zeta(y_r)\,dr$$

for $t\in[s,T]$.

We define the process $x_t^{z,s,x}$ for $t\in[0,T-s]$ as a solution of the equation

$$x_t=x+\int_0^t\sigma_{s+r}^z(x_r)\,d\mathbf{w}_r+\int_0^t b_{s+r}^z(x_r)\,dr.$$

Then for any nonnegative measurable function $F(z,x_{[0,T-s]})$ given on $Z\times C([0,T-s],E_d)$,

$$\mathsf{M}\{F(\zeta,y_{[s,T]})\,|\,\mathscr{F}_s\}=\Phi(\zeta,\xi)\qquad\text{(a.s.)},$$

where

$$\Phi(z,x)=\mathsf{M}F(z,x_{[0,T-s]}^{z,s,x}).$$

PROOF. Let $\tilde{\mathbf{w}}_t=\mathbf{w}_{t+s}-\mathbf{w}_s$, $\tilde{\mathscr{F}}_t=\mathscr{F}_{t+s}$, $\tilde{y}_t=y_{t+s}$, $\tilde{\sigma}_t^z(x)=\sigma_{t+s}^z(x)$, $\tilde{b}_t^z(x)=b_{t+s}^z(x)$. It is seen that

$$\tilde{y}_t=\xi+\int_0^t\tilde{\sigma}_r^\zeta(\tilde{y}_r)\,d\tilde{\mathbf{w}}_r+\int_0^t\tilde{b}_r^\zeta(\tilde{y}_r)\,dr,$$

in this case ξ, ζ are $\tilde{\mathscr{F}}_0$-measurable, and $\tilde{\mathbf{w}}_t$ is a Wiener process with respect to $\tilde{\mathscr{F}}_t$. By Theorem 4

$$\mathsf{M}\{F(\zeta,y_{[s,T]})\,|\,\mathscr{F}_s\}=\mathsf{M}\{F(\zeta,\tilde{y}_{[0,T-s]})\,|\,\tilde{\mathscr{F}}_0\}=\tilde{\Phi}(\zeta,\xi)\qquad\text{(a.s.)},$$

where $\tilde{\Phi}(z,x) = \mathsf{M}F(z,\tilde{x}^{z,x}_{[0,T-s]})$ and $\tilde{x}^{z,x}_t$ is a solution of the equation

$$x_t = x + \int_0^t \tilde{\sigma}^z_r(x_r)\, d\tilde{\mathbf{w}}_r + \int_0^t \tilde{b}^z_r(x_r)\, dr.$$

It remains to note that, by Corollary 3, the processes $x^{z,s,x}_t$, $\tilde{x}^{z,x}_t$ have identical finite-dimensional distributions. Therefore $\tilde{\Phi}(z,x) = \Phi(z,x)$, thus proving the theorem. □

The technique involving a time shift can be applied in the case where s is a Markov time. The following fact, which we suggest the reader should prove using the above technique, leads to the so-called "strong Markovian" property of solutions of stochastic equations.

8. Exercise

Let $\sigma_t(x) \equiv \sigma(x)$, $b_t(x) \equiv b(x)$ be independent of t and ω, let τ be a Markov time with respect to $\{\mathscr{F}_t\}$, and let x^x_t be a solution (it is given for each t) of the equation

$$dx_t = \sigma(x_t)\, d\mathbf{w}_t + b(x_t)\, dt, \qquad x_0 = x.$$

Prove that in this situation for any $x \in E_d$ and a nonnegative measurable function $F = F(x_{[0,\infty)})$ given on $C([0,\infty),E_d)$,

$$\mathsf{M}_x\{\theta_\tau F \mid \mathscr{F}_\tau\} = \mathsf{M}_{x^x_\tau} F \qquad (\{\tau < \infty\}\text{-a.s.}),$$

where x indicates that in computing the conditional expectation one needs to take $x^x_{[0,\infty)}$ for the argument F, and x^x_τ indicates that first $\mathsf{M}_y F \equiv \mathsf{M}F(x^y_{[0,\infty)})$ is to be found and, second, y is to be replaced by x^x_τ.

9. Remark. The assertions of Theorems 4 and 7 hold not only for nonnegative functions F. This property of F was necessary to make the expressions we dealt with meaningful. For example, Theorem 7 holds for any measurable function F for which $\mathsf{M}|F(\zeta,y_{[s,T]})| < \infty$. In fact, by Theorem 7

$$\mathsf{M}\{F_\pm(\zeta,y_{[s,T]}) \mid \mathscr{F}_s\} = \Phi_{(\pm)}(\zeta,\xi) \qquad \text{(a.s.)}, \tag{15}$$

where $\Phi_{(\pm)}(z,x) = \mathsf{M}F_\pm(z,x^{z,s,x}_{[0,T-s]})$. In this case the left side of (15) is finite with probability 1 for both the sign $+$ and the sign $-$. In particular, the functions $\Phi_{(+)}(z,x)$, $\Phi_{(-)}(z,x)$ are finite for those (z,x) which are values of $(\zeta(\omega),\xi(\omega))$ on some subset Ω which has complete probability. Having subtracted from (15) with the $+$ sign, the same with the $-$ sign, we find

$$\mathsf{M}\{F(\zeta,y_{[s,T]}) \mid \mathscr{F}_s\} = \Phi(\zeta,\xi) \qquad \text{(a.s.)}, \tag{16}$$

where $\Phi(z,x) = \mathsf{M}F(z,x^{z,s,x}_{[0,T-s]})$; in this case the function $\Phi(z,x)$ exists at any rate for those (z,x) which are necessary for Eq. (16) to be satisfied.

Theorem 7 enables us to deduce the well-known Kolmogorov's equation for the case where $\sigma_t(x)$ and $b_t(x)$ do not depend on ω.

Denote by $x_t^{s,x}$ a solution of the equation

$$x_t = x + \int_0^t \sigma_{s+r}(x_r)\,d\mathbf{w}_r + \int_0^t b_{s+r}(x_r)\,dr, \tag{17}$$

$$(a_t^{ij}(x)) = \frac{1}{2}\sigma_t(x)\sigma_t^*(x),$$

$$L = L(t,x) = \sum_{i,j=1}^{d} a_t^{ij}(x)\frac{\partial^2}{\partial x^i \partial x^j} + \sum_{i=1}^{d} b_t^i(x)\frac{\partial}{\partial x^i} - c_t(x),$$

$$F(s,x_{[0,T-s]}) = \int_0^{T-s} f_{s+t}(x_t)\exp\left[-\int_0^t c_{s+r}(x_r)\,dr\right]dt$$
$$+ g(x_{T-s})\exp\left[-\int_0^{T-s} c_{s+r}(x_r)\,dr\right],$$

$$v(s,x) = \mathsf{M}F(s,x_{[0,T-s]}^{s,x}).$$

10. Theorem. *Let $c_t(x)$, $f_t(x)$, $g(x)$ be nonrandom real-valued functions, $c_t(x) \geq 0$. Let $\sigma_t(x)$, $b_t(x)$, $c_t(x)$, $f_t(x)$, $g(x)$ be twice differentiable in x, where neither $\sigma_t(x)$ nor $b_t(x)$ depends on ω. Furthermore, let the foregoing functions and their first and second derivatives with respect to x be continuous with respect to (t,x) in a strip $[0,T] \times E_d$. In addition, let the product of the functions $\sigma_t(x)$, $b_t(x)$, $c_t(x)$, $f_t(x)$, $g(x)$ and their first and second derivatives and the function $(1 + |x|)^{-m}$ (functions and their derivatives) be bounded in this strip. Then the function $v(t,x)$ has the following properties:*

1. $|v(t,x)| \leq N(1 + |x|)^m$ *for all $x \in E_d$, $t \in [0,T]$, where N does not depend on (t,x);*
2. *$v(t,x)$ is once differentiable with respect to t, is twice differentiable with respect to x, and, in addition, the derivatives are continuous in the strip $[0,T] \times E_d$;*
3. *for all $t \in [0,T]$, $x \in E_d$*

$$\frac{\partial}{\partial t}v(t,x) + Lv(t,x) + f_t(x) = 0, \qquad v(T,x) = g(x). \tag{18}$$

Moreover, any function which has properties (1)–(3) *coincides with v in the strip $[0,T] \times E_d$.*

Proof. By assumption, $\|\sigma_t(0)\|$, $|b_t(0)|$ are continuous. Therefore they are bounded on $[0,T]$ and

$$\|\sigma_t(x)\| + |b_t(x)| \leq \|\sigma_t(0)\| + |b_t(0)| + K|x| \leq N(1 + |x|),$$

where N does not depend on t, x. Furthermore, $F(s,x_{[0,T-s]}^{s,x})$ is a random variable since $F(s,x_{[0,T-s]})$ is a measurable (even continuous) function on

$C([0, T - s], E_d)$. From this and the assumptions $|f_t(x)| \le N(1 + |x|)^m$, $|g(x)| \le N(1 + |x|)^m$ and $c_t(x) \ge 0$ we deduce the first property of the function v if we use estimates of moments of solutions of a stochastic equation (see Corollary 5.12).

Equation (17) makes sense, in general, only for $t \in [0, T - s]$. It will be convenient to assume further that the process $x_t^{s,x}$ is defined for $t \in [0,T]$ for all $s \in (-\infty,\infty)$, $x \in E_d$. As before, we define the process $x_t^{s,x}$ as a solution of Eq. (17), in which, having redefined the functions $\sigma_t(x)$, $b_t(x)$ if necessary, we extend these functions from the interval $[0,T]$ to $(-\infty,\infty)$ defining $\sigma_t(x) = \sigma_T(x)$, $b_t(x) = b_T(x)$ for $t \ge T$ and $\sigma_t(x) = \sigma_0(x)$, $b_t(x) = b_0(x)$ for $t \le 0$. By Theorem 8.6, the process $x_t^{s,x}$ is twice $\mathscr{L}B$-differentiable with respect to x. By virtue of the results obtained in Section 7 (see Lemmas 7.11 and 7.12), the above proves that the random variable $F(s,x_{[0,T-s]}^{s,x})$ is twice $\mathscr{L}$-differentiable with respect to x for each $s \in [0,T]$, and also that the function $v(s,x)$ has all second derivatives with respect to x for each $s \in [0,T]$.

In order to prove that the function $v(s,x)$ is continuous with respect to (s,x), we need only assume in (17) that $p = (s,x)$, $x = \xi_t(p)$, $\sigma_{s+t}(y) = \sigma_t(p,y)$, $b_{s+t}(y) = b_t(p,y)$, write $c_{s+t}(y) = c_t(p,y)$, $f_{s+t}(y) = f_t(p,y)$ in the expression for F, and, in addition, make use of Corollary 8.2 as well as the results from Section 7. Using similar notation, taking the first and second $\mathscr{L}B$-derivatives of $x_t^{s,x}$ with respect to x (see Remark 8.5), and the $\mathscr{L}$-derivatives of $F(x,s_{[0,T-s]}^{s,x})$ and applying Corollary 8.2 as well as the results from Section 7, we prove that the first and second derivatives of $v(s,x)$ with respect to x are continuous with respect to (s,x).

This implies continuity of $Lv(s,x) + f_s(x)$ with respect to (s,x). Hence, if the first relation in (18) has been proved, we have continuity of $(\partial/\partial t)v(t,x)$. It should be mentioned that the second relation in (18) is obvious. Therefore, it remains only to prove that the derivative $(\partial/\partial t)v(t,x)$ exists and the first equality in (18) is satisfied. Furthermore, it suffices to prove this fact not for $(\partial/\partial t)v(t,x)$ but only for the right derivative of the function $v(t,x)$ with respect to t for $t \in [0,T)$. Indeed, as is well known in analysis, if $f(t)$, $g(t)$ are continuous on $[0,T]$ and if the right derivative $f(t)$ is equal to $g(t)$ on $[0,T)$, then $f'(t) = g(t)$ on $[0,T]$. We fix x and take $t_2 > t_1$, t_1, $t_2 \in [0,T]$. Further, let $s = t_2 - t_1$. By Theorem 7 (see Remark 9),

$$\mathsf{M}\{F(t_2,x_{[s,T-t_1]}^{t_1,x})|\mathscr{F}_s\} = \Phi(x_s^{t_1,x}) \qquad \text{(a.s.)}, \tag{19}$$

where $\Phi(y) = \mathsf{M}F(t_2,x_{[0,T-t_2]}^{t_2,y}) = v(t_2,y)$. Furthermore, simple computations show that

$$F(t_1,x_{[0,T-t_1]}) = \int_0^s f_{t_1+t}(x_t)\exp\left[-\int_0^t c_{t_1+r}(x_r)\,dr\right]dt$$

$$+ F(t_2,x_{[s,T-t_1]})\exp\left[-\int_0^s c_{t_1+r}(x_r)\,dr\right].$$

From this and (19) we find

$$v(t_1,x) = \mathsf{M} \int_0^s f_{t_1+t}(x_t^{t_1,x})\Psi_t^{t_1,x}\,dt + \mathsf{M}v(t_2,x_s^{t_1,x})\Psi_s^{t_1,x}, \tag{20}$$

where $\Psi_t^{t_1,x} \equiv \exp[-\int_0^t c_{t_1+r}(x_r^{t_1,x})\,dr]$.

Next, let $w(y)$ be a smooth function with compact support equal to 1 for $|y - x| \leq 1$. Also, let $v_1(t_2,y) = v(t_2,y)w(y)$, $v_2(t_2,y) = v(t_2,y) - v_1(t_2,y)$. We represent the second term in (20) as the sum of two expressions starting from the equality $v = v_1 + v_2$. Using Ito's formula, we transform the expression which contains v_1. Note that derivatives of $v_1(t_2,y)$ are continuous and have compact support, and therefore bounded. We have

$$v(t_1,x) = v_1(t_2,x) + \mathsf{M}v_2(t_2,x_s^{t_1,x})\Psi_s^{t_1,x} + \mathsf{M}h_s^{t_1,x}, \tag{21}$$

where

$$h_s^{t_1,x} = \int_0^s \{f_{t_1+t}(x_t^{t_1,x}) + L(t_1 + t, x_t^{t_1,x})v_1(t_2,x_t^{t_1,x})\}\Psi_t^{t_1,x}\,dt.$$

It is seen that $v = v_1$ at a point x. We replace the expression $v_1(t_2,x)$ in (21) by the expression $v(t_2,x)$ and carry the latter into the left-hand side of (21). Further, we divide both sides of the equality by $s = t_2 - t_1$ and, in addition, we let $t_2 \downarrow t_1$. By the mean-value theorem, due to continuity of the expressions considered

$$\frac{1}{s} h_s^{t_1,x} \to f_{t_1}(x) + L(t_1,x)v_1(t_1,x) = f_{t_1}(x) + L(t_1,x)v(t_1,x).$$

Moreover, $|(1/s)h_s^{t_1,x}|$ does not exceed the summable quantity

$$N\left(1 + \sup_{t\in[0,T]} |x_t^{t_1,x}|\right)^q \tag{22}$$

for some suitable values of the constant N, q. Finally, $v_2(t_2,y) = 0$ for $|y - x| \leq 1$, and, by property 1, $|v_2(t_2,y)| \leq N(1 + |y|)^m$. Hence $|v_2(t_2,y)| \leq N|y - x|^{m+4}$ and by Corollary 5.12,

$$\left|\frac{1}{s}\mathsf{M}v_2(t_2,x_s^{t_1,x})\Psi_s^{t_1,x}\right| \leq \frac{N}{s}\mathsf{M}\sup_{t\leq s}|x_t^{t_1,x} - x|^{m+4} \leq Ns^{(m/2)+1} \to 0.$$

The arguments carried out above enable us to derive from (21) that the right derivative of the function $v(t,x)$ exists at a point $t = t_1$, and also prove that the derivative equals $[-f_{t_1}(x) - Lv(t_1,x)]$ for all $t_1 \in [0,T)$. As was explained above, this suffices to complete the demonstration of properties 1–3 for the function v.

We prove the last assertion of the theorem concerning uniqueness of solution of (18). Let $u(t,x)$ be a function having properties 1–3. In accord with Ito's formula for any $R > 0$

$$u(s,x) = \mathsf{M}\left\{u(s+\tau_R, x_{\tau_R}^{s,x})\Psi_{\tau_R}^{s,x} - \int_0^{\tau_R} \Psi_t^{s,x}\left[\frac{\partial}{\partial t} u(s+t, x_t^{s,x}) \right.\right.$$
$$\left.\left. + L(s+t, x_t^{s,x})u(s+t, x_t^{s,x})\right] dt\right\}$$
$$= \mathsf{M}\left\{u(s+\tau_R, x_{\tau_R}^{s,x})\Psi_{\tau_R}^{s,x} + \int_0^{\tau_R} f_{s+t}(x_t^{s,x})\Psi_t^{s,x}\, dt\right\}, \tag{23}$$

where τ_R equals the minimum of $T - s$ and the first exit time of $x_t^{s,x}$ from S_R. It is seen that $\tau_R \to T - s$ for $R \to \infty$. Moreover, the expression in the curly brackets under the sign of the last mathematical expectation in (23) is continuous with respect to τ_R and, in addition, it does not exceed a summable quantity of the type (22). Therefore, assuming in (23) that $R \to \infty$, using the Lebesgue theorem, we can interchange the sign of the limit and the sign of the expectation. Having done this and, further, having noted that $u(T,x) = g(x)$, we immediately obtain $u(s,x) = v(s,x)$, thus proving the theorem. □

11. Remark. The last assertion of the theorem shows that $v(s,x)$ depends neither on an initial probability space nor on a Wiener process. The function $v(s,x)$ can be defined uniquely by the functions $a_t(x)$, $b_t(x)$, $c_t(x)$, $f_t(x)$, $g(x)$, i.e., by the elements which belong to (18). The function $v(s,x)$ does not change if we replace the probability space, or take another Wiener process, perhaps, even a d_2-dimensional process with $d_2 \neq d_1$, or, finally, take another matrix $\sigma_t(x)$ of dimension $d \times d_2$, provided only that the matrix $\sigma_t(x)\sigma_t^*(x)$ does not change.

10. Ito's Formula with Generalized Derivatives

Ito's formula is an essential tool of stochastic integral theory. The classical formulation of the theorem on Ito's formula involves the requirement that the function to which this formula can be applied be differentiable a sufficient number of times. However, in optimal control theory there arises a necessity to apply Ito's formula to nonsmooth functions (see Section 1.5).

In this section, we prove that in some cases Ito's formula remains valid for functions whose generalized derivatives are ordinary functions. Moreover, we prove some relationships between functions having generalized derivatives and mathematical expectations. These relationships will be useful for our further discussion.

We fix two bounded regions $D \subset E_d$, $Q \subset E_{d+1}$ in spaces E_d and E_{d+1}, respectively. Let d_1 be an integer, $d_1 \geq d$, let $(\mathbf{w}_t, \mathscr{F}_t)$ be a d_1-dimensional Wiener process, let $\sigma_t = \sigma_t(\omega)$ be a matrix of dimension $d \times d_1$, let $b_t = b_t(\omega)$ be a d-dimensional vector, and, finally, let $c_t = c_t(\omega)$ be real-valued. Furthermore, let

$$a_t = \frac{1}{2}\sigma_t\sigma_t^*, \qquad \varphi_t = \int_0^t c_t\,dr,$$

$$L_t = \sum_{i,j=1}^d a_t^{ij}\frac{\partial^2}{\partial x^i\,\partial x^j} + \sum_{i=1}^d b_t^i\frac{\partial}{\partial x^i} - c_t.$$

Assume that σ_t, b_t, c_t are progressively measurable with respect to $\{\mathscr{F}_t\}$ and, in addition, for all $t \geq 0$

$$\mathsf{M}\int_0^t \|\sigma_r\|^2\,dr < \infty, \qquad \mathsf{M}\int_0^t |b_r|\,dr < \infty, \qquad c_t \geq 0.$$

Under the assumption made above, for each $x_0 \in E_d$ the process

$$x_t = x_0 + \int_0^t \sigma_r\,d\mathbf{w}_r + \int_0^t b_r\,dr$$

is well-defined.

1. Theorem. *Let s, x_0 be fixed, $x_0 \in E_d$, $s \in (-\infty,\infty)$. Also, let τ_Q be the first exit time of the process $(s + t, x_t)$ from a region Q, let τ be some Markov time (with respect to $\{\mathscr{F}_t\}$) such that $\tau \leq \tau_Q$, let τ_D be the first exit time of the process x_t from a region D, and, finally, let τ' be a Markov time not exceeding τ_D. Suppose that there exist constants K, $\delta > 0$ such that $\|\sigma_t(\omega)\| + |b_t(\omega)| + c_t(\omega) \leq K$, $(a_t\lambda,\lambda) \geq \delta|\lambda|^2$ for all $\lambda \in E_d$ and (ω,t), which satisfy the inequality $t < \tau \vee \tau'$.*

Then for any $u \in \bar{W}^2(D)$, $v \in \bar{W}^{1,2}(Q)$, $t \geq 0$

$$e^{-\varphi_{\tau'}}u(x_{\tau'}) - e^{-\varphi_t}u(x_t) = \int_t^{\tau'} e^{-\varphi_r}L_r u(x_r)\,dr + \int_t^{\tau'} e^{-\varphi_r}\operatorname{grad}_x u(x_r)\sigma_r\,d\mathbf{w}_r,$$

$$e^{-\varphi_\tau}v(s+\tau, x_\tau) - e^{-\varphi_t}v(s+t, x_t) = \int_t^{\tau} e^{-\varphi_r}\left(\frac{\partial}{\partial r} + L_r\right)v(s+r, x_r)\,dr + \int_t^{\tau} e^{-\varphi_r}\operatorname{grad}_x v(s+r, x_r)\sigma_r\,d\mathbf{w}_r \quad (1)$$

almost surely on the sets $\{\tau' \geq t\}$, $\{\tau \geq t\}$, respectively. Furthermore, for any $u \in W^2(D)$, $v \in W^{1,2}(Q)$

$$u(x_0) = -\mathsf{M}\int_0^{\tau'} e^{-\varphi_r}L_r u(x_r)\,dr + \mathsf{M}e^{-\varphi_{\tau'}}u(x_{\tau'}),$$

$$v(s,x_0) = -\mathsf{M}\int_0^{\tau} e^{-\varphi_r}\left(\frac{\partial}{\partial r} + L_r\right)v(s+r, x_r)\,dr + \mathsf{M}e^{-\varphi_{\tau}}v(s+\tau, x_\tau).$$

PROOF. We prove both the assertions of Theorem 1 in the same way via approximation of u, v by smooth functions. Hence we prove the first assertion only.

Let a sequence $v^n \in C^{1,2}(\bar{Q})$ be such that

$$\|v - v^n\|_{B(Q)} \to 0, \qquad \|v - v^n\|_{W^{1,2}(Q)} \to 0,$$
$$\| |\text{grad}_x (v - v^n)|^2 \|_{d+1,Q} \to 0.$$

Further, let

$$y_t = x_0 + \int_0^t \chi_{r<\tau}\sigma_r \, d\mathbf{w}_r + \int_0^t \chi_{r<\tau} b_r \, dr.$$

We note that $y_t = x_t$ for $t \leq \tau < \infty$, which can easily be seen for $t < \tau$, and which follows from the continuity property of y_t and x_t for $t = \tau < \infty$. We prove that the right side of Eq. (1) makes sense. Obviously, for $r < \tau$

$$\left|\left(\frac{\partial}{\partial r} + L_r\right) v(s + r, x_r)\right| \leq N \left[\sum_{i,j=1}^{d} |v_{x^i x^j}(s + r, x_r)| + \left|\frac{\partial}{\partial r} v(s + r, x_r)\right| + \sum_{i=1}^{d} |v_{x^i}(s + r, x_r)| + |v(s + r, x_r)|\right],$$

where N depends only on d, K. From this, using Theorem 2.4[11] we obtain

$$\mathsf{M} \int_0^\tau \left|\left(\frac{\partial}{\partial r} + L_r\right) v(s + r, x_r)\right| dr = \mathsf{M} \int_0^\tau \chi_Q(s + r, y_r) \left|\left(\frac{\partial}{\partial r} + L_r\right) v(s + r, y_r)\right| dr \leq N\|v\|_{W^{1,2}(Q)}. \tag{2}$$

Similarly,

$$\mathsf{M}\left|\int_0^\tau e^{-\varphi_r} \text{grad}_x v(s + r, x_r)\sigma_r \, d\mathbf{w}_r\right|^2 \leq N\mathsf{M} \int_0^\tau \chi_Q(s + r, y_r)|\text{grad}_x v(s + r, y_r)|^2 \, dr \leq N\| |\text{grad}_x v|^2 \|_{d+1,Q}. \tag{3}$$

Further, we apply Ito's formula to the expression $v^n(t, y_t)e^{-\varphi_t}$. Then, we have on the set $\{t \leq \tau\}$ almost surely

$$e^{-\varphi_\tau} v^n(s + \tau, x_\tau) - e^{-\varphi_t} v^n(s + t, x_t) = \int_t^\tau e^{-\varphi_r}\left(\frac{\partial}{\partial r} + L_r\right) v^n(s + r, x_r) \, dr + \int_t^\tau e^{-\varphi_r} \text{grad}_x v^n(s + r, x_r)\sigma_r \, d\mathbf{w}_r. \tag{4}$$

We pass to the limit in equality (4) as $n \to \infty$. Using estimates similar to estimates (2) and (3), we easily prove that the right side of (4) tends to the right side of (1).

The first assertion of Theorem 1 can be proved for the function u by an almost word-for-word repetition of the proof given. The slight difference is

[11] In Theorem 2.4, we need take for D any region such that $(-\infty,\infty) \times D \supset Q$.

that if for v^n the existence of the terms in (4) follows from the obvious boundedness of $\tau(\omega)$, then a similar formula for proving the first assertion of the theorem for u is valid since $\tau'(\omega) < \infty$ (a.s.) and even $\mathrm{M}\tau' < \infty$ (in Theorem 2.4, assume that $s = 0$, $g \equiv 1$). The theorem is proved. □

Henceforth, when we mention this theorem we shall call the assertions of the theorem Ito's formulas.

The assumption that the process x_t is nondegenerate is the most restrictive assumption of Theorem 1. However, we note that the formulation of the well-known Ito formula imposes no requirement for a process to be nondegenerate when only differentiable functions are being considered. In the next theorem the assumption about nondegeneracy will be dropped, and in Ito's formula instead of an equality an inequality will be proved.

Consider the case where σ_t, b_t, and c_t depend on the parameter $x \in E_d$. We fix $s \in E_1$. Furthermore, for $t \geq s$, $x \in E_d$ let there be given: $\sigma_t(x)$, a random matrix of dimension $d \times d_1$; $b_t(x)$, a random d-dimensional vector; $c_t(x)$ and $f_t(x)$, random variables. Assume that $\sigma_{s+t}(x)$, $b_{s+t}(x)$, $c_{s+t}(x)$, $f_{s+t}(x)$ are progressively measurable with respect to $\{\mathscr{F}_t\}$ for each x, and that $c_t(x)$, $f_t(x)$ are continuous with respect to x and bounded for $(\omega,t,x) \in \Omega \times Q$, where Q, as before, is a bounded region in E_{d+1}. Also, for all $t \geq s$, x and $y \in E_d$ let

$$\|\sigma_t(x) - \sigma_t(y)\| + |b_t(x) - b_t(y)| \leq K|x - y|,$$
$$\|\sigma_t(x)\| + |b_t(x)| \leq K(1 + |x|),$$

where K is a constant.

Under the above assumptions, for each $x \in E$ the solution $x_t^{s,x}$ of the equation

$$x_t = x + \int_0^t \sigma_{s+r}(x_r)\, d\mathbf{w}_r + \int_0^t b_{s+r}(x_r)\, dr$$

exists and is unique (see Theorem 5.7).

We denote by $\tau_Q^{s,x}$ the first exit time of $(s + t, x_t^{s,x})$ from the region Q;

$$a_t(x) = \frac{1}{2}\sigma_t(x)\sigma_t^*(x);$$

$$L_t(x) = \sum_{i,j=1}^{d} a_t^{ij}(x)\frac{\partial^2}{\partial x^i\, \partial x^j} + \sum_{i=1}^{d} b_t^i(x)\frac{\partial}{\partial x^i} - c_t(x);$$

$$\varphi_t^{s,x} = \int_0^t c_{s+r}(x_r^{s,x})\, dr.$$

2. Theorem. *Let* $(s,x) \in Q$ *and, in addition, let a function* $v \in C(\bar{Q})$ *belong to* $W^{1,2}(Q')$ *for each region* Q'*, which together with its closure lies in* Q*. Assume that the derivatives of* v *can be chosen so that for some set* $\Gamma \subset Q$*, for which* meas $(Q \backslash \Gamma) = 0$, *for all* ω *and* $(t,y) \in \Gamma$ *the inequality*

$$-\left[\frac{\partial}{\partial t} + L_t(y)\right]v(t,y) \geq f_t(y) \tag{5}$$

can be satisfied. Then for any Markov time τ (with respect to $\{\mathscr{F}_t\}$) not exceeding $\tau_Q^{s,x}$,

$$v(s,x) \geq \mathrm{M} e^{-\varphi_\tau} v(s+\tau, x_\tau) + \mathrm{M} \int_0^\tau e^{-\varphi_t} f_{s+t}(x_t)\,dt, \tag{6}$$

where $\varphi_t = \varphi_t^{s,x}$, $x_t = x_t^{s,x}$.

Proof. In proving Theorem 2, we drop the superscripts s, x. First, we note that in proving this theorem we can assume that $\tau \leq \tau_{Q'}$, where $Q' \subset \bar{Q}' \subset Q$. Indeed, for all such Markov times let our theorem have been proved. We take an arbitrary time $\tau \leq \tau_Q$. It is seen that $\tau_{Q'} \uparrow \tau_Q$ and $\tau \wedge \tau_{Q'} \uparrow \tau$ when the regions Q', while expanding, converge to Q. Substituting in (6) the variable $\tau \wedge \tau_{Q'}$ for τ, taking the limit as $Q' \uparrow Q$, and, finally, noting that v is continuous in $\bar{Q}$, φ_t and x_t are continuous with respect to t, and, in addition, τ and $f_{s+t}(x_t)$ for $t \leq \tau$ are bounded, we have proved the assertion of the theorem in the general case.

Thus, let $\tau \leq \tau_{Q'}$. Further, we apply a rather well-known method of perturbation of an initial stochastic equation (see Exercise 1.1.1). We consider some d-dimensional Wiener process $\tilde{\mathbf{w}}_t$ independent of $\{\mathscr{F}_t\}$. Formally, this can be done by considering a direct product of two probability spaces: an initial space and a space on which a d-dimensional Wiener process is defined.

We denote by x_t^n a solution of the equation

$$x_t^n = x + \int_0^t \sigma_{s+r}(x_r^n)\,d\mathbf{w}_r + \varepsilon_n \tilde{\mathbf{w}}_t + \int_0^t b_{s+r}(x_r^n)\,dr,$$

where $\varepsilon_n \neq 0$, $\varepsilon_n \to 0$ as $n \to \infty$.

It is convenient to rewrite the last equation in a different form. Let $\sigma_t^n(x)$ be a matrix of dimension $d \times (d_1 + d)$, such that the first d_1 columns of the matrix $\sigma_t^n(x)$ form a matrix $\sigma_t(x)$, and also the columns numbered $d_1 + 1, \ldots, d_1 + d$ form a matrix $\varepsilon_n I$, where I is a unit matrix of dimension $d \times d$. Furthermore, we take a $(d_1 + d)$-dimensional Wiener process $\bar{\mathbf{w}}_t = (\mathbf{w}_t^1, \ldots, \mathbf{w}_t^{d_1}, \tilde{\mathbf{w}}_t^1, \ldots, \tilde{\mathbf{w}}_t^d)$. Then

$$x_t^n = x + \int_0^t \sigma_{s+r}^n(x_r^n)\,d\bar{\mathbf{w}}_r + \int_0^t b_{s+r}(x_r^n)\,dr. \tag{7}$$

By Theorem 8.1, $\sup_{r \leq t} |x_r^n - x_r| \to 0$ as $n \to \infty$ in probability for each t. Therefore, there exists a subsequence $\{n_i\}$ such that $\sup_{r \leq t} |x_r^{n_i} - x_r| \to 0$ (a.s.) as $i \to \infty$ and for each t. In order not to complicate the notation, we assume that $\{n_i\} = \{n\}$.

Let $\tau_{Q'}^n$ be the first exit time of $(s + t, x_t^n)$ from Q'. It is not hard to show that $\underline{\lim}_{n \to \infty} \tau_{Q'}^n \geq \tau_{Q'}$ (a.s.). Hence, if we assume that

$$\tau^i = \tau \wedge \inf_{n \geq i} \tau_{Q'}^n,$$

then $\tau^i \leq \tau_{Q'}$ and $\tau^i \to \tau$ as $i \to \infty$ (a.s.).

Further, we apply Theorem 1 to v, Q', x_t^n, τ^i for $n \geq i$. Note that $\tau^i \leq \tau_Q^n$ for $n \geq i$. Moreover, $v \in W^{1,2}(Q')$. Next, it is seen that

$$\chi_{t < \tau_{Q'}^n}(|\sigma_{s+t}^n(x_t^n)| + |b_{s+t}(x_t^n)| + |c_{s+t}(x_t^n)|) \leq N,$$

where N does not depend on t, ω, n. Finally,

$$a_t^n \equiv \frac{1}{2}\sigma_t^n(x_t^n)[\sigma_t^n(x_t^n)]^* = a_t(x_t^n) + \frac{\varepsilon_n^2}{2} I,$$

$$(a_t^n\lambda,\lambda) \geq \frac{\varepsilon_n^2}{2}|\lambda|^2.$$

All the assumptions of Theorem 1 have been satisfied. Therefore, computing for the process x_r^n (see (7)) the operator L_r appearing in Theorem 1, and in addition, assuming that

$$\varphi_t^n = \int_0^t c_{s+r}(x_r^n)\,dr,$$

$$g_t^n(x) = \left[\frac{\partial}{\partial t} + L_t(x)\right]v(t,x) + \frac{\varepsilon_n^2}{2}\Delta v(t,x),$$

for $n \geq i$, we have

$$v(s,x) = -\mathsf{M}\int_0^{\tau^i} e^{-\varphi_r^n} g_{s+r}^n(x_r^n)\,dr + \mathsf{M}e^{-\varphi_{\tau^i}^n} v(s+\tau^i, x_{\tau^i}^n). \tag{8}$$

By the hypothesis of the theorem,

$$-\chi_\Gamma(t,x)g_t^n(x) \geq \chi_\Gamma(t,x)f_t(x) - \chi_\Gamma(t,x)\frac{\varepsilon_n^2}{2}\Delta v(t,x).$$

Furthermore, by Theorem 2.4,

$$\mathsf{M}\int_0^{\tau^i} \chi_{Q\setminus\Gamma}(s+r, x_r^n)\,dr \leq N\|\chi_{Q\setminus\Gamma}\|_{d+1,H_\infty} = 0.$$

Therefore, in integrating over r in the first expression in the right side of (8), we can assume that $(s + r, x_r^n) \in \Gamma$. From (8) we find

$$v(s,x) \geq \mathsf{M}\int_0^{\tau^i} e^{-\varphi_r^n} f_{s+r}(x_r^n)\,dr + \mathsf{M}e^{-\varphi_{\tau^i}^n} v(s+\tau^i, x_{\tau^i}^n)$$
$$-\frac{\varepsilon_n^2}{2}\mathsf{M}\int_0^{\tau^i} e^{-\varphi_r^n}\Delta v(s+r, x_r^n)\,dr.$$

Because τ^i does not exceed the diameter T of the region Q', $\sup_{r\leq T}|x_r^n - x_r| \to 0$ as $n \to \infty$, $f_{s+r}(y)$ and $c_{s+r}(y)$ are continuous with respect to y, and $\tau^i \uparrow \tau$ as $i \to \infty$, we conclude that in the last expression for $v(s,x)$ the first two terms in the right side as $n \to \infty$, then as $i \to \infty$, yield the right side of Eq. (6).

Therefore, for proving the theorem it remains only to show that

$$\lim_{n\to\infty} \varepsilon_n^2\mathsf{M}\int_0^{\tau^n} |\Delta v(s+r,x_r^n)|\,dr = 0.$$

Making use of Theorem 2.2, we assume $s = 0$, $c_t \equiv 1$, $F(c,a) = c$, $b_t = b_{s+t}(x_t^n)$, $r_t \equiv 1$, $p = d$, $\sigma_t = \sigma_{s+t}^n(x_t^n)$. Note that, as was noted before, $|b_t| \leq$

$N \cdot 1 = Nc_t$ for $t < \tau^n_{Q'}$, where N does not depend on n, and, moreover,

$$(\det a_t)^{1/(d+1)} = (\det a^n_{s+t}(x^n_t))^{1/(d+1)} \geq \left(\frac{\varepsilon_n^2}{2}\right)^{d/(d+1)}.$$

Therefore

$$\begin{aligned}\varepsilon_n^2 \mathsf{M} \int_0^{\tau^n} |\Delta v(s+r, x^n_t)|\,dr &\leq \varepsilon_n^2 e^T \mathsf{M} \int_0^{\tau^n} e^{-r}|\Delta v(s+r, x^n_r)|\,dr \\ &\leq \varepsilon_n^2 e^T \mathsf{M} \int_0^{\tau^n_{Q'}} e^{-r}|\Delta v(s+r, x^n_r)|\,dr \\ &\leq 2^{d/(d+1)} \varepsilon_n^{2/(d+1)} e^T \mathsf{M} \int_0^{\tau^n_{Q'}} e^{-r}(\det a^n_{s+r}(x^n_r))^{1/(d+1)} \\ &\quad \times (\chi_{Q'}|\Delta v|)(s+r, x^n_r)\,dr \\ &\leq 2^{d/(d+1)} \varepsilon_n^{2/(d+1)} e^T N \|\chi_{Q'}|\Delta v|\|_{d+1, H_\infty},\end{aligned}$$

where N does not depend on n. The last expression tends to zero as $n \to \infty$, since $v \in W^{1,2}(Q')$. Therefore, the norm of that expression is finite. The theorem is proved. □

3. Remark. It is seen from the proof that if for all (t,ω) the function $f_t(x)$ is upper semicontinuous, $\underline{\lim}_{x_n \to x} f_t(x_n) \geq f_t(x)$, the assertion of the theorem still holds.

4. Corollary. *If $\sigma_t(x)$, $b_t(x)$, $c_t(x)$ do not depend on ω and in addition, $L_t(x)v(t,x) + \partial v(t,x)/\partial t$ is a bounded continuous function of $(t,x) \in Q$, we have in the notation of the theorem*

$$v(s,x) = \mathsf{M} e^{-\varphi_\tau} v(s+\tau, x_\tau) - \mathsf{M} \int_0^\tau e^{-\varphi_r} \left[L_{s+r}(x_r) v(s+r, x_r) + \frac{\partial v}{\partial \tau}(s+r, x_r) \right] dr.$$

5. Exercise to Theorem 1

(Compare [44, p. 39].) Let $d \geq 2$, $\alpha \in (0,1)$, $\mu = [(d-1)/(1-\alpha)] - 1$, $u(x) = |x|^\alpha$, $\sigma(x) = \sqrt{2a(x)}$, where $a^{ij}(x) = \delta^{ij} + \mu(x^i x^j/|x|^2)$. We take as D a sphere S_R, and also, we take as x_t some (possibly "weak") solution of the equation $dx_t = \sigma(x_t)\,dw_t$, $x_0 = 0$. Let $\sigma_t = \sigma(x_t)$, $b_t = 0$, $c_t = 0$.

Show that second derivatives of u are summable with respect to D to the power $p = \alpha d/(2-\alpha)$. (Note that $p \to d$ as $\alpha \to 1$.) Also, show that $Lu(x_t) = 0$ (a.s.) and that Ito's formula is not applicable to $u(x_t)$.

6. Remark. In the case where $Q = (0,T) \times S_R$, we have $\tau^{s,x} = 0$ for $s = 0$ in the notations introduced before Theorem 2. This suggests that it would be useful to have in mind that if $Q = (0,T) \times S_R$, one can take in Theorem 2 instead of $\tau^{s,x}$ (in Theorem 1 instead of τ_Q) the minimum between $T - s$ and the first exit time of the process $x^{s,x}_t$ (respectively, the first exit time of the process x_t) from S_R. For $s = 0$ this minimum is not in general equal to zero. Thus we can derive meaningful assertions from Theorems 1 and 2.

In order to prove the validity of the remark made above, it suffices to repeat word-for-word the proof of Theorems 1 and 2.

Notes

Section 1. The notations and definitions given in this section are of common usage. Definition 2 as well as the concept of an exterior norm are somewhat special.

Sections 2, 3, 4. The results obtained in these sections generalize the corresponding results obtained in [32, 34, 36, 40]. Estimates of stochastic integrals having a jumplike part can be found in Pragarauskas [62].

Sections 5, 7, 8, 9. These sections contain more or less well-known results related to the theory of Ito's stochastic integral equations; see Dynkin [11], Liptser and Shiryayev [51], and Gikhman and Skorokhod [24]. The introduction of the spaces $\mathscr{L}$, $\mathscr{L}B$ is our idea.

Section 6. The existence of a solution of a stochastic equation containing measurable coefficients not depending on time was first proved in [28] by the method due to Skorokhod [70]. In this section we use Skorokhod's method in the case when the coefficients may depend on time. For the problem of uniqueness of a solution of a stochastic equation as well as the problem of constructing the corresponding Markov process, see [24, 28, 38]; also see S. Anoulova and G. Pragarauskas: On weak Markov solutions of stochastic equations, *Litovsky Math. Sb.* 17(2) (1977), 5–26, also see the references listed in this paper.

Section 10. The results obtained in this section are related to those in [28, 34].