Chapter II

Stochastic Differential Equations

§1. General Problems of the Theory of Stochastic Differential Equations

In the present section we introduce the notion of a stochastic differential equation and prove some general theorems concerning the existence and uniqueness of solutions of these equations. For this purpose it is necessary to generalize the notion of a stochastic integral introduced above. Generally speaking, our approach to stochastic differential equations is based on the following considerations.

Assume that we are dealing with a motion of a system S in the phase space $\mathcal{R}^m$ and let $\xi(t)$ denote the location of this system in $\mathcal{R}^m$ at time t ($\xi(t)=(\xi^1(t),\ldots,\xi^m(t))$). Assume also that a displacement of system S located at time t at point x during the time interval $(t, t+\Delta t)$ can be represented in the form

$$\xi(t+\Delta t)-\xi(t)=A(x, t+\Delta t)-A(x, t)+\delta. \tag{1}$$

Here $A(x, t)$ is, in general, a random function; the difference $A(x, t+\Delta t)-A(x, t)$ characterizes the action of an "external field of forces" at point x on S during the time period $(t, t+\Delta t)$ and δ is a quantity which is of a higher order of smallness in a certain sense than the difference $A(x, t+\Delta t)-A(x, t)$. If $A(x, t)$ as a function of t is absolutely continuous, then relation (1) can be replaced by the ordinary differential equation

$$\frac{d\xi}{dt}=A'_t(\xi(t), t). \tag{2}$$

Equation (2) defines the motion of S in $\mathcal{R}^m$ for $t>t_0$ under the initial condition $\xi(t_0)=\xi_0$ while $A'_t(x, t)$ determines the "velocity field" in the phase space at time t.

It is obvious that equation (2) cannot describe motions such as Brownian, i.e., motions which do not possess a finite velocity in the phase space or motions which possess discontinuities in the phase space. To obtain an equation which will describe a motion of systems of this kind it is expedient to replace relation (1) with an equation of the integral type. For this purpose, we visualize that the time interval $[t_0, t]$ is subdivided into subintervals by the subdividing points

$t_1, t_2, \ldots, t_n = t$. It then follows from (1) that

$$\xi(t)-\xi(t_0)=\sum_{i=0}^{n-1} A(\xi(t_i), t_{i+1})-A(\xi(t_i), t_i)+\sum_{i=1}^{n-1} \delta_i.$$

Since δ_i are of a small order, it is natural to assume that $\sum_{i=0}^{n-1} \delta_i \to 0$ as $n \to \infty$. In this case the last equality formally becomes

$$\xi(t)-\xi(t_0)=\int_{t_0}^{t} A(\xi(s), ds), \tag{3}$$

and the expression

$$\int_0^t A(\xi(s), ds)$$

can be called a stochastic integral in the random field $A(x, t)$ along the random curve $\xi(s)$, $s \in [t_0, t]$; the integral should be interpreted as the limit, in a certain sense, to be defined more precisely, of sums of the form

$$\sum_{i=0}^{n-1} [A(\xi(t_i), t_{i+1})-A(\xi(t_i), t_i].$$

Relation (3) is called a stochastic differential equation and is written in the form

$$d\xi = A(\xi(t), dt), \qquad \xi(t_0)=\xi_0, \qquad t \geqslant t_0.$$

Under sufficiently general assumptions, for example, if $A(x, t)$ is a quasi-martingale for each $x \in \mathscr{R}^m$, one can assume that

$$A(x, t)=\alpha(x, t)+\beta(x, t), \tag{4}$$

where $\beta(x, t)$ as a function of t is a local martingale and process $\alpha(x, t)$ is representable as the difference of two monotonically nondecreasing natural processes. In this connection it makes sense to suppose that the right-hand side of equation (3) can be represented according to formula (4) and introduce further restrictions on functions $\alpha(x, t)$ and $\beta(x, t)$ in various ways. For example, we may assume that the function $\alpha(x, t)$ appearing in expression (4) is an absolute continuous function of t while $\beta(x, t)$—as a function of t—is a local square integrable martingale. (Some more general assumptions concerning $\beta(x, t)$ are considered below.)

In what follows, equation (3) will be written in the form

$$\xi(t)=\xi_0+\int_{t_0}^{t} \alpha(\xi(s), s)\, ds+\int_{t_0}^{t} \beta(\xi(s), ds) \tag{5}$$

or

$$d\xi=\alpha(\xi(t), t)\, dt+\beta(\xi(t), dt), \qquad \xi(t_0)=\xi_0.$$

In the case when $\beta(x, t) \equiv 0$, equation (6) is called *an ordinary differential equation* (with a random right-hand side).

Often fields $\beta(x, t) = \{\beta^1(x, t), \ldots, \beta^m(x, t)\}$ of the form

$$\beta^k(x, t) = \int_0^t \sum_{j=1}^r \gamma_j^k(x, s)\, d\mu^j(s), \qquad k = 1, \ldots, m, \tag{6}$$

are considered. Here $\mu^j(s)$ are local mutually orthogonal square integrable martingales, $j = 1, \ldots, r$, and $\gamma_j^k(x, s)$ are random functions satisfying conditions which assure the existence of corresponding integrals. In this case, the second integral in equation (5) may be defined as the vector-valued integral with components

$$\int_{t_0}^t \beta^k(\xi(s), ds) = \int_{t_0}^t \sum_{j=1}^r \gamma_j^k(\xi(s), s)\, d\mu^j(s), \qquad k = 1, \ldots, m,$$

and the theory of stochastic integrals described in Section 2 of Chapter I can be utilized. However, if we confine ourselves to functions $\beta(x, t)$ of type (6) a substantial amount of generality is lost. This can be seen from the fact that the joint characteristic of processes $\beta^k(x, t)$ and $\beta^k(y, t)$ defined by formula (6) is of the form

$$\langle \beta^k(x, \cdot), \beta^k(y, \cdot) \rangle_t = \int_0^t \sum_{i=1}^m \gamma_j^k(x, s) \gamma_j^k(y, s)\, d\langle \mu^j, \mu^j \rangle_s,$$

while in the general case it is given by a function $\Gamma^k(x, y, t)$ which for a fixed t is an arbitrary nonnegative-defined kernel of arguments x and y:

$$\sum_{i,j=1}^N \Gamma^k(x_i, y_j, t) z_i z_j \geq 0 \quad \text{for all } z_j \in \mathcal{R}^1, \quad j = 1, \ldots, n, \quad n = 1, 2, \ldots.$$

For example (for simplicity we consider here the one-dimensional case) let functions $\gamma_j(x, t) = c_j(x, t)$, $j = 1, \ldots, m$, be nonrandom and $\mu_j(t) = w_j(t)$ be independent Wiener processes. In this case the correlation function $R(x, y, t)$ of the field

$$\beta(x, t) = \int_0^t \sum_{j=1}^m c_j(x, s)\, dw_j(s)$$

equals

$$R(x, y, t) = \mathsf{E}\beta(x, t)\beta(y, t) = \int_0^t \sum_{j=1}^m c_j(x, s) c_j(y, s)\, ds.$$

On the other hand, if we set $\beta(x, t) = w(x, t)$, where $w(x, t)$ is an arbitrary Gaussian field with independent increments in t, its correlation function

$R_w(x, y, t) = \mathsf{E}w(x, t)w(y, t)$ is then an arbitrary nonnegative-defined kernel (for a fixed t).

Thus, the restrictions, when considering stochastic integrals along a process $\xi(t)$ imposed by fields of the type (6) lead to a substantial narrowing of the class of problems under consideration. Therefore it is expedient to introduce a direct definition and investigate properties of the stochastic integral

$$\int_0^T \beta(\xi(s), ds)$$

by interpreting it in the simplest cases as the limit in probability of the sums

$$\sigma = \sigma(\xi) = \sum_{k=1}^{n} \beta(\xi(s_{k-1}), s_k) - \beta(\xi(s_{k-1}), s_{k-1}).$$

The sums σ are called *integral sums.*

It is appropriate to observe that remarks about the insufficient generality of random fields as given by relation (6) are not fully justified in the case when the stochastic processes are represented by equation (5), with $\alpha(x, t) = a(x, t)$ being a nonrandom function and $\beta(x, t)$ being a function with independent increments in t. Indeed, the increment $\Delta\xi(t)$ of a solution of equation (5) at each time t depends on $\xi(t)$ and on the value of the field $\beta(x, t)$ at the point $x = \xi(t)$ and is independent of the nature of the relationship between $\beta(x, t)$ and $\beta(y, t)$ at the point $y \neq \xi(t)$ (provided the probabilistic characteristics of the field $\beta(x, t)$ as a function of x are sufficiently smooth). Therefore one may expect that solutions of equations (5) will be stochastically equivalent for any two fields $\beta(x, t) = \beta_1(x, t)$ and $\beta(x, t) = \beta_2(x, t)$ under the condition that the joint distributions of the sequence of vectors

$$\{\beta_i(x, t_1), \beta_i(x, t_2), \ldots, \beta_i(x, t_N)\} \qquad \forall x \in \mathcal{R}^m, \quad \forall N = 1, 2, \ldots,$$

coincide for $i = 1$ and $i = 2$ and that the fields $\beta_i(x, t)$ possess independent increments in t.

For example, let $w(x, t)$ be an arbitrary Gaussian field possessing independent increments in t, $B(x, t) = \mathsf{E}w^k(x, t)w^j(x, t) = \{B_{jk}(x, t)\}$ and the functions $B_{jk}(x, t)$ be differentiable with respect to t, $b_{jk}(x, t) = (d/dt)B_{kj}(x, t)$. Denote by $\sigma(x, t)$ a symmetric matrix such that $\sigma^2(x, t) = b(x, t)$ and introduce independent Wiener processes $w_j(t)$, $j = 1, \ldots, n$. Set

$$\beta_1(x, t) = \int_0^t \sigma(x, s)\, dw(s), \qquad w(t) = (w_1(t), \ldots, w_n(t)).$$

Then

$$\mathsf{E}\beta_1(x, t)\beta_1(x, t) = \int_0^t \sigma(x, s)\sigma(x, s)\, ds = \int_0^t b(x, s)\, ds,$$

and one can expect that the solutions of the differential equations

$$d\xi = a(\xi, t)\, dt + w(\xi, dt),$$

$$d\xi = a(\xi, t)\, dt + \sigma(\xi, t)\, dw(t)$$

will be stochastically equivalent, although the fields $w(x, t)$ and $\beta_1(x, t)$ in general are not.

Analogous observations can be made also in the case when $\beta(x, t)$ is an arbitrary field with independent increments in t with finite moments of the second order. Assume that the increment $\beta(x, t+\Delta t)-\beta(x, t)$ possesses the characteristic function

$$\begin{aligned}&\mathsf{E}\exp\{i(z, \beta(x, t+\Delta t)-\beta(x, t))\}\\&=\exp\left\{-\frac{1}{2}\int_t^{t+\Delta t}(b(x, s)z, z)\,ds\right.\\&\qquad\left.+\int_t^{t+\Delta t}ds\int_{\mathcal{R}^m}[e^{i(z, c(x,s,u))}-1-i(z, c(x, s, u))]\Pi(s, du)\right\}\end{aligned}$$

(if $\beta(x, t)$ possess finite moments of the second order and is absolutely continuous in t one can reduce an arbitrary characteristic function of a process with independent increments to this form). In this case, one can expect for sufficiently smooth functions $a(x, t)$, $b(x, t)$ and $c(x, t, u)$ that solutions of the stochastic equations

$$d\xi = a(\xi, t)\,dt+\beta(\xi, dt)$$

and

$$d\xi = a(\xi, t)\,dt+\sigma(\xi, t)\,dt+\int_{\mathcal{R}^m}c(\xi, t, u)\tilde{\nu}(dt, du)$$

will be stochastically equivalent. Here $\sigma(x, t)$ is a symmetric matrix, $\sigma^2(x, t)=b(x, t)$, $\tilde{\nu}(t, A)$ is a centered Poisson measure with Var $\tilde{\nu}(t, A)=\int_0^t \Pi(s, A)\,ds$.

It is also clear that if the increments in t of the field $\beta(x, t)$ are dependent then the remarks above concerning the possibility of replacing the field $\beta(x, t)$ in equation (5) by a simpler field without restricting the class of obtained solutions are no longer valid.

The preceding outline of a definition of a stochastic differential equation is expedient to extend in yet another direction. At present, "feedback" systems play an important part in a number of scientific–engineering problems. For such systems the "exterior field of forces" acting on the system at a given time depends not only on the instantaneous location of the system in the phase space but also on its phase trajectory in "the past":

$$\xi(t+\Delta t)-\xi(t)=\alpha(\xi|_{t_0}^t, t+\Delta t)-\alpha(\xi|_{t_0}^t, t)+\delta_t, \tag{7}$$

where $\alpha(\varphi|_{t_0}^t, s)$, $s\geq t>t_0$ is a family of random functionals with values in $\mathcal{R}^m$ defined on a certain space of functions $\varphi(u)$, $u\in[t_0, t]$, with values in $\mathcal{R}^m$.

The notation $\alpha(\varphi|_{t_0}^t, s)$ will be inconvenient for our further discussions mainly due to the absence of a fixed region in which the arguments of the functional $\alpha(\cdot, s)$ vary. In order to avoid this difficulty one can proceed as follows.

Introduce the space $\mathcal{D}_T^m$ ($\mathcal{D}^m[a, b]$) of functions $\varphi(s)$ defined on $(-\infty, T]$ (on $[a, b]$), continuous from the right with values in $\mathcal{R}^m$ possessing—at each point of the domain of definition—right-hand and left-hand limits (and in the case of the space $\mathscr{D}_T^m$ also possessing the limit as $s \to -\infty$). Let $\mathscr{D}^m = \mathscr{D}_0^m$. Denote by θ_t $(t \leqslant T)$ the mapping of $\mathscr{D}_t^m$ into $\mathscr{D}^m$ defined by the relation:

$$(\theta_t\varphi)(s) = \varphi(t+s), \qquad s \leqslant 0.$$

Next let $\alpha(\varphi, t) = \alpha(\varphi, t, \omega)$ be a random function defined on $\mathscr{D}^m \times [0, T] \times \Omega$. Relation (7) can be rewritten as follows,

$$\xi(t+\Delta t) - \xi(t) = \alpha(\theta_t\xi, t+\Delta t) - \alpha(\theta_t\xi, t) + \delta_t,$$

and equation (5) can be represented by equation

(8) $$\xi(t) = \xi_0 + \int_{t_0}^t \alpha(\theta_s\xi, s)\, ds + \int_{t_0}^t \beta(\theta_s\xi, ds), \qquad t > t_0.$$

Here it becomes necessary to define the process $\xi(t)$ over the whole "past," i.e., up to the time t_0. In this connection one should adjoin to equation (8) relation

(9) $$\xi(t) = \varphi(t), \qquad t \leqslant t_0,$$

which will be called from now on the *initial condition* for the stochastic differential equation (8).

The stochastic line integral. Let $\{\mathfrak{F}_t, t \in [0, T]\}$ be a current of σ-algebras on a fixed probability space $\{\Omega, \mathfrak{S}, \mathsf{P}\}$, $(\mathfrak{F}_t \subset \mathfrak{S})$, and $\beta(\varphi, t)$ be a random function adopted to $\{\mathfrak{F}_t\}$ with values in $\mathscr{R}^m$.

Two variants of theorems are considered below. One variant will refer to random processes with continuous sample functions (with probability 1), the other to processes with sample functions without discontinuities of the second kind (mod P). In this connection two sets of assumptions are introduced.

Let $\mathscr{C}_T^m$ $(\mathscr{C}^m, \mathscr{C}^m[a, b])$ be the subspace of the space $\mathscr{D}_T^m$ $(\mathscr{D}^m, \mathscr{D}^m[a, b])$ consisting of continuous functions. The space $\mathscr{C}^m$ is endowed with the uniform norm

$$\|\varphi\| = \sup_{s \leqslant 0} |\varphi(s)|.$$

The space $\mathscr{D}^m$ will be assumed to be a metric space with the metric $\rho_{\mathscr{D}^m}$ of the space of functions without discontinuities of the second kind (Volume I, Chapter VI, Section 5). In order to simplify the study of discontinuous processes a simpler metric will be utilized in $\mathscr{D}^m$. (Further assumptions on the equations under consideration will be formulated in terms of this metric.) This metric is generated by the seminorm $\|\varphi\|_*$, defined by relation

(10) $$\|\varphi\|_* = \{\int_{-\infty}^0 |\varphi(s)|^2 |K(ds)|\}^{1/2},$$

where $K(\cdot)$ is a finite measure defined on Borel sets on the half-line $(-\infty, 0]$ and $K(-\infty, 0] = K < \infty$.

If, for example, *stochastic differential equations with a lag*, i.e., equations of the form

$$d\xi(t)=\alpha(\xi(t-h_1),\ldots,\xi(t-h_r),t)\,dt$$
$$+\beta(\xi(t-h_1),\xi(t-h_2),\ldots,\xi(t-h_r),dt),$$

are discussed, then the functions depending on values of $\varphi(s)$ at a finite number of points, i.e., functions of the form $\beta(\varphi(0),\varphi(-h_1),\ldots,\varphi(-h_r),t)$ should be chosen as functions $\beta(\varphi,t)$. In this case, it is natural to identify functions $\varphi(s)$ taking on the same values at points $0,-h_1,\ldots,-h_r$ and metrize $\mathscr{D}^m$ by means of the metric

$$\|\varphi-\psi\|_*=[(\varphi(0)-\psi(0))^2+(\varphi(-h_1)-\psi(-h_1))^2+\cdots+(\varphi(-h_r)-\psi(-h_r))^2]^{1/2},$$

i.e., using the seminorm (10) which corresponds to measure K concentrated at points $0=h_0,-h_1,\ldots,-h_r$ and taking on values $K(\{-h_k\})=1$ at these points.

Returning to functions $\beta(\varphi,t)$ we shall first assume that they satisfy one of the following two sets of conditions:

$\beta 1)$ a) Function $\beta(\varphi,s)=\beta(\varphi,s,\omega)$ is defined on $\mathscr{D}^m\times[0,T]\times\Omega$ and for each $t\leq T$ its contraction on the segment $s\in[0,t]$ is $\mathfrak{B}_{\mathscr{D}^m}\times\mathfrak{T}_t\times\mathfrak{F}_t$-measurable;

b) $\beta(\varphi,t)$ is a square integrable $\mathfrak{F}_t$-martingale for a fixed φ, with sample functions belonging to $\mathscr{D}^m[0,T]$ with probability 1 and the characteristics of the components of this $\mathfrak{F}_t$-martingale are continuous with probability 1.

Here $\mathfrak{B}_{\mathscr{D}^m}$ is the minimal σ-algebra of the subsets of $\mathscr{D}^m$ containing cylindrical sets in $\mathscr{D}^m$ and $\mathfrak{T}_t$ is the σ-algebra of Borel sets on the interval $[0,t]$.

$\beta 2)$ Function $\beta(\varphi,s)$ satisfies conditions which are obtained from $\beta 1)$ if $\mathscr{D}^m$, $\mathfrak{B}_{\mathscr{D}^m}$, and $\mathscr{D}^m[0,T]$ are replaced by $\mathscr{C}^m$, $\mathfrak{B}_{\mathscr{C}^m}$, and $\mathscr{C}^m[0,T]$, respectively.

A random function $\beta(\varphi,t)$ satisfying conditions $\beta 1)(\beta 2))$ is called a *martingale field* in $\mathscr{D}^m$ (in $\mathscr{C}^m$) or simply a field.

If $\beta(\varphi,t)$ is a martingale field in $\mathscr{D}^m$ then a random function $\Lambda(\varphi,t)$ exists which is, for a fixed φ, a natural integrable monotonically nondecreasing process such that $\Lambda(\varphi,0)=0$ and for any $\Delta=(t,t+\Delta t]$

$$\mathsf{E}\{|\beta(\varphi,\Delta)|^2\,|\,\mathfrak{F}_t\}=\mathsf{E}\{\Lambda(\varphi,\Delta)\,|\,\mathfrak{F}_t\},$$

where $\Lambda(\varphi,\Delta)=\Lambda(\varphi,t+\Delta t)-\Lambda(\varphi,t)$.

We say that the field $\beta(\varphi,t)$ is *linearly bounded in a seminorm* or in *a norm* if

$$\Lambda(\varphi,\Delta)\leq(1+\|\varphi\|_*^2)\Lambda_0(\Delta) \tag{11}$$

or, correspondingly,

$$\Lambda(\varphi,\Delta)\leq(1+\|\varphi\|^2)\Lambda_0(\Delta).$$

Here $\Lambda_0(t)$ is a continuous integrable monotonically nondecreasing process adopted to the current of σ-algebras $\{\mathfrak{F}_t, t\in[0, T]\}$. If $\Lambda(\varphi, t)$ is a continuous function of t for each φ, then the condition of linear boundedness in the seminorm is equivalent to the requirement:

There exists a process $\Lambda_0(t)$ satisfying the preceding conditions such that for any $\Delta\subset[0, T]$

$$\mathsf{E}\{|\beta(\varphi, \Delta)|^2\,|\,\mathfrak{F}_t\}\leqslant(1+\|\varphi\|_*^2)\mathsf{E}\{\Lambda_0(\Delta)\,|\,\mathfrak{F}_t\}. \tag{12}$$

It is trivial to verify that (11) implies (12). The converse follows easily from Theorem 21 in Section 1. An analogous remark is valid also for the fields which are linearly bounded in the norm.

Similar remarks hold concerning the martingale $\beta(\varphi, t)-\beta(\psi, t)$. If for an arbitrary $N>0$ a monotonically nondecreasing, continuous and integrable process $\Lambda_N(t)$, $t\in[0, T]$, exists; adopted to the current of σ-algebras $\{\mathfrak{F}_t, t\in[0, T]\}$, which is independent of φ and ψ and such that

$$\mathsf{E}\{|\beta(\varphi, \Delta)-\beta(\psi, \Delta)|^2\,|\,\mathfrak{F}_t\}\leqslant\|\varphi-\psi\|_*^2\mathsf{E}\{\Lambda_N(\Delta)|\mathfrak{F}_t\} \tag{13}$$

for all $\varphi, \psi\in\mathscr{D}^m$ satisfying the conditions $\|\varphi\|_*\leqslant N$ and $\|\psi\|_*\leqslant N$, we then say that $\beta(\varphi, t)$ satisfies *a local Lipschitz condition* (relative to the seminorm). If there exists a process $\Lambda(t)$ such that one can set $\Lambda_N(t)=\Lambda(t)$ for all $N>0$, we then say that $\beta(\varphi, t)$ satisfies a *uniform Lipschitz condition* (relative to the seminorm). Analogous terminology is used also in the case when the seminorm $\|\varphi-\psi\|_*$ is replaced by the norm $\|\varphi-\psi\|$ in equality (13).

We now present the definition of a stochastic line integral

$$\int_0^T\beta(\theta_t\xi, dt).$$

(This definition will be somewhat generalized below.)

The following assumptions are made concerning the random processes $\xi(t)$, $t\in(-\infty, T]$:

ξ1) The process $\xi(t)$, $t\in[0, T]$, is adopted to the current of σ-algebras $\{\mathfrak{F}_t, t\in[0, T]\}$, the variables $\xi(s)$ are $\mathfrak{F}_0$-measurable for $s<0$, and the sample functions of the process $\xi(t)$ belong to $\mathscr{D}_T^m$ with probability 1; or

ξ2) The process $\xi(t)$, $t\leqslant T$, satisfies condition ξ1) and the sample functions of this process belong with probability 1 to $\mathscr{C}_T^m$.

Let δ be a subdivision of the interval $[0, T]$ with the subdividing points

$$t_0=0<t_1<t_2<\cdots<t_n=T,$$
$$|\delta|=\max_{1\leqslant k\leqslant n}\Delta t_k,\qquad \Delta t_k=t_k-t_{k-1}.$$

In what follows Δ or Δ_k denote the semiinterval $(t, t+\Delta t]$ or $(t_k, t_{k+1}]$.

Theorem 1. *Let $\xi(t)$, $t\in(-\infty, T]$, satisfy the condition $\xi 1)$, and the random function $\beta(\varphi, t)$ satisfy $\beta 1)$ and also the local Lipschitz condition* (13). *Then the limit*

$$\int_0^T \beta(\theta_t\xi, dt) \underset{\text{Def}}{=} \mathsf{P}\text{-}\lim_{|\delta|\to 0} \sum_{k=1}^n \beta(\theta_{t_{k-1}}\xi, t_k)-\beta(\theta_{t_{k-1}}\xi, t_{k-1}) \tag{14}$$

exists.

Definition. The limit on the right-hand side of relation (14) (provided it exists) is called *the stochastic line integral* or *the stochastic integral in the field $\beta(\varphi, t)$ along the curve $\xi(t)$.*

Proof of Theorem 1. Let

$$\tau_N=(\inf\{t: |\xi(t)|>N\})\vee 0$$

(we assume that $\inf\{\varnothing\}=T$), $\tau'_L=\inf\{t: \Lambda_N(t)>L\}$, and $\xi_N(t)=\xi(t)$ for $t<\tau_N$, $\xi_N(t)=0$ if $t\geqslant\tau_N$, $\tilde\beta(\varphi, t)=\tilde\beta_{LN}(\varphi, t)=\beta(\varphi, t\wedge\tau'_L)$. Using the theorem on the stopping of a martingale, it is easy to verify that for $\|\varphi\|\leqslant N$, $\|\psi\|\leqslant N$,

$$\mathsf{E}\{|\tilde\beta(\varphi, \Delta)-\tilde\beta(\psi, \Delta)|^2\,|\,\mathfrak{F}_t\}\leqslant\|\varphi-\psi\|_*^2\mathsf{E}\{\tilde\Lambda_N(\Delta)\,|\,\mathfrak{F}_t\}, \tag{15}$$

where

$$\tilde\Lambda_N(\Delta)=\Lambda_N[(t+\Delta t)\wedge\tau_L]-\Lambda_N(t\wedge\tau_L).$$

Consider two subdivisions δ_1 and δ_2 of the interval $[0, T]$, where δ_2 is a refinement of δ_1 ($\delta_2<\delta_1$). Denote the subdividing points of δ_1 by t_k ($k=0, 1, \ldots, n$) and that of δ_2 by t_{kj} ($t_k=t_{k0}<t_{k1}<\cdots<t_{ks_k}=t_{k+1}$).

Set

$$\Delta t_k=t_k-t_{k-1}, \qquad \Delta t_{kj}=t_{kj}-t_{kj-1},$$

$$\Delta_k=(t_{k-1}, t_k], \qquad \Delta_{kj}=(t_{kj-1}, t_{kj}],$$

$$\sigma_1(\xi)=\sum_{k=1}^n \beta(\theta_{t_{k-1}}\xi, \Delta_k)=\sum_{k=1}^n \beta(\theta_{t_{k-1}}\xi, t_k)-\beta(\theta_{t_{k-1}}\xi, t_{k-1}),$$

let $\sigma_2(\xi)$ be the integral sum which is analogous to $\sigma_1(\xi)$ but is constructed using subdivision δ_2, and let $\tilde\sigma_i(\xi)$ be integral sums constructed from the field $\tilde\beta(\varphi, t)$ using subdivisions δ_i ($i=1, 2$). Then

$$\mathsf{P}\{|\sigma_1(\xi)-\sigma_2(\xi)|>\varepsilon\}\leqslant\mathsf{P}\{\tau_N\vee\tau'_L<T\}+\mathsf{P}\{|\tilde\sigma_1(\xi_N)-\tilde\sigma_2(\xi_N)|>\varepsilon\}.$$

Since the sample functions of the processes $\xi(t)$ ($t\in(-\infty, T]$) and $\Lambda_N(t)$ ($t\in[0, T]$) are bounded, the probability $\mathsf{P}\{\tau_N\vee\tau'_L<T\}$ can be made as small as desired for N and $L=L(N)$ sufficiently large.

We now bound the second summand in the right-hand side of the last inequality. Observe that

$$\tilde{\sigma}_1(\xi_N)-\tilde{\sigma}_2(\xi_N)=\sum_{k=0}^{n-1}\sum_{r=1}^{s_k}\tilde{\beta}(\theta_{t_{k-1}}\xi_N,\Delta_{kr})-\tilde{\beta}(\theta_{t_{kr-1}}\xi_N,\Delta_{kr}).$$

Using (15) we obtain

$$\begin{aligned}\mathsf{E}|\tilde{\sigma}_1(\xi_N)-\tilde{\sigma}_2(\xi_N)|^2&=\sum_{k=0}^{n-1}\sum_{r=1}^{s_k}\mathsf{E}|\tilde{\beta}(\theta_{t_{k-1}}\xi_N,\Delta_{kr})-\tilde{\beta}(\theta_{t_{kr-1}}\xi_N,\Delta_{kr})|^2\\&\leqslant\mathsf{E}\sum_{k=0}^{n-1}\sum_{r=1}^{s_k}\|\theta_{t_{k-1}}\xi_N-\theta_{t_{kr-1}}\xi_N\|_*^2\tilde{\Lambda}_N(\Delta_{kr}).\end{aligned}\tag{16}$$

Here we utilized the fact that

$$\begin{aligned}&\mathsf{E}|\tilde{\beta}(\theta_{t_{k-1}}\xi_N,\Delta_{kr})-\tilde{\beta}(\theta_{t_{kr-1}}\xi_N,\Delta_{kr})|^2\\&\quad=\mathsf{E}\mathsf{E}\{\cdots|\mathfrak{F}_{t_{kr-1}}\}\\&\quad=\mathsf{E}[\mathsf{E}\{|\tilde{\beta}(\varphi,\Delta_{kr})-\tilde{\beta}(\psi,\Delta_{kr})|^2|\mathfrak{F}_{t_{kr-1}}\}]_{\varphi=\theta_{t_{k-1}}\xi_N,\psi=\theta_{t_{kr-1}}\xi_N}.\end{aligned}$$

Note that the sum appearing under the expectation sign in the right-hand side of inequality (16) is uniformly bounded. This sum is bounded above by $4N^2K\tilde{\Lambda}_N(T)\leqslant 4N^2KL$, where $K=K(-\infty,0]$. Furthermore,

$$\|\theta_{t_{k-1}}\xi_N-\theta_{t_{kr-1}}\xi_N\|_*^2=\int_{-\infty}^0|\xi_N(t_{k0}+s)-\xi_N(t_{kr-1}+s)|^2K(ds),$$

so that

$$\begin{aligned}\mathsf{E}|\tilde{\sigma}_1(\xi_N)-\tilde{\sigma}_2(\xi_N)|^2=\mathsf{E}\int_{-\infty}^0\Big(\sum_{k,r}&|\xi_N(t_{k-1}+s)\\&-\xi_N(t_{kr-1}+s)|^2\Lambda_N(\tilde{\Delta}_{kr})\Big)K(ds).\end{aligned}\tag{17}$$

Using a method common in integral calculus it is easy to verify that the sum appearing under the sign of the integral in equation (17) tends to 0 as $|\delta|\to 0$ for all $s\leqslant 0$.

Indeed let $\varepsilon_1>0$ be an arbitrary given number. The function $\xi_N(u)$ possesses only a finite number of jumps of the size at least $\sqrt{\varepsilon_1/L}$ on the interval $[s,s+T]$. Let this be the points $s_1,\ldots,s_m$. We enclose each one of these points by intervals i_r of length h_0, where $h_0m<\varepsilon_1$. Omitting the intervals i_r from the segment $[s,s+T]$ we obtain a closed set S. We now choose ε_2 such that $|\xi_N(s')-\xi_N(s'')|^2<2\varepsilon_1/L$ provided $|s'-s''|<\varepsilon_2$, s' and $s''\in S$. It is easy to verify that such an ε_2 exists. Indeed, assuming otherwise, we construct sequences of points s'_n, s''_n, $n=1,2,\ldots$, such that $|s'_n-s''_n|<1/n$, $\lim s'_n=\lim s''_n=s_0\in S$, $|\xi_N(s'_n)-\xi_N(s''_n)|>2\varepsilon/L$, which is possible only if $|\xi_N(s_0-)-\xi_N(s_0)|\geqslant 2\varepsilon_1/L$ since the one-sided limits of function $\xi_N(s)$ exist. But the last inequality contradicts the fact that $s_0\in S$. If $|\delta_1|<h_0/2$

then each one of the intervals $\Delta_k = [t_{k-1}, t_k]$ is either located in the interior of S or it contains one of the end-points of i_r as an interior point or is located in the interior of i_r. Denote by I_1, I_2 and I_3 the corresponding sets of intervals Δ_k. Let $|\delta_1| < ((h_0/2) \wedge \varepsilon_2)$. Then

$$\begin{aligned} z &= \sum_{k,r} |\xi_N(t_{k0}+s) - \xi_N(t_{kr-1}+s)|^2 \tilde{\Lambda}_N(\Delta_{k_r}) \\ &\leqslant \sum_{I_1} + \sum_{I_2} + \sum_{I_3} \\ &\leqslant \frac{2\varepsilon_1}{L} \sum_{I_1} \tilde{\Lambda}_N(\Delta_{kr}) + 2m \cdot 4N^2 \cdot \max_{\Delta_k \in I_2} \tilde{\Lambda}_N(\Delta_k) + 4N^2 \sum_{\Delta_k \in I_3} \tilde{\Lambda}_N(\Delta_k) \\ &\leqslant 2\varepsilon_1 + 8mN^2 \max_{\Delta_k \in I_2} \tilde{\Lambda}_N(\Delta_k) + 4N^2 \sum_{r=1}^{m} \tilde{\Lambda}_N(i_r). \end{aligned}$$

Since the function $\tilde{\Lambda}_N(t)$ is continuous, one can, by fixing an arbitrary ε_0, choose ε_1 and then choose h_0 and ε_2 such that $z \leqslant \varepsilon_0$ for all δ_1 satisfying $|\delta_1| < (h_0/2 \wedge \varepsilon_2)$ (for a given ω). Thus $z \to 0$ as $|\delta_1| \to 0$ with probability 1.

Noting that it is permissible in inequality (17) to approach the limit under the sign of the integral, we obtain

$$\mathsf{E}|\tilde{\sigma}_1(\xi_N) - \tilde{\sigma}_2(\xi_N)|^2 \to 0 \quad \text{as } |\delta_1| \to 0.$$

Thus, $\mathsf{P}\{|\tilde{\sigma}_1(\xi_N) - \tilde{\sigma}_2(\xi_N)| > \varepsilon\} \to 0$ as $|\delta_1| \to 0$. This easily implies that for arbitrary subdivisions δ_1 and δ_2 of the interval $[0, T]$ (i.e., when δ_2 is not necessarily a subdivision of δ_1) we have

$$\mathsf{P}\{|\sigma_1(\xi) - \sigma_2(\xi)| > \varepsilon\} \to 0 \quad \text{as } |\delta_1|, |\delta_2| \to 0.$$

The theorem is thus proved. □

The preceding theorem can be somewhat refined in the case when integration over continuous processes is considered.

Theorem 2. *Assume that the process $\xi(t)$, $t \in (-\infty, T]$, satisfies condition $\xi 2$) and the field $\beta(\varphi, t)$ satisfies condition $\beta 2$) and the local Lipschitz condition (with respect to a uniform norm), i.e., for $\|\varphi\| \leqslant N$, $\|\psi\| \leqslant N$, and $t \in [0, T]$*

$$\mathsf{E}\{|\beta(\varphi, \Delta) - \beta(\psi, \Delta)|^2 \mid \mathfrak{F}_t\} \leqslant \|\varphi - \psi\|^2 \mathsf{E}\{\Lambda_N(\Delta) \mid \mathfrak{F}_t\}, \tag{18}$$

where $\Lambda_N(t)$ is a continuous monotonically nondecreasing integrable process adopted to $\{\mathfrak{F}_t, t \in [0, T]\}$. Then the stochastic line integral (14) *exists.*

Proof. Using the same argument as in the proof of Theorem 1 we obtain

$$\mathsf{E}|\tilde{\sigma}_1(\xi_N) - \tilde{\sigma}_2(\xi_N)|^2 \leqslant \mathsf{E} \sum_{k,r} \|\theta_{t_{k-1}}\xi_N - \theta_{t_{kr-1}}\xi_N\|^2 \tilde{\Lambda}_N(\Delta_{kr});$$

moreover, the sum in the right-hand side of the last inequality is uniformly bounded in ω. It follows from the assumptions imposed on the structure of function $\xi(t)$ that

$$\|\theta_{t_{k-1}}\xi_N-\theta_{t_{kr-1}}\xi_N\|^2=\sup_{s\le 0}|\xi(t_k+s)-\xi(t_{kr-1}+s)|^2\to 0$$

with probability 1 as $|\delta_1|\to 0$ uniformly in k and r. Therefore

$$\mathsf{E}|\tilde\sigma_1(\xi_N)-\tilde\sigma_2(\xi_N)|^2\to 0$$

as $|\delta_1|\to 0$. This, as in the case of Theorem 1, implies the required assertion. □

Remark. If the conditions of Theorems 1 and 2 are satisfied and $\sup_{-\infty<t<T}|\xi(t)|\le C$, where C is an absolute constant, then the convergence in relation (14) is not only in probability but in the mean square as well, and the stochastic integral possesses finite moments of the second order.

We shall now derive several bounds on the integrals introduced above.

Lemma 1. *Let the field $\beta(\varphi,t)$ satisfy condition $\beta 1$) and the local Lipschitz condition, let $\xi_k(t)$, $k=1,2$, satisfy condition $\xi 1$), and let*

$$\|\theta_T\xi_k\|=\sup\{|\xi_k(t)|,\ t\le T\}\le N,\qquad k=1,2.$$

Then

$$(19)\quad \mathsf{E}\{|\textstyle\int_0^T\beta(\theta_t\xi_1,dt)-\int_0^T\beta(\theta_t\xi_2,dt)|^2|\mathfrak{F}_0\}\le\mathsf{E}\{\int_0^T\|\theta_t\xi_1-\theta_t\xi_2\|_*^2\Lambda_N(dt)|\mathfrak{F}_0\}.$$

Proof. As in the proof of Theorem 1 we obtain the inequality

$$\mathsf{E}\{|\sigma(\xi_1)-\sigma(\xi_2)|^2|\mathfrak{F}_0\}\le\mathsf{E}\left\{\int_{-\infty}^0\left(\sum_{k=1}^n|\xi_1(t_k+s)-\xi_2(t_k+s)|^2\Lambda_N(\Delta_k)\right)K(ds)|\mathfrak{F}_0\right\}.$$

The sum appearing under the integral sign is uniformly bounded (in ω) and tends to the limit

$$\int_0^T|\xi_1(t+s)-\xi_2(t+s)|^2\Lambda_N(dt)$$

as $|\delta|\to 0$. This fact easily implies the validity of inequality (19). □

Remark. If $\beta(\varphi,t)$ and $\xi_k(t)$, $k=1,2$, satisfy the conditions of Theorem 2 and $\|\theta_T\xi_k(t)\|\le N$, then

$$(20)\quad \mathsf{E}\{|\textstyle\int_0^T\beta(\theta_t\xi_1,dt)-\int_0^T\beta(\theta_t\xi_2,dt)||\mathfrak{F}_0\}\le\mathsf{E}\{\int_0^T\|\theta_t\xi_1-\theta_t\xi_2\|^2\Lambda_N(dt)|\mathfrak{F}_0\}.$$

The proof of this assertion is analogous to the preceding one.

The following lemma can be proven in an analogous manner:

Lemma 2. 1. *If the conditions of Theorem* 1 *are satisfied and if*

a) $\|\theta_T\xi\| \leq N$, *where* N *is an absolute constant,*

b) *there exists a continuous monotonically nondecreasing and integrable process* $\Lambda_0(t)$ *adopted to the current of* σ*-algebras* $\{\mathfrak{F}_t, t \in [0, T]\}$ *such that*

$$\mathsf{E}\{|\beta(\varphi, \Delta)|^2 | \mathfrak{F}_t\} \leq (1 + \|\varphi\|_*^2)\mathsf{E}\{\Lambda_0(\Delta t) | \mathfrak{F}_t\}, \tag{21}$$

then

$$\mathsf{E}\{|\int_0^T \beta(\theta_t\xi, dt)|^2 | \mathfrak{F}_0\} \leq \mathsf{E}\{\Lambda_0(T) + \int_0^T \|\theta_t\xi\|_*^2 \Lambda_0(dt) | \mathfrak{F}_0\}. \tag{22}$$

2. *If the conditions of Theorem* 2 *and condition* a) *are satisfied and if*

c) $$\mathsf{E}\{|\beta(\varphi, \Delta)|^2 | \mathfrak{F}_t\} \leq (1 + \|\varphi\|^2)\mathsf{M}\{\Lambda_0(\Delta) | \mathfrak{F}_t\}, \tag{23}$$

then

$$\mathsf{E}\{(|\int_0^T \beta(\theta_t\xi, dt)|)^2 | \mathfrak{F}_0\} \leq \mathsf{E}\{\Lambda_0(T) + \int_0^t \|\theta_t\xi\|^2 \Lambda_0(dt) | \mathfrak{F}_0\}. \tag{24}$$

Lemma 3. *Let* $\beta(\varphi, t)$ *and* $\xi(t)$ *satisfy the conditions of either Theorems* 1 *or* 2. *Assume also that* τ *is random time on* $[0, T]$, $\beta_\tau(\varphi, t) = \beta(\varphi, t \wedge \tau)$, $\xi_\tau(t) = \xi(t)$ *for* $t < \tau$, *and* $\xi_\tau(t) = \xi(\tau -)$ *for* $t \geq \tau$.

Then

$$\int_0^t \beta(\theta_s\xi, ds) = \int_0^t \beta_\tau(\theta_s\xi_\tau, ds) = \int_0^t \beta_\tau(\theta_s\xi, ds) \tag{25}$$

with probability 1 *on the set* $t \leq \tau$.

First we note that if the conditions of either Theorems 1 or 2 are satisfied for the time interval $[0, T]$, then they are also fulfilled for the narrower interval $[0, t]$, $t < T$, and in accordance with either Theorems 1 or 2 one can uniquely (mod P) define the integral

$$\int_0^t \beta(\theta_s\xi, ds).$$

Moreover, functions $\beta_\tau(\varphi, t)$ and $\xi_\tau(t)$ also satisfy the conditions of these theorems. Therefore the quantities appearing in relation (25) are well defined. Equalities (25) now follow from the fact that the sums in terms of which the integrals in equalities (25) are defined coincide on the set $t \leq \tau$.

Remark. We emphasize that (25) holds (mod P) also for $t = \tau$.

Lemma 4. *Let* $\beta(\varphi, t)$ *satisfy condition* $\beta 1$) *and also the uniform Lipschitz condition and let* $\xi_k(t)$, $k = 1, 2$, *satisfy condition* $\xi 1$), *and, moreover,*

$$\mathsf{E} \int_0^T \|\theta_t\xi\|_*^2 \Lambda(dt) < \infty.$$

Then inequality (19) *is valid with* $\Lambda_N = \Lambda$.

Proof. Let $\tau_N=\inf\{t:\bigcup_{k=1}^2|\xi_k(t)|\geq N\}$ and $\xi_k^N(t)=\xi_k(t)$ for $t<\tau_N$, $\xi_k^N(t)=\xi_k(\tau_N-)$ for $t\geq\tau_N$.

It follows from Lemma 3 that

$$\int_0^T\beta(\theta_s\xi_k^N,ds)=\int_0^T\beta(\theta_s\xi_k,ds)$$

for all N sufficiently large. In view of Fatou's lemma

$$\begin{aligned}\mathsf{E}\{|\int_0^T\beta(\theta_s\xi_1,ds)-\int_0^T\beta(\theta_s\xi_2,ds)|^2\}&\leq\underline{\lim}\,\mathsf{E}|\int_0^T\beta(\theta_s\xi_1^N,ds)-\int_0^T\beta(\theta_s\xi_2^N,ds)|^2\\&\leq\underline{\lim}\,\mathsf{E}\int_0^T\|\theta_t\xi_1^N-\theta_t\xi_2^N\|_*^2\Lambda(dt).\end{aligned}$$

Taking into account that $\xi_k^N(t)\to\xi_k(t)$ uniformly in t with probability 1, we obtain

$$\mathsf{E}\{|\int_0^T\beta(\theta_s\xi_1,ds)-\int_0^T\beta(\theta_s\xi_2,ds)|^2\}\leq\mathsf{E}\int_0^T\|\theta_t\xi_1-\theta_t\xi_2\|_*^2\Lambda(dt).\qquad(26)\qquad\square$$

Lemma 5. *If the field $\beta(\varphi,t)$ is linearly bounded and satisfies the local Lipschitz condition, the process $\xi(t)$ satisfies condition $\xi1$), and if, moreover,*

$$\mathsf{E}\int_0^T\|\theta_T\xi\|_*^2\Lambda_0(dt)<\infty,$$

then inequality (22) *is satisfied even without the restriction* $\|\theta_T\xi\|\leq N$.

The proof is analogous to the proof of Lemma 4.

Remark. Inequalities analogous to (22) and (26) are valid also in the case when $\beta(\varphi,t)$ and $\xi_k(t)$, $k=1,2$, satisfy the conditions of Theorem 2 and if, in addition, the field $\beta(\varphi,t)$ is linearly bounded (or correspondingly satisfies the uniform Lipschitz condition) and

$$\mathsf{E}\int_0^T\|\theta_t\xi_k\|^2\Lambda_0(dt)<\infty\qquad(\mathsf{E}\int_0^T\|\theta_t\xi_k\|^2\Lambda(dt)<\infty).$$

Lemma 6. *Assume that the random field $\beta(\varphi,t)$ satisfies the conditions of Theorem* 1 *and processes $\xi_k(t)$, $k=1,2$, that of Lemma* 1. *Then for any $\varepsilon>0$ and $N>0$*

$$\mathsf{P}\{|\int_0^T\beta(\theta_t\xi_1,dt)-\int_0^T\beta(\theta_t\xi_2,dt)|>\varepsilon\}\leq\frac{N}{\varepsilon^2}+\mathsf{P}\{\int_0^T\|\theta_t\xi_1-\theta_t\xi_2|_*^2\Lambda_N(dt)>N\}.\qquad(27)$$

Proof. Let δ be a subdivision of the interval $[0,T]$ with the subdividing points t_k, $k=1,2,\ldots,n$. The sequence of sums

$$\sum_{j=1}^k\beta(\theta_{t_{j-1}}\xi_1,\Delta_j)-\beta(\theta_{t_{j-1}}\xi_2,\Delta_j),\qquad k=0,1,\ldots,n,$$

forms a square integrable martingale and in view of Lemma 10 in Section 1,

$$\mathsf{P}\left\{\left|\sum_{j=1}^{n} \beta(\theta_{t_{j-1}}\xi_1, \Delta_j) - \sum_{j=1}^{n} \beta(\theta_{t_{j-1}}\xi_2, \Delta_j)\right| \geq \varepsilon\right\}$$

$$\leq \frac{N}{\varepsilon^2} + \mathsf{P}\left\{\sum_{j=1}^{n} \|\theta_{t_{j-1}}\xi_1 - \theta_{t_{j-1}}\xi_2\|_*^2 \Lambda_N(\Delta_j) \geq N\right\}.$$

Approaching the limit as $|\delta| \to 0$ and using the results obtained in the proof of Theorem 1 we obtain inequality (27). □

Remark. If the conditions of the remark following Lemma 1 are satisfied and $\|\theta_T\xi_1\| \leq N$, $\|\theta_T\xi_2\| \leq N$, then

(28)
$$\mathsf{P}\{|\int_0^T \beta(\theta_t\xi_1, dt) - \int_0^T \beta(\theta_t\xi_2, dt)| > \varepsilon\}$$
$$\leq \frac{N}{\varepsilon^2} + \mathsf{P}\{\int_0^T \|\theta_t\xi_1 - \theta_t\xi_2\| \Lambda_N(dt) > N\}.$$

Inequality (27) allows us to extend the definition of a stochastic line integral to a wider class of processes $\xi(t)$ than those appearing in Theorem 1. However, since, in what follows, stochastic line integrals are utilized only in the theory of those stochastic differential equations, whose solutions result in processes with sample functions belonging to $\mathscr{D}_T^m$ or $\mathscr{C}_T^m$, it is sufficient to confine ourselves to the definition of the integral given above and to classes of processes $\xi(t)$ introduced above for which these integrals exist.

The stochastic line integral as a function of the upper limit of integration. Let $\beta(\varphi, t)$ and $\xi(t)$ satisfy the conditions of either Theorems 1 or 2. If $0 \leq a < b \leq T$, then the corresponding conditions are satisfied when the interval $[0, T]$ is replaced by $[a, b]$.

Thus one can define the stochastic integral

$$\int_a^b \beta(\theta_s\xi, ds).$$

Clearly this integral is an $\mathfrak{F}_b$-measurable random variable and for $0 \leq a < b < c \leq T$

(29)
$$\int_a^b \beta(\theta_s\xi, ds) + \int_b^c \beta(\theta_s\xi, ds) = \int_a^c \beta(\theta_s\xi, ds) \quad (\text{mod } \mathsf{P}).$$

Set

$$\eta(t) = \int_0^t \beta(\theta_s\xi, ds).$$

The process $\eta(t)$ is adopted to the current of σ-algebras $\{\mathfrak{F}_t, t \in [0, T]\}$ and is uniquely determined with probability 1 for each t.

As before one can utilize the fact that the process $\eta(t)$ is not unique, and in all cases we interpret $\eta(t)$ as the separable modification of this process.

Lemma 7. *Let the conditions of Theorem* 1 *be satisfied. Then:*

a) *The process* $\eta(t)$, $t\in[0, T]$, *is a local square integrable martingale and possesses a modification with sample functions belonging to* $\mathscr{D}[0, T]$ *with probability* 1.

b) *If* $\sup_{\infty<t\leq T}|\xi(t)|\leqslant N$, $N>0$, *and condition* b) *of Lemma* 2 *is satisfied, then* $\eta(t)$ *is a square integrable martingale and*

$$\mathsf{E}\{|\textstyle\int_{t_1}^{t_2}\beta(\theta_s\xi, ds)|^2\,|\,\mathfrak{F}_{t_1}\}\leqslant \mathsf{E}\{\int_{t_1}^{t_2}(1+\|\theta_s\xi\|_*^2)\Lambda_0(dt)\,|\,\mathfrak{F}_{t_1}\}. \tag{30}$$

c) *If condition* b) *of Lemma* 2 *is satisfied with* $\Lambda_0(\Delta)=C_0\,\Delta t$, *where* C_0 *is an absolute constant and*

$$\textstyle\int_0^T \mathsf{E}\|\theta_s\xi\|_*^2\,ds<\infty,$$

then $\eta(t)$ *is a square integrable martingale; moreover,*

$$\mathsf{E}\{|\textstyle\int_{t_1}^{t_2}\beta(\theta_s\xi, ds)|^2\,|\,\mathfrak{F}_{t_1}\}\leqslant C_0\int_{t_1}^{t_2}(1+\mathsf{E}\{\|\theta_s\xi\|_*^2\,|\,\mathfrak{F}_{t_1}\})\,ds. \tag{31}$$

d) *Set*

$$\textstyle\int_0^\tau \beta(\theta_s\xi, ds)\underset{\text{Def}}{=}\eta(\tau),$$

$\beta_\tau(\varphi, t)=\beta(\varphi, t\wedge\tau)$, *where* τ *is a random time on* $\{\mathfrak{F}_t, t\in[0, T]\}$; *then*

$$\textstyle\int_0^\tau \beta(\theta_s\xi, ds)=\int_0^T\beta_\tau(\theta_s\xi, ds)\quad(\text{mod}\,\mathsf{P}). \tag{32}$$

e) *For any* $\varepsilon>0$ *and* $N>0$

$$\mathsf{P}\{\sup_{0\leqslant t\leqslant T}|\textstyle\int_0^t\beta(\theta_s\xi, ds)|>\varepsilon\}\leqslant\dfrac{N}{\varepsilon^2}+\mathsf{P}\{\int_0^T(1+\|\theta_s\xi\|_*^2)\Lambda_0(ds)>N\}. \tag{33}$$

Proof. First we shall prove assertion b). Inequality (30) follows directly from Lemma 2. Since in this case the sum $\sum_{a\leqslant t_k\leqslant b}\beta(\theta_{t_k}\xi, \Delta_k)$ is uniformly integrable, one can approach the limit in the equality

$$\mathsf{E}\Big\{\sum_{a\leqslant t_k\leqslant b}\beta(\theta_{t_k}\xi, \Delta_k)\,|\,\mathfrak{F}_a\Big\}=0$$

as $|\delta|\to 0$ and obtain

$$\mathsf{E}\{\textstyle\int_a^b\beta(\theta_s\xi, ds)\,|\,\mathfrak{F}_a\}=0.$$

Therefore $\eta(t)$ is a square integrable martingale. Assertion c) is proved analogously.

To prove assertion d) we first assume that $\sup_t |\xi(t)| \le N$. Set

$$\sigma(t) = \sum_{k=1}^{j} \beta(\theta_{t_{k-1}}\xi, \Delta_k) + \beta(\theta_{t_j}\xi, t) - \beta(\theta_{t_j}\xi, t_j),$$

if $t \in (t_j, t_{j+1}]$. Then by definition,

$$\text{P-lim}\, \sigma(\tau) = \int_0^T \beta_\tau(\theta_s\xi, ds).$$

On the other hand, $|\eta(\tau) - \sigma(\tau)| \le \sup_{0\le t\le T} |\eta(t) - \sigma(t)|$, and since $\eta(t) - \sigma(t)$ is a square integrable separable martingale, in view of the remark following Theorem 2 we have

$$\mathsf{E} \sup |\eta(t) - \sigma(t)|^2 \le 4\mathsf{E}|\eta(T) - \sigma(T)|^2 \to 0.$$

Thus

$$\eta(\tau) = \text{P-}\lim_{\lambda\to 0} \sigma(\tau) = \int_0^T \beta_\tau(\theta_s\xi, ds),$$

which proves equality (32) in this particular case.

To study the general case we shall introduce the random time τ_N, which is the time of the first exit of the process $\xi(t)$ outside the sphere of radius N (here $\tau_N = T$ provided $|\xi(t)| \le N$ for all $t < T$) and set $\xi_N(t) = \xi(t)$ for $t < \tau$ and $\xi_N(t) = \xi(\tau_N -)$ for $t \ge \tau$ and $\beta_{\tau_N}(\varphi, t) = \beta(\varphi, t \wedge \tau_N)$.

Utilizing Lemma 3 and formula (32) we have

$$\eta(t\wedge\tau_N) = \int_0^{t\wedge\tau_N} \beta(\theta_s\xi, ds) = \int_0^{t\wedge\tau_N} \beta_{\tau_N}(\theta_s\xi_N, ds) = \int_0^t \beta_{\tau_N}(\theta_s\xi_N, ds),$$

so that in view of b) $\eta(t\wedge\tau)$ is a square integrable martingale. Since $\lim_{N\to\infty} \tau_N = T$, it follows that $\eta(t)$ is a local square integrable martingale. The existence of a modification with sample functions belonging to $\mathscr{D}[0, T]$ with probability 1 follows from the general properties of local martingales.

Next, since $\tau_N = T$ with probability 1 for N sufficiently large, $\eta(\tau) = \text{P-lim}\, \eta(\tau\wedge\tau_N)$ for any random time τ. Consequently,

$$\eta(\tau) = \text{P-lim}\, \eta(\tau\wedge\tau_N) = \text{P-lim} \int_0^{\tau\wedge\tau_N} \beta_{\tau\wedge\tau_N}(\theta_s\xi_{\tau\wedge\tau_N}, ds) = \text{P-lim} \int_0^T \beta_{\tau\wedge\tau_N}(\theta_s\xi, ds).$$

On the other hand, in view of Lemma 3

$$\mathsf{P}\Big\{\int_0^T \beta_{\tau\wedge\tau_N}(\theta_s\xi, ds) - \int_0^T \beta_\tau(\theta_s\xi, ds) \ne 0\Big\} \le \mathsf{P}\{\tau_N < \tau\},$$

and since $\mathsf{P}(\tau_N < \tau) \to 0$ as $N \to \infty$ we arrive at formula (32).

Finally inequality 33) follows from the fact that $\eta(t) = \int_0^t \beta(\theta_s\xi, ds)$ is a local square integrable martingale and the remark following Lemma 10 in Section 1. □.

Remark. If the conditions of Theorem 2 are satisfied, then a separable modification of the process $\eta(t)$ is a continuous process and the inequalities (30), (31), and (33) are satisfied when $\|\theta_s\xi\|_*$ is replaced in these inequalities by $\|\theta_s\xi\|$.

We shall now compute the characteristic of the stochastic integral $\eta(t)$.

Lemma 8. *Assume that the conditions of Theorem* 1 *are satisfied and that*

$$\langle\beta(\varphi,\cdot),\beta(\psi,\cdot)\rangle_t=\int_0^t b(\varphi,\psi,s)\,ds$$

and

$$|b(\varphi,\varphi,t)-2b(\varphi,\psi,t)+b(\psi,\psi,t)|\leq\lambda_N(t)\|\varphi-\psi\|_*^2$$

for $\|\varphi\|\leq N$ *and* $\|\psi\|\leq N$, *where* $\lambda_N(t)$ *is a nonnegative random process adopted to the current of* σ-*algebras* $\{\mathfrak{F}_t,\ t\in[0,T]\}$ *and integrable with probability* 1 *on the interval* $[0,T]$. *Set*

$$\eta_i(t)=\int_0^t\beta(\theta_s\xi_i,ds),$$

where $\xi_i(t)$ $(i=1,2)$ *also satisfy the conditions of Theorem* 1. *Then*

$$\langle\eta_1(\cdot),\eta_2(\cdot)\rangle_t=\int_0^t b(\theta_s\xi_1,\theta_s\xi_2,s)\,ds. \tag{34}$$

Proof. We introduce "integral sums with a variable limit of summation"

$$\sigma_i(t)=\sum_{k=1}^{j}\beta(\theta_{t_{k-1}}\xi_i,\Delta_k)+\beta(\theta_{t_j}\xi,t)-\beta(\theta_{t_j}\xi,t_j)\quad\text{for } t\in(t_j,t_{j+1}),$$

where $\Delta_k=(t_{k-1},t_k]$. It is easy to verify that

$$\langle\sigma_1,\sigma_2\rangle_t=\sum_{k=1}^{j}\int_{t_{k-1}}^{t_k}b(\theta_{t_{k-1}}\xi_1,\theta_{t_{k-1}}\xi_2,s)\,ds+\int_{t_j}^{t}b(\theta_{t_j}\xi_1,\theta_{t_j}\xi_2,s)\,ds.$$

Now observe that

$$\begin{aligned}&\Delta\langle(\beta(\varphi,\cdot)-\beta(\psi,\cdot)),(\beta(\varphi,\cdot)-\beta(\psi,\cdot))\rangle_t\\&=\int_t^{t+\Delta t}[b(\varphi,\varphi,s)-2b(\varphi,\psi,s)-b(\psi,\psi,s)]\,ds\\&\leq\|\varphi-\psi\|_*\int_t^{t+\Delta t}\lambda_N(s)\,ds\end{aligned}$$

and

$$\Delta\langle\beta(\varphi,\cdot),\beta(\psi_1,\cdot)-\beta(\psi_2,\cdot)\rangle_t=\int_t^{t+\Delta t}[b(\varphi,\psi_1,s)-b(\varphi,\psi_2,s)]\,ds.$$

This implies (for almost all s)

$$\begin{aligned}&|b(\varphi,\psi_1,s)-b(\varphi,\psi_2,s)|\\&\leq[b(\varphi,\varphi,s)[b(\psi_1,\psi_1,s)-2b(\psi_1,\psi_2,s)+b(\psi_2,\psi_2,s)]^{1/2}\\&\leq[b(\varphi,\varphi,s)\lambda_N(s)]^{1/2}\|\psi_1-\psi_2\|_*,\end{aligned}$$

and

$$|b(\varphi_1, \psi_1, s) - b(\varphi_2, \psi_2, s)|$$
$$\leq \lambda_N(s)^{1/2}(\|\varphi_1 - \varphi_2\|_* b(\psi_1, \psi_1, s)^{1/2} + \|\psi_1 - \psi_2\|_* b(\varphi_2, \varphi_2, s)^{1/2})$$

for almost all s. Thus the function $b(\varphi, \psi, s)$ is (for almost all s) a continuous function of arguments φ and ψ (relative to the seminorm $\|\cdot\|_*$). It is easy to verify—using the same arguments as in the proof of Theorem 1—that

$$\langle \sigma_1, \sigma_2 \rangle_t \to \int_0^t b(\theta_s \xi_1, \theta_s \xi_2, s)\, ds$$

with probability 1 as $|\delta| \to 0$. Therefore the convergence of $\sigma_i(t)$ to $\eta_i(t)$ implies equality (34). □

Remark. If we assume that the conditions of Theorem 2 and those of Lemma 8 are satisfied and that $\|\varphi - \psi\|_*$ is replaced by $\|\varphi - \psi\|$ in these conditions, then equality (34) is still valid.

The existence and uniqueness theorems of solutions of stochastic differential equations. Let a current of σ-algebras $\{\mathfrak{F}_t, t \in [0, T]\}$ and random function $\alpha(\varphi, t)$ and $\beta(\varphi, t)$ with values in $\mathcal{R}^m$, adopted to $\{\mathfrak{F}_t\}$ with $\varphi \in \mathcal{D}^m$, $t \in [0, T]$, be given. A random time τ, $0 < \tau \leq T$, on $\{\mathfrak{F}_t\}$ and a random process $\xi(t)$ defined for $t \in [0, \tau)$, $\mathfrak{F}_t \times \mathfrak{T}_t$-progressively measurable and satisfying with probability 1 relations

$$\xi(t) = \varphi(t) \quad \text{for } t < 0,$$
$$\xi(t) = \xi(0) + \int_0^t \alpha(\theta_s \xi, s)\, ds + \int_0^t \beta(\theta_s \xi, ds), \qquad t \geq 0, \tag{35}$$

for each $t < \tau$, is called a *solution of the stochastic differential equation*

$$d\xi = \alpha(\theta_t \xi, t)\, dt + \beta(\theta_t \xi, dt), \qquad t \geq 0, \tag{36}$$

satisfying the "initial condition"

$$\xi(t) = \varphi(s), \qquad s \leq 0.$$

It is assumed here that integrals in the right-hand side of (35) are well defined, the first as a Lebesgue integral and the second as a stochastic integral.

The random variable τ is called the *lifetime* of the process $\xi(t)$ (i.e., of a solution of a stochastic differential equation).

Equation (36) is called *regular* on $[0, T]$ provided it possesses a unique solution of the whole time interval $[0, T]$ (i.e., if a unique solution of equation (36) with $\tau = T$ exists).

We now introduce some general assumptions on functions $\alpha(\varphi, t)$ and $\beta(\varphi, t)$ under which the right-hand side of equality (35) will be well defined for a sufficiently wide class of processes $\xi(t)$. Observe that there is no point to consider the most general classes of processes $\xi(t)$ since the right-hand side of equation (35) represents a process which possesses a continuous modification or a modification with sample functions in $\mathscr{D}_T^m$ and hence the process $\xi(t)$ must also possess these properties.

We shall first discuss the function $\alpha(\varphi, t)$. We introduce two sets of assumptions:

$\alpha 1)$ a) The function $\alpha(\varphi, s) = \alpha(\varphi, s, \omega)$ is defined on $\mathscr{D}^m \times [0, T] \times \Omega$ and is $\mathfrak{B}_{\mathscr{D}^m} \times \mathfrak{T}_t \times \mathfrak{F}_t$-measurable for $s \in [0, t]$.

$\alpha 1)$ b) $\alpha(\varphi, \cdot) \in \mathscr{D}^m[0, T]$ with probability 1 for a fixed ω.

$\alpha 1)$ c) For a fixed ω the family of functions $\{\alpha(t, \cdot), t \in [0, T]\}$ of argument φ is uniformly continuous on $\mathscr{D}^m$ relative to metric $\rho_{\mathscr{D}}$.

$\alpha 2)$ The function $\alpha(\varphi, s)$ satisfies assumptions $\alpha 1)$ provided $\mathscr{D}^m$, $\mathfrak{B}_{\mathscr{D}^m}$, and $\mathscr{D}[0, T]$ are replaced by $\mathscr{C}^m$, $\mathfrak{B}_{\mathscr{C}^m}$, and $\mathscr{C}^m[0, T]$, respectively.

We say that the function $\alpha(\varphi, t)$ is *linearly bounded* (relative to a uniform norm or a seminorm if in the succeeding inequalities the norm $\|\cdot\|$ can be replaced by the seminorm $\|\cdot\|_*$) provided a continuous monotonically nondecreasing process $\lambda_0(t)$ exists adopted to $\{\mathfrak{F}_t, t \in [0, T]\}$ such that $\lambda_0(T) < \infty$ with probability 1 and

$$(37) \qquad \left|\int_a^b \alpha(\varphi, t)\, dt\right| \leq (1 + \|\varphi\|) \int_a^b \lambda_0(t)\, dt.$$

We say that $\alpha(\varphi, t)$ satisfies the *local Lipschitz condition* (in a uniform norm or seminorm) if for any $N > 0$ there exists a monotonically nondecreasing process $\lambda_N(t)$ adopted to $\{\mathfrak{F}_t, t \in [0, T]\}$ such that

$$(38) \qquad \left|\int_a^b [\alpha(\varphi, t) - \alpha(\psi, t)]\, dt\right| \leq \|\varphi - \psi\| \int_a^b \lambda_N(t)\, dt$$

for all φ and ψ satisfying $\|\varphi\| \leq N$ and $\|\psi\| \leq N$. If we choose the process $\lambda(t)$ which is independent of N in place of $\lambda_N(t)$ then we shall refer to the process $\alpha(\varphi, t)$ as one satisfying the *uniform Lipschitz condition.*

The class of processes $\alpha(\varphi, t)$ which satisfy conditions $\alpha 2)$, (37) and (38) will be denoted by $S_\alpha^c(\lambda_0, \lambda_N)$. We denote by $S_\alpha(\lambda_0, \lambda_N)$ the class of processes which satisfy $\alpha 1)$ and the conditions obtained from (37) and (38) when the uniform norm $\|\cdot\|$ is replaced by $\|\cdot\|_*$ in the corresponding inequalities. In the case when we shall be dealing with random functions $\alpha(\varphi, t)$ satisfying only one of the inequalities (37) or (38), for instance (37), we shall write $\alpha(\varphi, t) \in S_0^c(\lambda_0, \cdot)$ and analogously in other cases.

Note that $\varphi_t = \theta_t\psi$ ($\psi \in \mathscr{D}^m$, $t \in [0, T]$), with values in $\mathscr{R}^m$ is a Borel function in argument t.

Indeed if B is a cylinder in $\mathscr{D}^m$ with the basis $B = \prod_{i=1}^n B_i$ in coordinates $(s_1, \ldots, s_n)$, $s_k \leq 0$, then

$$\{t\colon \varphi_t \in B\} = \bigcap_{i=1}^n \{t\colon \varphi(t + s_i) \in B_i\}.$$

Since the sets $\{z: \varphi(z) \in B_i\} = Z_i$ are Borel sets provided $B_i \in \mathcal{R}^m$ are such, the set $\{t: \varphi_t \in B\} = \bigcap_{i=1}^n \{Z_i - s_i\}$, where $Z - s$ denotes the set $\{z: z + s \in Z\}$, will also be a Borel set. Thus, if $g(\varphi, t)$ is a $\mathfrak{B}_{\mathcal{D}^m} \times \mathfrak{T}$-measurable function in the arguments (φ, t) where $\mathfrak{B}_{\mathcal{D}^m}$ is the minimal σ-algebra generated by the cylinders in $\mathcal{D}^m$ and $\mathfrak{T}$ is a σ-algebra of Borel sets in $[0, T]$, then $g(\theta_t\varphi, t)$ will be a Borel function in argument t.

Consequently, if the field $\alpha(\varphi, t)$ satisfies conditions $\alpha 1)$ or $\alpha 2)$, then the integral

$$\int_0^T \alpha(\theta_t\psi, t)\, dt$$

exists with probability 1.

Next, if $\alpha(\varphi, t) \in S_\alpha(\lambda_0, \cdot\,)$, then we have for $0 \leqslant a < b \leqslant T$,

$$\left|\int_a^b \alpha(\theta_t\psi, t)\, dt\right| \leqslant \int_a^b (1 + \|\theta_t\psi\|_*)\lambda_0(t)\, dt \pmod{\mathsf{P}}. \tag{39}$$

The proof follows easily from the fact that $\theta_t\psi$, $t \in [0, T]$ is a continuous function with values in $\mathcal{D}^m$ with respect to metric $\rho_{\mathcal{D}}$ in $\mathcal{D}^m$, and $\alpha(\varphi, t)$ is a continuous function of argument φ (uniformly in t).

Analogously, if $\alpha(\varphi, t) \in S_\alpha^c(\lambda_0, \cdot\,)$, then

$$\left|\int_a^b \alpha(\theta_t\psi, t)\, dt\right| \leqslant \int_a^b (1 + \|\theta_t\psi\|)\lambda_0(t)\, dt. \tag{40}$$

If, however, $\alpha(\varphi, t) \in S_\alpha(\,\cdot\,, \lambda_N)$ and $\|\psi_1\| \vee \|\psi_2\| \leqslant N$, then

$$\left|\int_a^b \alpha(\theta_t\psi, t)\, dt - \int_a^b \alpha(\theta_t\psi_2, t)\, dt\right| \leqslant \int_a^b \|\theta_t(\psi_1 - \psi_2)\|_* \lambda_N(t)\, dt, \tag{41}$$

and an analogous inequality is valid for $\alpha(\varphi, t) \in S_\alpha^c(\,\cdot\,, \lambda_N)$.

As far as the integral $\int_0^t \beta(\theta_s\xi, ds)$ is concerned, conditions for its existence and its properties were discussed in the preceding subsections.

We introduce the notation for the classes of fields $\beta(\varphi, t)$ analogous to the notation for the classes of functions $\alpha(\varphi, t)$ introduced above. Namely, we write $\beta(\varphi, t) \in S_\beta(\lambda_0, \lambda_N)$ (or $\beta(\varphi, t) \in S_\beta^c(\lambda_0, \lambda_N)$) provided the field $\beta(\varphi, t)$ satisfies conditions $\beta 1)$ $(\beta 2))$, is linearly bounded, satisfies the local Lipschitz condition relative to a seminorm (norm), and the dominating processes $\Lambda_0(t)$ and $\Lambda_N(t)$ are absolutely continuous with $\lambda_0(t) = \Lambda_0'(t)$ and $\lambda_N(t) = \Lambda_N'(t)$.

Set

$$A(\varphi, t) = \int_0^t \alpha(\varphi, s)\, ds + \beta(\varphi, t),$$

and write $A(\varphi, t) \in S(\lambda_0, \lambda_N)$ $(S^c(\lambda_0, \lambda_N))$ provided

$$\alpha(\varphi, t) \in S_\alpha(\lambda_0, \lambda_N) \quad \text{and} \quad \beta(\varphi, t) \in S_\beta(\lambda_0, \lambda_N)$$
$$(\alpha(\varphi, t) \in S_\alpha^c(\lambda_0, \lambda_N) \quad \text{and} \quad \beta(\varphi, t) \in S_\beta^c(\lambda_0, \lambda_N)).$$

Let the process $\xi(t)$, $t \in [0, T]$, be adopted to a current of σ-algebras $\{\mathfrak{F}_t, t \in [0, T]\}$ with sample functions belonging with probability 1 to $\mathcal{D}^m[0, T]$. We complete the definition of $\xi(t)$ for $t \leqslant 0$ by setting $\xi(t) = \varphi(t)$ for $t \leqslant 0$, where $\varphi(t)$ is a given function in $\mathscr{D}^m$. Assume that $\alpha(\varphi, t)$ satisfies condition $\alpha 1$ a) and that $\beta(\varphi, t) \in S_\beta(\,\cdot\,, \lambda_N)$. In what follows these conditions will always be assumed to be valid unless specified otherwise.

Define a new process $\eta(t)$, $t \in (-\infty, T]$, by setting

$$\eta(t) = \varphi(t), \qquad t \leqslant 0$$

$$\eta(t) = \varphi(0) + \int_0^t A(\theta_s\xi, ds), \qquad t \geqslant 0,$$

where

$$\int_0^t A(\theta_s\xi, ds) \underset{\text{Def}}{=} \int_0^t \alpha(\theta_s\xi, s)\, ds + \int_0^t \beta(\theta_s\xi, ds), \qquad t \in [0, T]. \tag{42}$$

The stochastic line integral in the right-hand side of (42) is interpreted here as a modification with sample functions belonging to $\mathscr{D}^m[0, T]$. We shall denote by I the correspondence $\xi \to \eta$, i.e., $\eta(t) = I(t, \xi)$.

Lemma 9. If $A(\varphi, t) \in S(C, \lambda_N)$ where C is an absolute constant and $\sup_{0 \leqslant t \leqslant T} \mathsf{E}|\xi(t)|^2 < \infty$, then

$$\mathsf{E}\{\sup_{0 \leqslant h \leqslant a} |I(t+h, \xi) - I(t, \xi)|^2 \mid \mathfrak{F}_t\} \leqslant C'[a(1 + \|\varphi\|^2) + \int_t^{t+a} z(s)\, ds], \tag{43}$$

where C' depends on C, K, and T only, and

$$z(s) = \sup_{0 \leqslant u \leqslant s} \mathsf{E}\{|\xi(u)|^2 \mid \mathfrak{F}_t\}.$$

Proof. Since

$$\sup_{0 \leqslant h \leqslant a} |I(t+h, \xi) - I(t, \xi)|^2$$
$$\leqslant 2 \sup_{0 \leqslant h \leqslant a} \{|\int_t^{t+h} \alpha(\theta_s\xi, s)\, ds + \int_t^{t+h} \beta(\theta_s\xi, ds)|^2\},$$

and taking into account the fact that for separable square integrable martingales $\eta(t)$

$$\mathsf{E}\Big\{\sup_{t \leqslant s \leqslant t+a} |\eta(s)|^2 \mid \mathfrak{F}_t\} \leqslant 4\mathsf{E}\{|\eta(t+a)|^2 \mid \mathfrak{F}_t\Big\},$$

and utilizing inequalities (22) and (40), we obtain

$$\mathsf{E}\Big\{\sup_{0 \leqslant h \leqslant a} |I(t+h, \xi) - I(t, \xi)|^2 \mid \mathfrak{F}_t\Big\}$$
$$\leqslant 2(2C^2a^2 + 4C^2a + 2C^2\mathsf{E}\{(\int_t^{t+a} \|\theta_s\xi\|_*\, ds)^2 \mid \mathfrak{F}_t\}$$
$$+ 4C^2 \int_t^{t+a} \mathsf{E}\{\|\theta_s\|_*^2 \mid \mathfrak{F}_t\}\, ds$$
$$\leqslant C'(a + \int_t^{t+a} \mathsf{E}\{\|\theta_s\xi\|_*^2 \mid \mathfrak{F}_t\}\, ds),$$

where $C'=4C^2(T+2)$. On the other hand,

$$\mathsf{E}\{\|\theta_s\xi\|_*^2 \mid \mathfrak{F}_t\} = \int_{-\infty}^0 \mathsf{E}\{|\xi(s+u)|^2 \mid \mathfrak{F}_t\} K(du) \leq K(\|\varphi\|^2 + z(s)),$$

which together with the preceding inequality proves the lemma. □

Remark. If $A(\varphi, t) \in S^c(C, \lambda_N)$ then

$$\mathsf{E} \sup_{0 \leq h \leq a} |I(t+h, \xi) - I(t, \xi)|^2 \leq C'[(1+\|\varphi\|)^2 a + \int_t^{t+a} Z(s)\, ds], \tag{44}$$

where $Z(s) = \mathsf{E} \sup_{0 \leq t \leq s} |\xi(t)|^2$.

The proof of inequality (44) is analogous to the proof of the preceding lemma.

The following lemma is proved in the same manner as Lemma 8.

Lemma 10. *If* $\sup_{0 \leq t \leq T} \mathsf{E}|\xi_k(t)|^2 < \infty$, $k = 1, 2$, *and* $A(\varphi, t) \in S(\cdot, C)$, *then*

$$\mathsf{E} \sup_{0 \leq t \leq a} |I(t, \xi_1) - I(t, \xi_2)|^2 \leq C'' \int_0^a v(t)\, dt,$$

where C'' is a constant depending on C and T only and

$$v(t) = \sup_{0 \leq s \leq t} \mathsf{E}|\xi_1(s) - \xi_2(s)|^2.$$

If, however, $A(\varphi, t) \in S^c(\cdot\ C)$, *then*

$$\mathsf{E} \sup_{0 \leq t \leq a} |I(t, \xi_1) - I(t, \xi_2)|^2 \leq C''' \int_0^a V(t)\, dt,$$

where $V(t) = \sup_{0 \leq s \leq t} \mathsf{E}|\xi_1(s) - \xi_2(s)|^2$.

To simplify the writing, we introduce the following notation.

Denote by $H^*(H^c)$ the space of random processes satisfying conditions ($\xi 1$ ($\xi 2$)), and by H_2^* (H_2^c) the subspace H^* (H^c) consisting of processes satisfying the additional condition

$$\|\xi(\cdot)\|_2 = \left\{ \sup_{0 \leq t \leq T} \mathsf{E}|\xi(t)|^2 \right\}^{1/2} < \infty \quad \left(\mathsf{E} \sup_{0 \leq t \leq T} |\xi(t)|^2 < \infty\right).$$

We also note the following elementary lemma which will be used repeatedly below.

Lemma 11. *If $z(t)$ is a function bounded on* $[0, T]$ *and*

$$z(t) \leq A + B \int_0^t z(s)\, ds, \qquad B > 0,$$

then

$$z(t) \leq A\, e^{Bt}.$$

Proof. Indeed, clearly

$$\begin{aligned} z(t) &\leqslant A + B \int_0^t \left(A + B \int_0^{t_1} z(s)\, ds\right) dt_1 \\ &\leqslant A + ABt + AB^2 \frac{t^2}{2} + \cdots + AB^n \frac{t^n}{n!} \\ &\quad + B^{n+1} \int_0^t \int_0^{t_1} \cdots \int_0^{t_n} z(s)\, ds\, dt_n \cdots dt_1. \end{aligned}$$

Approaching the limit as $n \to \infty$ we obtain the required assertion. □

Theorem 3. *Let $A(\varphi, t) \in S(C, C)$. A stochastic differential equation* (36) *under an arbitrary initial condition $\varphi \in \mathscr{D}_0^m$ is regular in H_2^*, i.e., it possesses a unique solution in H_2^* defined for all $t \in [0, T]$. This solution has the properties*

$$\mathsf{E}\{\sup_{0 \leqslant s \leqslant T} |\xi(s)|^2 \,|\, \mathfrak{F}_0\} \leqslant A(1 + \|\varphi\|^2), \tag{45}$$

$$\mathsf{E}\{\sup_{t \leqslant s \leqslant t+h} |\xi(s) - \xi(t)|^2 \,|\, \mathfrak{F}_t\} \leqslant B(1 + \|\varphi\|^2 + \sup_{a \leqslant u \leqslant t} |\xi(u)|^2) h, \tag{46}$$

where A and B are constants which depend on C, T and K only.

Proof. Introduce the norm

$$\|\xi(\cdot)\|_2 = \{\sup_{0 \leqslant t \leqslant T} \mathsf{E}|\xi(t)|^2\}^{1/2}$$

in the space H_2^* and consider H_2^* as a subset of the space H_2 of random functions $\xi(t)$, $t \in [0, T]$ with values in $\mathscr{R}^m$ adopted to the current of σ-algebras $\{\mathfrak{F}_t, t \geqslant 0\}$ such that $\|\xi(\cdot)\|_2 < \infty$, where the norm is defined by the preceding relation. The space H_2 is a complete space (as opposed to the space H_2^*). As it follows from inequality (43) operator I maps H_2^* into itself and Lemma 9 shows that some power of operator I is a contracting operator. Starting with an arbitrary process $\xi_0(t) \in H_2^*$ ($\xi_0(0) = \varphi(0)$) we construct the successive approximations

$$\xi_{n+1}(t) = I(t, \xi_n), \qquad t \in [0, T],$$

and set

$$v_n(t) = \mathsf{E} \sup_{s \leqslant t} |\xi_{n+1}(s) - \xi_n(s)|, \qquad n = 1, 2, \ldots,$$

$$v_0(t) = \sup_{0 \leqslant s \leqslant t} \mathsf{E}|\xi_1(s) - \xi_0(s)|^2, \qquad V_0 = v_0(T).$$

It follows from Lemma 9 that

$$v_1(t) \leqslant C'' V_0 t, \ldots, v_n(t) \leqslant (C'') V_0 \frac{t^n}{n!}.$$

If we define

$$\varepsilon_n = \left[V_0 \frac{(C''T)^M}{n!} \right]^{1/3}$$

then the Chebyshev inequality implies that $\mathsf{P}\{\sup_{0\leq t\leq T} |\xi_{n+1}(t)-\xi_n(t)| > \varepsilon_n\} \leq \varepsilon_n$, and since the series $\sum_{n=1}^{\infty} \varepsilon_n$ is convergent so is the series

$$\sum_{n=1}^{\infty} \sup_{0\leq t\leq T} |\xi_{n+1}(t)-\xi_n(t)|$$

with probability 1. Thus $\lim \xi_n(t) = \xi(t)$ exists with probability 1 and, moreover, uniformly in $t \in [0, T]$. We complete the definition of $\xi(t)$ for $t<0$ by setting $\xi(t)=\varphi(t)$. Under this definition $\xi(t)$ satisfies the conditions $\xi 1)$ and $\mathsf{E} \sup_{0\leq t\leq T} |\xi(t)|^2 \leq \infty$. Moreover, $\mathsf{E} \sup_{0\leq t\leq T} |\xi(t)-\xi_n(t)|^2 \to 0$. Indeed,

$$\begin{aligned}
\mathsf{E} \sup_{0\leq t\leq T} |\xi(t)-\xi_n(t)|^2 &\leq \mathsf{E} \lim_{m\to\infty} \sup_t |\xi_{n+m}(t)-\xi_n(t)|^2 \\
&\leq \lim_{m\to\infty} \mathsf{E}\left(\sum_{k=n}^{n+m-2} \sup_t |\xi_{k+1}-\xi_k| \cdot \sqrt{k(k-1)} \cdot \frac{1}{\sqrt{k(k-1)}} \right)^2 \\
&\leq \sum_n^{\infty} k(k-1) v_k(T) \cdot \sum_{k=n}^{\infty} \frac{1}{k(k-1)} \\
&\leq \frac{V_0 C''^2 T^2}{n-1} \cdot \sum_n^{\infty} \frac{T^{k-2}}{(k-2)!} \to 0 \quad \text{as } n\to\infty.
\end{aligned}$$

It is now easy to justify the limit transition in the relation

$$\xi_{n+1}(t) = \varphi(0) + \int_0^t \alpha(\theta_s\xi_n, s)\, ds + \int_0^t \beta(\theta_s\xi_n, ds).$$

Indeed, it follows from the uniform convergence of $\xi_n(t)$ to $\xi(t)$ that $\theta_s\xi_n(u)$ converges to $\theta_s\xi(s)$ uniformly in u. Therefore $\alpha(\theta_s\xi_n, s) \to \alpha(\theta_s\xi, s)$ for all $s \in [0, T]$ with probability 1. Hence with probability 1

$$\int_0^t \alpha(\theta_s\xi_n, s)\, ds \to \int_0^t \alpha(\theta_s\xi, s)\, ds.$$

Furthermore, in view of the above,

$$\mathsf{E}\left|\int_0^t \beta(\theta_s\xi_n, ds) - \int_0^t \beta(\theta_s\xi, ds)\right|^2 \leq C^2 \int_0^t \mathsf{E}|\xi_n(s)-\xi(s)|^2\, ds \to 0.$$

Thus, $\xi(t)$ is a solution of the equation

$$\xi(t) = \varphi(0) + \int_0^t \alpha(\theta_s\xi, s)\, ds + \int_0^t \beta(\theta_s\xi, ds)$$

with probability 1 for each t. Since the functions in the right-hand and left-hand sides of the equality are continuous from the right, this equality is valid with probability 1 for all $t \in [0, T]$.

We now verify inequalities (45) and (46). Set $z(t) = \mathsf{E}\{\sup_{0 \le s \le t} |\xi(s)|^2 \mid \mathfrak{F}_0\}$. Taking Lemma 9 into account, we obtain

$$\begin{aligned} z(t) &\le 2|\varphi(0)|^2 + 2\mathsf{E} \sup |I(s, \xi)|^2 \\ &\le 2|\varphi(0)|^2 + 2C_1^*[t(1 + \|\varphi\| + K \int_0^t z(s)\, ds] \\ &\le C_2(\|\varphi\|^2 + t + \int_0^t z(s)\, ds), \end{aligned}$$

where C_2 is a new constant which depends on C, K, and T only. The last inequality implies that

$$\frac{z(t) + 1}{\|\varphi\|^2 + t + \int_0^t z(s)\, ds + (1/C_2)} \le C_2$$

or

$$z(t) \le A(\|\varphi\|^2 + 1)\, e^{C_2 t}.$$

Analogously for

$$z_1(t) = \mathsf{E}\{\sup_{a \le s \le t} |\xi(s) - \xi(a)|^2 \mid \mathfrak{F}_a\}$$

we have

$$z_1(t) \le C_3(t(1 + \|\varphi\|^2) + \int_0^t z_1(s)\, ds);$$

this implies that

$$z_1(t) \le (1 + \|\varphi\|^2 + 2a \sup_{0 \le u \le a} |\xi(u)|^2)(e^{C_3(t-a)} - 1).$$

We shall now prove the uniqueness of the solution of equation (36) in H_2^*. If there were two solutions $\xi(t)$ and $\eta(t)$ then in view of Lemma 10

$$V(t) \le C'' \int_0^t V(s)\, ds,$$

where

$$V(t) = \mathsf{E} \sup_{0 \le s \le t} |\xi(s) - \eta(s)|^2,$$

$$V(t) \le C''T \sup_{0 \le s \le T} \mathsf{E}|\xi(s) - \eta(s)|^2 = C'''.$$

Integrating the inequality just obtained, we have

$$\begin{aligned} V(t) &\le C''^2 \int_0^t dt_1 \int_0^{t_1} V(t_2)\, dt_2 \le \cdots \le C''^n \int_0^t \int_0^{t_1} \cdots \int_0^{t_{n-1}} V(t_n)\, dt_n\, dt_{n-1} \ldots dt_1 \\ &\le C'''C''^n \frac{t^n}{n!}. \end{aligned}$$

Thus, $V(t) = 0$.

The theorem is proved. $\square$

Remark 1. Let the function $\alpha(\varphi, t)$ and $\beta(\varphi, t)$ satisfy the conditions of Theorem 3. Consider the equation

$$\xi(t) = \varphi(t) + \int_0^t \alpha(\theta_s\xi, s)\,ds + \int_0^t \beta(\theta_s\xi, ds), \tag{47}$$

where the function $\varphi(t)$ possesses the following properties: its contraction to the half-line $(-\infty, 0]$ is a fixed function in $\mathscr{D}_0^m$, and the contraction to the interval $(0, T]$ belongs to H_2^*.

It is clear that the part of the proof of Theorem 3 related to the existence and uniqueness of the solution for equation (36) in H_2^* is carried over without alterations to equation (47). Thus the following result is valid:

Under the previous assumptions equation (47) *possesses a unique solution in* H_2^*.

Remark 2. If $A(\varphi, t) \in S^c(C, C)$ and $\varphi(t) \in H_2^c$, then equation (47) possesses a unique solution in H_2^c defined for all $t \in [0, T]$ which also satisfies inequalities (45) and (46).

The proof differs only slightly from the proof of Theorem 3; one must just proceed from an initial approximation $\xi_0(t)$ such that $\mathsf{E}\sup_{0 \leq t \leq T} |\xi_0(t)|^2 < \infty$.

In order to generalize the theorem on the existence and uniqueness of the solution for equation (36) the following result is required.

Theorem 4. *Let $A_i(\varphi, t) \in S(\cdot, \lambda_N)$ or $A_i(\varphi, t) \in S^c(\cdot, \lambda_N)$, and for $\|\varphi\| \leq N$ and $\|\psi\| \leq N$ with $t < \tau$, where τ is an $\mathfrak{F}_t$-random time, let the following relations be satisfied*:

$$\alpha_1(\varphi, t) = \alpha_2(\varphi, t) = \alpha(\varphi, t),$$
$$\beta_1(\varphi, t) = \beta_2(\varphi, t) = \beta(\varphi, t).$$

Then if $\xi_i(t)$, $i = 1, 2$, are solutions of equations

$$d\xi_i(t) = \alpha_i(\theta_t\xi_i, t)\,dt + \beta_i(\theta_t\xi_i, dt), \qquad t > 0,$$
$$\xi_i(s) = \varphi(s), \qquad s \leq 0,$$

such that $\sup_{0 \leq t \leq T} |\xi_i(t)| < \infty$ *with probability* 1, *then for* $\sup_{s \leq \tau} |\xi_i(s)| \leq N$, $i = 1, 2$, *we have with probability* 1 *for all* $t < \tau$

$$\xi_1(t) = \xi_2(t).$$

Proof. We shall prove only the case when $A(\varphi, t) \in S(\cdot, \lambda_N)$. The second case is proven analogously. Let

$$\sigma = \inf\{t\colon t \geq \tau, |\xi_1(t)| > N, |\xi_2(t)| \geq N, \lambda_N(t) \geq L\},$$

if the set in the braces is nonvoid, and $\sigma = T$ otherwise; also set

$$\alpha_\sigma(\varphi, t) = \alpha(\varphi, t) \quad \text{for } t < \sigma, \qquad \alpha_\sigma(\varphi, t) = 0 \quad \text{for } t \geq \sigma,$$
$$\beta_\sigma(\varphi, t) = \beta(\varphi, t \wedge \sigma).$$

Then for $\|\varphi\| \leq N$ and $\|\psi\| \leq N$

$$\begin{aligned}\left|\int_t^{t+\Delta t} \alpha_{i\sigma}(\varphi, s)\, ds - \int_t^{t+\Delta t} \alpha_{i\sigma}(\psi, s)\, ds\right| &= \left|\int_{t\wedge\sigma}^{(t+\Delta t)\wedge\sigma} [\alpha_i(\varphi, s) - \alpha_i(\psi, s)]\, ds\right| \\ &\leq \|\varphi - \psi\|_* \int_{t\wedge\sigma}^{(t+\Delta t)\wedge\sigma} \lambda_N(s)\, ds \\ &\leq L\|\varphi - \psi\|_* \Delta t,\end{aligned}$$

and analogously applying the theorem on the characteristic of a stopping martingale we obtain

$$\begin{aligned}\mathsf{E}\{|\beta_{i\sigma}(\varphi, \Delta) - \beta_{i\sigma}(\psi, \Delta)|^2 \mid \mathfrak{F}_t\} &\leq \mathsf{E}\left\{\int_{t\wedge\sigma}^{(t+\Delta t)\wedge\sigma} \lambda_N(s)\, ds \|\varphi - \psi\|_* \mid \mathfrak{F}_t\right\} \\ &\leq L\|\varphi - \psi\|_*^2 \Delta t.\end{aligned}$$

Set $\xi(t) = \xi_1(t) - \xi_2(t)$, $\chi(t) = 1$ for $t < \sigma$ and $\chi(t) = 0$ for $t \geq \sigma$. Then

$$\begin{aligned}\mathsf{E}\chi(t)|\xi(t)|^2 &\leq 2\mathsf{E}\chi(t)\left|\int_0^t [\alpha_1(\theta_s\xi_1, s) - \alpha_2(\theta_s\xi_2, s)]\, ds\right|^2 \\ &\quad + 2\mathsf{E}\chi(t)\left|\int_0^t [\beta_1(\theta_s\xi_1, ds) - \beta_2(\theta_s\xi_2, ds)]\right|^2 \\ &= 2\mathsf{E}\chi(t)\left|\int_0^{t\wedge\sigma} [\alpha_{1\sigma}(\theta_s\xi_1, s) - \alpha_{1\sigma}(\theta_s\xi_2, s)]\, ds\right|^2 \\ &\quad + 2\mathsf{E}\chi(t)\left|\int_0^{t\wedge\sigma} \beta_{1\sigma}(\theta_s\xi_1, ds) - \beta_{1\sigma}(\theta_s\xi_2, ds)\right|^2 \\ &\leq 2\mathsf{E}\chi(t)L^2T \int_0^{t\wedge\sigma} \|\theta_s\xi_1 - \theta_s\xi_2\|_*^2\, ds + 2\mathsf{E}L \int_0^{t\wedge\sigma} \|\theta_s\xi_1 - \theta_s\xi_2\|_*^2\, ds \\ &\leq L'\mathsf{E}\int_0^t \chi(s)\|\theta_s\xi\|_*^2\, ds.\end{aligned} \tag{48}$$

Set $z(t) = \sup\{\mathsf{E}\chi(s)|\xi(s)|^2, s \leq t\}$. Since

$$\begin{aligned}\mathsf{E}\chi(s)\|\theta_s\xi\|_*^2 &= \mathsf{E}\chi(s) \int_{-\infty}^0 |\xi(s+u)|^2 K(du) \\ &\leq \mathsf{E}\int_{-\infty}^0 \chi(s+u)|\xi(s+u)|^2 K(du) \leq Kz(s),\end{aligned}$$

it follows from inequality (48) that

$$z(t) \leq L'K \int_0^t z(s)\, ds, \qquad 0 \leq t \leq T.$$

This implies that $z(t) = 0$ for all $t \in [0, T]$ or that $\chi(t)|\xi_1(t) - \xi_2(t)| = 0$ for each t with probability 1. Noting that $\xi_1(t)$ and $\xi_2(t)$ are continuous from the right we deduce that $\chi(t)|\xi_1(t) - \xi_2(t)| = 0$ for all t with probability 1 or that $\xi_1(t) = \xi_2(t)$ for all $t < \sigma$ with probability 1. For L approaching ∞ and taking into account the boundedness of $\lambda_N(t)$ we obtain that $\xi_1(t) = \xi_2(t)$ for all

$$t < \inf\{t: t \geq \tau, |\xi_1(t)| \geq N, |\xi_2(t)| \geq N\}. \quad \square$$

Theorem 5. *Assume that $A(\varphi, t)\in S(\lambda_0, \lambda_N)$ (or that $A(\varphi, t)\in S^c(\lambda_0, \lambda_N)$. The stochastic differential equation* (36) *possesses in H^* (H^c) a unique solution which is defined for all $t\in[0, T]$.*

Proof. As in the proof of Theorem 4 we shall confine ourselves to the case when $A(\varphi, t)\in S(\lambda_0, \lambda_N)$. We first establish the existence of a solution for equation (36).

Let $p>0$. Introduce the functions $\alpha'_p(\varphi, t)$ and $\beta'_p(\varphi, t)$ such that $\alpha'_p(\varphi, t)\in S_\alpha(\lambda_0, \lambda')$, $\beta'_p(\varphi, t)\in S_\beta(\lambda_0, \lambda')$ and $\alpha'_p(\varphi, t)=\alpha(\varphi, t)$, $\beta'_p(\varphi, t)=\beta(\varphi, t)$ for $\|\varphi\|\leqslant r(p)$, where $\lambda'=\lambda'(t)$ does not depend on N and $r(p)$ is a quantity to be defined more precisely below. Set $\tau_p=\inf\{t\colon \lambda_0(t)\geqslant p, \lambda'(t)\geqslant l_p\}$ if the set in braces is nonempty and set $\tau_p=T$ otherwise. Also let

$$\alpha_p(\varphi, t)=\alpha'_p(\varphi, t) \quad \text{for } t<\tau_p, \qquad \alpha_p(\varphi, t)=0 \quad \text{for } t\geqslant\tau_p,$$

$$\beta_p(\varphi, t)=\beta'_p(\varphi, t\wedge\tau_p\} \quad \text{for } t<\tau_p, \qquad \beta_p(\varphi, t)=0 \quad \text{for } t\geqslant\tau_p.$$

(The values of the constants l_p will also be defined more precisely below.) In this case

$$\alpha_p(\varphi, t)=\alpha(\varphi, t), \qquad \beta_p(\varphi, t)=\beta(\varphi, t)$$

for $\|\varphi\|\leqslant r(p)$ and $t<\tau_p$ and $A_p(\varphi, t)=\int_0^t \alpha_p(\varphi, s)\, ds+\beta_p(\varphi, t)\in S(p, l_p)$ (an analogous fact was established in the course of the proof of the preceding theorem). Moreover, $\alpha_p(\varphi, t)$ and $\beta_p(\varphi, t)$ are $\mathfrak{F}_{\tau_p\wedge t}$-measurable functions and $\beta_p(\varphi, t)$ is for a fixed φ a martingale relative to $\{\mathfrak{F}_{\tau_p\wedge t}\}$. It follows from Theorem 3 that the equations

$$d\xi_p=A_p(\theta_t\xi_p, dt), \quad t\in[0, T], \qquad \xi_p(s)=\varphi(s), \quad s<0, \tag{49}$$

possess on the interval $[0, T]$ a solution $\xi_p(t)\in H_2^*$. In view of Theorem 4 we have for $r(p')>r(p)$ on the set $\Omega_p=\{\omega\colon \tau_p=T, \|\theta_T\xi\|\leqslant\tau(p)\}$ equality $\xi_p(t)=\xi_{p'}(t)$ for all t. Moreover, $\mathsf{P}(\bar\Omega_p)\leqslant\mathsf{P}(\tau_p<T)+\mathsf{P}(\|\theta_T\xi\|\geqslant r(p))$. It then follows from Theorem 3 and the Chebyshev inequality that

$$\mathsf{P}(\|\theta_T\xi\|\geqslant r(p))\leqslant\frac{K(p)}{r^2(p)},$$

where $K(m)$ is a function which depends on m, $\|\varphi\|$, and T only (and does not depend on l_p).

Let $K(p)/r^2(p)\to 0$ as $p\to\infty$. Then for p sufficiently large

$$\mathsf{P}(\|\theta_T\xi\|\geqslant r(p))\leqslant\frac{\varepsilon}{2}.$$

We now find values p and l_p such that

$$\mathsf{P}(\tau_p<T)=\mathsf{P}(\{\lambda_0(T)\geqslant p\}\cup\{\lambda'(T)\geqslant l_p\})<\frac{\varepsilon}{2}.$$

Then $\mathsf{P}(\tilde{\Omega}_p)<\varepsilon$. This implies that processes $\xi_p(t)$ converge to a limit $\xi(t)$ with probability 1 and, moreover, $\xi(t)=\xi_p(t)$ with probability 1 for all $t\in[0, T]$ starting with some value of $p=p_0(\omega)$. In particular, the sample functions of the process $\xi(t)$ can be assumed to have left-hand limits and be continuous from the right with probability 1 for all $t\in[0, T]$. Since $\xi_p(t)$ is $\mathfrak{F}_{\tau_p\wedge t}$-measurable, it follows that $\xi(t)$ is measurable with respect to the σ-algebra $\mathfrak{F}_t$. Finally, if $\omega\in\Omega_p$, then $\alpha_p(\varphi, t)=\alpha(\varphi, t)$, $\beta_p(\varphi, t)=\beta(\varphi, t)$, and in view of Theorem 5 we have on the set Ω_p

$$\xi(t)=\varphi(0)+\int_0^t A(\theta_s\xi, s)\,ds$$

with probability 1.

Consequently the last equality is valid with probability 1 on the whole set Ω. The existence of the solution is thus verified. The uniqueness of the solution easily follows from Theorems 3 and 4. □

Assume now that $A(\varphi, t)\in S(\cdot, \lambda_N)$ (or $A(\varphi, t)\in S^c(\cdot, \lambda_N)$). Construct fields $\alpha_p(\varphi, t)$, $\beta_p(\varphi, t)$ which coincide with $\alpha(\varphi, t)$ and $\beta(\varphi, t)$ correspondingly for $\|\varphi\|\leqslant p$, which satisfy the uniform Lipschitz condition with the dominating function $\lambda_p(t)$, and which vanish for $\|\varphi\|\geqslant p+1$. Let $\xi_p(t)$ be a solution of the equations

$$d\xi_p=A_p(\theta_t\xi_p, dt),$$

$$\xi_p(t)=\varphi(t), \qquad t\leqslant 0,$$

and $\tau_p=\inf\{t\colon \|\theta_t\xi_p\|\geqslant p,\ t\in[0, T]\}$ or $\tau_p=T$ if the set of values of t in the braces is void. The sequence of random times τ_p is monotonically nondecreasing. Set $\tau_\infty=\lim \tau_p$. As before $\xi_p(t)=\xi_{p'}(t)$ for $p'>p$ and $t<\tau_p$. Therefore the limit $\lim \xi_p(t)=\xi(t)$ exists with probability 1 for all $t<\tau_\infty$ and coincides with probability 1 with a function $\xi_p(t)$. Thus $\xi(t)$ possesses the left-hand limit and is continuous from the right for all $t<\tau_\infty$ (with probability 1).

Similarly, as in the proof of the preceding theorem, one can verify that $\xi(t)$ satisfies equation (36) for $t<\tau_\infty$.

Theorem 6. *If $A(\varphi, t)\in S(\cdot, \lambda_N)$ (or $(A(\varphi, t)\in S^c(\cdot, \lambda_N))$ then a random time τ_∞ exists on $\{\mathfrak{F}_t, t\in[0, T]\}$ and a random process $\xi(t)$ adopted to $\{\mathfrak{F}_t, t\in[0, T]\}$ defined for $t<\tau_\infty$ such that $\xi(t)$ satisfies* (36) *for all $t<\tau_\infty$, and the sample functions of $\xi(t)$ possess with probability* 1 *left-hand limits and are continuous from the right for all $t<\tau_\infty$. Moreover,* $\mathsf{P}(\tau_\infty>0)=1$.

Theorem 7. *If under the conditions of Theorem* 5 *one can set $\lambda_0=C$, then the solution of equation* (36) *belongs to H_2^* (H_2^c) and*

$$\mathsf{E}\sup_{0\leqslant t\leqslant T}|\xi(t)|^2\leqslant B(1+\|\varphi\|^2), \tag{50}$$

$$\mathsf{E}\sup_{t\leqslant s\leqslant t+h}|\xi(s)-\xi(t)|^2\leqslant B(1+\|\varphi\|^2)h, \tag{51}$$

where B is a constant which depends on C, K, and T only.

Proof. Let τ_p be a random time introduced in the proof of Theorem 5 and let $\xi_p(t)$ be a solution of equation (49). Then $\xi_p(t) \in H_2^*$ and $\mathsf{E}\sup_{0 \leq t \leq T}|\xi_p(t)|^2 < \infty$. It follows from Lemma 8 that

$$(52)\qquad \mathsf{E}\{\sup_{0 \leq h \leq a} |\xi_p(t+h) - \xi_p(t)|^2 \,|\, \mathfrak{F}_t\} \leq C'[a(1+\|\varphi\|^2) + \int_t^{t+a} \sup \mathsf{E}\{|\xi(s)|^2 \,|\, \mathfrak{F}_t\}\, ds],$$

where C' depends on C, K, and T only.

Set $z(t) = \mathsf{E}\{\sup_{0 \leq s \leq t}|\xi_p(s)|^2\}$. It follows from the preceding inequality that

$$z(t) \leq 2|\varphi(0)|^2 + 2C'[T(1+\|\varphi\|^2) + \int_0^t z(s)\, ds]$$

so that in view of Lemma 11

$$z(t) \leq C''(1+\|\varphi\|^2),$$

where C''—like C'—depends on C, K, and T only. Analogously one obtains the inequality

$$(53)\qquad \mathsf{E}\{\sup_{t \leq s \leq t+h} |\xi_p(s)|^2 \,|\, \mathfrak{F}_t\} \leq C''(1+\|\theta_t\xi_p\|^2),$$

which together with (52) yields:

$$(54)\qquad \mathsf{E}\{\sup_{0 \leq h \leq a} |\xi_p(t+h) - \xi_p(t)|^2 \,|\, \mathfrak{F}_t\} \leq C'''(1+\|\varphi\|^2 + \sup_{0 \leq s \leq t} |\xi(s)|^2)a.$$

In particular,

$$(55)\qquad \mathsf{E} \sup_{0 \leq h \leq a} |\xi_p(t+h) - \xi_p(t)|^2 \leq C^{\mathrm{IV}}(1+\|\varphi\|^2)a.$$

Taking into account that $\xi_p(t)$ converges with probability 1 as $p \to \infty$ to a solution $\xi(p)$ of equation (36) and utilizing Fatou's lemma we obtain the required result from inequalities (53) and (55). □

Remark. Analogous results can be obtained for equation (47). In this case one should consider the auxiliary equations

$$\xi_p(t) = \varphi(t) + \int_0^t \alpha_p(\theta_s\xi, s)\, ds + \int_0^t \beta_p(\theta_s\xi, ds).$$

If it is assumed that

$$(56)\qquad \mathsf{E} \sup_{0 \leq s \leq T} |\varphi(s)|^2 = v < \infty,$$

then, in the same manner as in Theorem 7, the following assertion is obtained:

if the assumptions of Theorem 7 and condition (56) *are fulfilled, then the following bound is valid for a solution of equation* (47):

$$\mathsf{E}\sup_{0\le t\le T}|\xi(t)|^2\le B(1+\|\varphi\|^2+v),$$

where B depends on C, K, and T only.

Bounds on the moments of solutions of stochastic differential equations. Consider equation (36) satisfying the conditions of Theorem 5. Assume first that for a fixed φ, $\beta(\varphi, t)\in \mathcal{LM}_2^c$. Furthermore, let the characteristic $\langle\beta^k,\beta^k\rangle_t$ of the process $\beta^k(\varphi, t)$ be absolutely continuous with respect to the Lebesgue measure:

$$\langle\beta^k,\beta^k\rangle_t=\int_0^t\beta^{kk}(\varphi,s)\,ds,\qquad k=1,\dots,m,\tag{57}$$

and

$$|\beta^{kk}(\varphi,s)|\le\lambda^2(s)(1+\|\varphi\|^2).\tag{58}$$

It follows from here that the function $\langle\beta^k,\beta^j\rangle_t$ is also absolutely continuous with respect to the Lebesgue measure. Hence

$$\langle\beta^k,\beta^j\rangle_t=\int_0^t\beta^{kj}(\varphi,s)\,ds,$$

and inequality (51) in Section 1 implies that

$$|\beta^{kj}(\varphi,s)|\le\lambda^2(s)(1+\|\varphi\|^2).$$

Here $\lambda(t)$ is a nonnegative random process bounded on $[0, T]$ and adopted to the current of σ-algebras $\{\mathfrak{F}_t, t\in[0,T]\}$. Also assume that

$$|\alpha(\varphi,s)|\le\lambda(s)(1+\|\varphi\|).\tag{59}$$

We now apply Itô's formula to the function $f(x)=|x|^r$ and to the process $\xi(t)$ satisfying the equation

$$\begin{aligned} d\xi(t)&=\alpha(\theta_t\xi,t)\,dt+\beta(\theta_t\xi,dt), && t\in(0,T],\\ \xi(t)&=\varphi(t), && t\le 0.\end{aligned}\tag{60}$$

Since

$$\nabla f(x)=r|x|^{r-2}x,$$
$$\nabla^2 f(x)=r(r-2)|x|^{r-1}(x\times x)+r|x|^{r-2}E,$$

where E is the unit matrix and $x\times x$ is the matrix with elements $x_{jk}=x^jx^k$

$(j, k = 1, 2, \ldots, m)$, then

$$|\xi(t)|^r = |\varphi(0)|^r + \int_0^t L_c(\xi, s)\, ds + \xi_r(t),$$

where

$$L_c(\xi, s) = r|\xi(s)|^{r-2}\Big[(\xi(s)\alpha(\theta_s\xi, s)) + \sum_{k=1}^{m} \beta^{kk}(\theta_s\xi, s)\Big]$$

$$+ r(r-2)|\xi(s)|^{r-4} \sum_{j,k=1}^{m} \beta^{jk}(\theta_s\xi, s)\xi^j(s)\xi^k(s)$$

and

$$\xi_r(t) = r \int_0^t |\xi(s)|^{r-2}(\xi(s)\beta(\theta_s\xi, ds)).$$

Set

$$\rho(t) = \sup_{0 \le s \le t} |\xi(s)|, \qquad \tau = \inf\{t: \rho(t) > N_1, \lambda(t) > N\},$$

where N_1 and N are constants. Here it is assumed that $\inf \varnothing = T$.

The process $\xi_r(t)$ is the local square integrable martingale with characteristic

$$\langle \xi_r, \xi_r \rangle_t = r^2 \int_0^t |\xi(s)|^{2r-4} \sum_{j,k=1}^{m} \xi^j(s)\xi^k(s)\beta^{jk}(\theta_s\xi, s)\, ds,$$

and in view of Lemma 2 in Section 3 of Chapter I, $\xi_r(t \wedge \tau)$ possesses moments of all orders. Consequently,

$$\mathsf{E}\rho^{2r}(t \wedge \tau) \le 3\Big[|\varphi(0)|^{2r} + \mathsf{E} \sup_{0 \le u \le t} \big(\int_0^{u \wedge \tau} L_c(\xi, s)\, ds\big)^2 + \mathsf{E} \sup_{0 \le u \le t} |\xi_r(u \wedge \tau)|^2\Big].$$

We now bound the terms on the right-hand side of the inequality obtained. We have

$$\sup_{0 \le u \le t} \big(\int_0^{u \wedge \tau} L_c(\xi, s)\, ds\big)^2 \le T \int_0^{t \wedge \tau} c_1\lambda^2(s)[c_2 + \|\varphi\|^{2r} + \rho^{2r}(s)]\, ds,$$

where c_1 and c_2 are constants which depend on r and m only. In the last step we have used the inequality $\|\theta_s\varphi\| \le \|\varphi\| + \rho(s)$, which is self-evident. Furthermore, it follows from the inequality for martingales (Section 1, (4)) that

$$\mathsf{E} \sup_{0 \le u \le t} |\xi_r(u \wedge \tau)|^2 \le 4\mathsf{E}|\xi_r(t \wedge \tau)|^2$$

$$\le 4mr^2\mathsf{E} \int_0^{t \wedge \tau} \lambda^2(s)|\xi(s)|^{2r-2}(1 + \|\theta_s\xi\|^2)\, ds$$

$$\le c_3\mathsf{E} \int_0^{t \wedge \tau} \lambda^2(s)(c_4 + \|\varphi\|^{2r} + \rho^{2r}(s))\, ds,$$

where the constants c_3 and c_4 depend on r and m only. Thus

$$\mathsf{E}\rho^{2r}(t\wedge\tau)\leqslant 3\left[|\varphi(0)|^{2r}+c'N^2(1+\|\varphi\|^{2r})t+c''N^2\int_0^{t\wedge\tau}\mathsf{E}\rho^{2r}(s)\,ds\right].$$

and the constants c' and c'' depend only on r, m, and T. The inequality obtained and Lemma 11 imply that

$$\mathsf{E}\rho^{2r}(t\wedge\tau)\leqslant 3[|\varphi(0)|^{2r}+c'N^2(1+\|\varphi\|^{2r})]\,e^{3c''N^2t}. \tag{61}$$

Here $\tau=\tau(N_1, N)$. Let $N_1\to\infty$. Since the sample functions of process $\xi(t)$ are bounded with probability 1, it is easy to verify that

$$\rho(t\wedge\tau)=\sup_{0\leqslant s\leqslant T\wedge\tau}|\xi(s)|\to\sup_{t\leqslant T\wedge\tau_N}|\xi(s)|,$$

where $\tau_N=\inf\{t\colon \lambda(t)>N\}$. Moreover, in view of Fatou's lemma and inequality (61)

$$\mathsf{E}\rho^{2r}(t\wedge\tau_N)\leqslant\mathsf{E}\lim_{N_1\to\infty}\rho^{2r}(t\wedge\tau)\leqslant C'(\|\varphi\|^{2r},N^2)\,e^{3c''N^2t},$$

where $C'(\|\varphi\|^{2r}, N^2)$ is a constant which depends linearly on $\|\varphi\|^{2r}$, N^2, and T.
We have thus proved the following theorem:

Theorem 8. *A process $\xi(t)$ satisfying equation* (60) *and conditions* (57)–(59) *for $t\leqslant\tau_N=\inf\{t\colon\lambda(t)>N\}$ possesses moments of all orders.*

Corollary. *If*

$$|\alpha(\varphi,t)|\leqslant C(1+\|\varphi\|),$$
$$\mathsf{E}\{|\beta(\varphi,\Delta)|^2\,|\,\mathfrak{F}_t\}\leqslant C(1+\|\varphi\|^2)\,\Delta t,$$

where C is an absolute constant, then the solution of equation (60) *possesses finite moments of all orders for all $t\in[0,T]$.*

Consider now the problem of the existence of moments of solutions for a stochastic equation of the form

$$d\xi=\alpha(\theta_t\xi,t)\,dt+\beta(\theta_t\xi,dt)+\zeta(\theta_t\xi,dt), \tag{62}$$

where $\alpha(\varphi,t)$ and $\beta(\varphi,t)$ satisfy conditions (57)–(59) and

$$\zeta(\varphi,t)=\int_0^t\int_{\mathcal{R}^q}\gamma(\varphi,s,u)\mu(ds,du),$$

where $\mu(\cdot,\cdot)$ is a local martingale measure associated with an integer-valued measure $\nu(t,A)$ on $\mathcal{R}^q$ with characteristic $\pi(t,A)$ which is absolutely continuous with respect to the Lebesgue measure.

Also assume that

(63) $$\int_{\mathcal{R}^q} \gamma^2(\varphi, t, u)\pi(t, du) \leq \lambda^2(t)(1+\|\varphi\|^2)$$

and for any $N>0$, for $\|\varphi_i\| \leq N$, $i=1, 2$,

$$\mathsf{E}\{\int_t^{t+\Delta t}\int |\gamma(\varphi_1, s, u)-\gamma(\varphi_2, s, u)|^2 \pi(ds, du)\,|\,\mathfrak{F}_t\} \leq \|\varphi_1-\varphi_2\|^2 \mathsf{E}\{\int_t^{t+\Delta t} \lambda_N^2(s)\, ds\,|\,\mathfrak{F}_t\},$$

where $\lambda(s)$ and $\lambda_N(s)$ satisfy

(65) $$\int_0^T \lambda^2(s)\, ds < \infty, \qquad \int_0^T \lambda_N^2(s)\, ds < \infty$$

with probability 1. Under these assumptions the conditions of Theorem 5 are easily seen to be fulfilled and equation (62) possesses a unique solution on the interval $t \in [0, T]$.

Let $\tau = \tau(N_1, N) = \inf\{t: |\xi(t)| > N_1, \lambda(t) > N\}$. Then

$$\alpha_\tau(\varphi, t) = \alpha(\varphi, t \wedge \tau), \qquad \beta_\tau(\varphi, t) = \beta(\varphi, t \wedge \tau),$$

$$\zeta_\tau(\varphi, t) = \zeta(\varphi, t \wedge \tau)$$

are linearly bounded by an absolute constant and the process $\xi_\tau(t)$ satisfying the equation

$$d\xi_\tau = \alpha_\tau(\theta_t\xi_\tau, t)\, dt + \beta_\tau(\theta_t\xi_\tau, dt) + \zeta_\tau(\theta_t\xi_\tau, dt), \qquad t>0,$$

$$\xi_\tau(t) = \varphi(t), \qquad t \leq 0,$$

possesses finite second-order moments and, moreover, $\xi_\tau(t) = \xi(t)$ for $t \leq \tau$. In this connection consider first equation (62). Assume that for functions α, β, and ζ the corresponding majorant equals $\lambda(t) = N$ (although in the preceding discussion it could have happened that $\lambda(\tau) > N$, this is of no consequence since $\lambda(t) < N$ for $t < \tau$ and in the succeeding inequalities the value of the function $\lambda(t)$ at a single point is irrelevant).

We now utilize the generalized Itô formula (Chapter I, Section 3, (45)) with $f(x) = |x|^r$. We obtain

(66) $$|\xi(t)|^r = |\varphi(0)|^r + \int_0^t [L_c(\xi, s) + L_d(\xi, s)]\, ds + \xi_r(t) + \eta_r(t),$$

where $L_c(\xi, t)$ and $\xi_r(t)$ are as defined above while

$$L_d(\xi, t) = \int_{\mathcal{R}^q} \{|\xi(t) + \gamma(\theta_t\xi, t, u)|^r - |\xi(t)|^r$$

$$- r(\gamma(\theta_t\xi, t, u), \xi(t))|\xi(t)|^{r-2}\}\pi(t, du),$$

$$\eta_r(t) = \int_0^t \int_{\mathcal{R}^q} (|\xi(s) + \gamma(\theta_s\xi, s, u)|^r - |\xi(s)|^r)\mu(ds, du).$$

In addition to the preceding conditions we shall also assume that

$$\int_{\mathscr{R}^q} |\gamma(\varphi, t, u)|^{2r} \pi(t, du) \leq \lambda^2(t)(1+\|\varphi\|)^{2r}. \tag{67}$$

Taking into account the results in Section 2 of Chapter 1 concerning the finiteness of the moments of stochastic integrals it is easy to verify that the terms on the right-hand side of equation (66) possess finite moments of the second order.

Proceeding as above we obtain the inequality

$$\mathsf{E}\rho^{2r}(t) \leq 4\left(|\varphi(0)|^{2r} + \tilde{C}N^2(1+\|\varphi\|^{2r})t + \tilde{\tilde{C}}N^2 \int_0^t \mathsf{E}\rho^{2r}(s)\, ds\right),$$

where

$$\rho(t) = \sup_{0\leq s\leq t} |\xi(s)|,$$

and the bound

$$\mathsf{E}\rho^{2r}(t) \leq \tilde{C}(\|\varphi\|^{2r})\, e^{\tilde{\tilde{C}}N^2 t}.$$

Thus the following theorem is proved.

Theorem 9. *If for a stochastic equation* (62) *conditions* (57)–(59), (67), *and* (69) *are satisfied, then the solution of this equation possesses for $t<\tau$ finite moments up to the* $(2r)$th *order inclusively. If we also assume that $\lambda(t)=N$, where N does not depend on chance, then,*

$$\mathsf{E} \sup_{0\leq t\leq T} |\xi(t)|^{2r} \leq C_1(1+\|\varphi\|^{2r}), \tag{68}$$

where C_1 is a constant which depends on N, T, and the dimensionality of the space only.

Corollary. *Under the conditions of Theorem 9 and with $\lambda(t)=N$ we have*

$$\mathsf{E} \sup_{0\leq s\leq t} |\xi(s)-\varphi(0)|^{2r} \leq C_2(1+\|\varphi\|^{2r})t, \tag{69}$$

where C_2 is a constant.

Indeed, the preceding arguments yield inequality

$$\mathsf{E}\rho_1^{2r}(t) \leq \tilde{C}N^2(1+\|\varphi\|^2)t + \tilde{\tilde{C}}N^2 \int_0^t \mathsf{E}\rho^{2r}(s)\, ds,$$

where $\rho_1(t)=\sup_{0\leq s\leq t}|\xi(s)-\varphi(0)|^2$. The bound (69) follows from this inequality and (66).

Continuous dependence on a parameter of solutions of stochastic equations. Consider the equation of the form

$$\eta_u(t) = \varphi_u(t) + \int_0^t A_u(\theta_s\eta_u, ds), \qquad t \geq 0, \tag{70}$$
$$\eta_u(s) = \varphi(s), \qquad s < 0,$$

where u is a scalar parameter, $u \in [0, u_0]$, the field

$$A_u(\psi, t) = \int_0^t \alpha_u(\psi, s)\, ds + \beta_u(\psi, t)$$

and the function $\varphi_u(t)$ both depend on the parameter u and the initial condition ($\varphi(s)$ for $s < 0$) does not depend on u.

Theorem 10. *Assume that $A_u(\varphi, t) \in S(C, C)$ and, moreover, let*

a) $\sup\limits_{0 \leq t \leq T} \mathsf{E}|\varphi_u(t)|^2 \leq C,$

b) $\lim\limits_{u \to 0} \sup\limits_{0 \leq t \leq T} \mathsf{E}|\varphi_u(t) - \varphi_0(t)|^2 = 0 \qquad \forall t \in [0, T],$

c) $\mathsf{E}\{|\Delta\beta_u(\psi, t) - \Delta\beta_0(\psi, t)|^2 | \mathfrak{F}_t\} \leq \mathsf{E}\{\int_t^{t+\Delta t} \gamma_u(\psi, s)\, ds \,|\mathfrak{F}_t\}$

and, for arbitrary $N > 0$ and $t \in [0, T]$, let

$$\lim_{u \to 0} \mathsf{P}\{\sup_{\|\psi\| \leq N} |\alpha_u(\psi, t) - \alpha_0(\psi, t)| + \gamma_u(\psi, t)) > \varepsilon\} = 0.$$

Then

$$\lim_{u \to 0} \sup_{0 \leq t \leq T} \mathsf{E}|\eta_u(t) - \eta_0(t)|^2 = 0.$$

Proof. The variables $\xi_u(t)$ possess finite moments of the second order since equations (70) satisfy the conditions of Theorem 7. We represent the difference $\eta_u(t) - \eta_0(t)$ in the form

$$\eta_u(t) - \eta_0(t) = \sigma_u(t) + \int_0^t A_u(\theta_s\eta_u, ds) - \int_0^t A_u(\theta_s\eta_0, ds),$$

where

$$\sigma_u(t) = \varphi_u(t) - \varphi_0(t) + \int_0^t A_u(\theta_s\eta_0, ds) - \int_0^t A_0(\theta_s\eta_0, ds).$$

It is easy to verify that

$$\begin{aligned} \mathsf{E}|\eta_u(t) - \eta_0(t)|^2 &\leq 3\mathsf{E}|\sigma_u(t)|^2 + 3C^2(T+1)\mathsf{E}\int_0^t \|\theta_s(\eta_u - \eta_0)\|_*^2\, ds \\ &= 3\mathsf{E}|\sigma_u(t)|^2 + C' \int_0^t \int_{-s}^0 \mathsf{E}|\eta_u(s+u) + \eta_0(s+u)|^2 K(du)\, ds. \end{aligned}$$

Set $v_u(t) = \sup\limits_{0 \leq s \leq t} \mathsf{E}|\eta_u(s) - \eta_0(s)|^2$. It follows from the last inequality that

$$v_u(t) \leq 3 \sup_{0 \leq s \leq t} \mathsf{E}|\sigma_u(t)|^2 + C'K \int_0^t v_u(s)\, ds.$$

In view of Lemma 11 we have

$$v_u(t) \leqslant C'' \sup_{0 \leqslant s \leqslant t} \mathsf{E}|\sigma_u(s)|^2,$$

where C'' does not depend on u. Furthermore,

$$\begin{aligned}\sup_{0 \leqslant t \leqslant T} \mathsf{E}|\sigma_u(t)|^2 \leqslant 3\Big\{ & \sup_{0 \leqslant t \leqslant T} \mathsf{E}|\varphi_u(t)-\varphi_0(t)|^2 \\ & + \mathsf{E}\big(\textstyle\int_0^T |\alpha_u(\theta_s\eta_0, s)-\alpha_0(\theta_s\eta_0, s)|\, ds\big)^2 \\ & + \sup_{0 \leqslant t \leqslant T} \mathsf{E}\,|\textstyle\int_0^t \beta_u(\theta_s\eta_0, ds)-\int_0^t \beta_0(\theta_s\eta_0, ds)|^2\Big\} \\ = {} & 3(I_1+I_2+I_3).\end{aligned}$$

By the condition, the quantity $I_1 \to 0$. Next

$$I_2 \leqslant T\mathsf{E} \int_0^t |\alpha_u(\theta_s\eta_0, s)-\alpha_0(\theta_s\eta_0, s)|^2\, ds,$$

where the integrand is dominated by the quantity

$$4C^2(1+\textstyle\int_{-\infty}^0 |\eta_0(s+u)|^2 K(du)),$$

which is independent of u and is integrable with respect to the measure $d\mathsf{P}\times ds$. On the other hand,

$$|\alpha_u(\theta_s\eta_0, s)-\alpha_0(\theta_s\eta_0, s)| \to 0$$

in probability for each s and hence in measure $d\mathsf{P}\times ds$. Therefore $I_2 \to 0$ as $u \to 0$. Finally,

$$I_3 \leqslant 4\mathsf{E}\,|\textstyle\int_0^T \beta_u(\theta_s\eta_0, ds)-\int_0^T \beta_0(\theta_s\eta_0, ds)|^2 = 4\mathsf{E}\int_0^T \gamma_u(\theta_s\eta_0, s)\, ds,$$

and as in the case of quantity I_2 it is easy to verify that $I_3 \to 0$ as $u \to 0$. □

Remark. We shall strengthen the assumptions of Theorem 10 and assume that in addition to the conditions stipulated in the theorem

$$\lim_{u\to 0} \mathsf{E} \sup_{0 \leqslant t \leqslant T} |\varphi_u(t)-\varphi_0(t)|^2 = 0.$$

In this case

$$\lim_{u\to 0} \mathsf{E} \sup_{0 \leqslant t \leqslant T} |\eta_u(t)-\eta_0(t)|^2 = 0. \tag{72}$$

The proof of this assertion is analogous to the proof of Theorem 10.

Theorem 11. *Consider the stochastic equation.*

$$d\xi_u = A_u(\theta_s\xi_u, dt), \qquad \xi_u(s) = \varphi(s), \quad s \le 0 \quad (u \in [0, u_0]),$$

satisfying the conditions of Theorem 5 *and for all* $N > 0$ *let*

$$\overline{\lim_{p\to\infty}} \sup_u \left[\mathsf{P}\{\sup_{0\le t\le T} \lambda_0^u(t) \ge p\} + \mathsf{P}\{\sup_{0\le t\le T} \lambda_N^u(t) \ge p\}\right] = 0$$

and also let condition c) *of Theorem* 10 *be satisfied. Then, for any* $\varepsilon > 0$

$$\mathsf{P}\{\sup_{0\le t\le T} |\xi_u(t) - \xi_0(t)| > \varepsilon\} \to 0 \quad \text{as } u \to 0.$$

Proof. Let

$$\tau_p = \inf\{t\colon \lambda_0^u(t) \ge p, \lambda_N^u(t) \ge p, |\xi_u(t)| \ge N\} \qquad (\inf\varnothing = \mathrm{T}),$$

$$\alpha_u^p(\varphi, t) = \alpha_u(\varphi, t) \quad \text{for } t < \tau_p,$$

$$\alpha_u^p(\varphi, t) = 0 \quad \text{for } \tau \ge \tau_p, \qquad \beta_u^p(\varphi, t) = \beta_u(\varphi, t \wedge \tau_p),$$

$$A_u^p(\varphi, t) = \int_0^t \alpha_u^p(\varphi, s)\, ds + \beta_u^p(\varphi, t).$$

Theorem 10 (or the corresponding remarks for this theorem) is applicable to equations

$$d\xi_u^p = A_u^p(\theta_t\xi_u^p, dt), \qquad \xi_u^p(s) = \varphi(s), \quad s \le 0.$$

On the other hand, in view of Theorem 4 $\xi_u^p(t) = \xi_u(t)$ with probability 1 for all $t < \tau_p$. Therefore for any $\varepsilon > 0$

$$\mathsf{P}\{\sup_{0\le t\le T} |\xi_u(t) - \xi_u^p(t)| > \varepsilon\} \le \mathsf{P}\{\tau_p < T\}.$$

Furthermore,

$$\begin{aligned}
\mathsf{P}\{\sup_{0\le t\le T} |\xi_u(t) - \xi_0(t)| > \varepsilon\} &\le \mathsf{P}\left\{\sup_{0\le t\le T} |\xi_0(t) - \xi_0^p(t)| > \frac{\varepsilon}{3}\right\} \\
&\quad + \mathsf{P}\left\{\sup_{0\le t\le T} |\xi_0^p(t) - \xi_u^p(t)| > \frac{\varepsilon}{3}\right\} \\
&\quad + \mathsf{P}\left\{\sup_{0\le t\le T} |\xi_u^p(t) - \xi_u(t)| > \frac{\varepsilon}{3}\right\} \\
&\le 2\mathsf{P}\{\tau_p < T\} + \mathsf{P}\left\{\sup_{0\le t\le T} |\xi_0^p(t) - \xi_u^p(t)| > \frac{\varepsilon}{3}\right\}.
\end{aligned}$$

The uniform stochastic boundedness of the processes $\lambda_0^u(t)$ and $\lambda_N^u(t)$ and Theorem 10 imply that one can first choose a sufficiently large value of p and an N such that $\mathsf{P}\{\tau_p < T\} < \varepsilon/3$ and then choose a $\delta > 0$ such that $\mathsf{P}\{\sup_{0\le t\le T} |\xi_0^p(t) - \xi_u^p(t)| > \varepsilon/3\}$ for $u \in [0, \delta]$. Consequently, for $u < \delta$

$$\mathsf{P}\{\sup_{0\le t\le T} |\xi_u(t) - \xi_0(t)| > \varepsilon\} < \varepsilon. \quad \square$$

Finite-difference approximations of solutions of stochastic equations. Consider the equation

$$(73) \qquad \begin{aligned} d\xi &= A(\theta_t\xi, dt), & t &\in [0, T], \\ \xi(t) &= \varphi(t), & t &\le 0, \end{aligned}$$

where $A(\varphi, t) \in S(\lambda_0, \lambda_N)$. Introduce an arbitrary subdivision $\delta = (0, t_1, t_2, \ldots, t_n = T)$ of the interval $[0, T]$ and define random processes $\xi_\delta(t)$, $\zeta_\delta(t)$, $t \in [0, T]$, using the following recursive relations:

$$\begin{aligned} &\xi_\delta(t) = \zeta_\delta(t) = \varphi(t) && \text{for } t \le 0, \\ &\zeta_\delta(t) = \xi_\delta(t_k) && \text{for } t \in [t_k, t_{k-1}),\ k = 0, \ldots, n-1, \\ &\xi_\delta(t) = \xi_\delta(t_k) + \textstyle\int_{t_k}^t A(\theta_s\zeta_\delta, ds) && \text{for } t \in (t_k, t_{k+1}]. \end{aligned}$$

The process $\xi_\delta(t)$ can be expressed in terms of the process $\zeta_\delta(t)$ in the following manner:

$$\xi_\delta(t) = \varphi(0) + \textstyle\int_0^t A(\theta_s\zeta_\delta, ds) \qquad \forall t \in [0, T].$$

This process $\xi_\delta(t)$ is called a *finite-difference approximation* of a solution of equation (73). We shall now show that $\xi_\delta(t)$ converges to process $\xi(t)$ as $|\delta| \to 0$.

Assume first that $A(\varphi, t) \in S(C_0, C)$. It follows directly from the general properties of the stochastic line integrals and the recurrent relations defining $\xi_\delta(t)$ and $\zeta_\delta(t)$ that the moments of the second order of the variables $\xi_\delta(t)$ and $\zeta_\delta(t)$ are finite. Furthermore, $\mathsf{E} \sup_{0\le t\le T} |\xi_\delta(t)|^2 < \infty$. Set

$$z_\delta(t) = \mathsf{E}\{\sup_{o\le t'\le T} |\xi(t') - \xi_\delta(t')|^2 \,|\, \mathfrak{F}_0\}.$$

Clearly

$$z_\delta(t) \le 2\mathsf{E}\{\sup_{0\le t'\le t} |\textstyle\int_0^{t'} [\alpha(\theta_s\xi, s) - \alpha(\theta_s\zeta_\delta, s)]\, ds|^2$$

$$+ \sup_{0\le t'\le t} |\textstyle\int_0^{t'} \beta(\theta_s\xi, ds) - \beta(\theta_s\zeta_\delta, ds)|^2 \,|\, \mathfrak{F}_0\}.$$

We shall now bound the terms appearing in the right-hand side of the inequality obtained using the methods which were repeatedly utilized above. We obtain

$$z_\delta(t) \leq C^2(8+2T)\mathsf{E}\{\int_0^t \|\theta_s(\xi - \zeta_\delta)\|_*^2 \, ds \mid \mathfrak{F}_0\};$$

this implies that

$$\begin{aligned} z_\delta(t) &\leq C'\mathsf{E}\{\int_0^t \int_{-\infty}^0 (|\xi(s+u) - \xi_\delta(s+u)|^2 \\ &\qquad + |\xi_\delta(s+u) - \zeta_\delta(s+u)|^2) K(du) \, ds \mid \mathfrak{F}_0\} \\ &= C'[z'(t) + z''(t)], \end{aligned}$$

where $C' = C'(C, T)$. Also, inequality

$$z'(t) \leq K \int_0^t z_\delta(s) \, ds$$

can be easily verified. On the other hand,

$$z''(t) \leq K \int_0^t \sup_{0 \leq t' \leq s} \mathsf{E}\{|\xi_\delta(t') - \zeta_\delta(t')|^2 \mid \mathfrak{F}_0\} \, ds.$$

Let

$$w(t) = \mathsf{E}\{|\xi_\delta(t) - \zeta_\delta(t)|^2 \mid \mathfrak{F}_0\}.$$

If $t \in (t_k, t_{k-1}]$, then

$$\begin{aligned} w(t) &= \mathsf{E}\{\mathsf{E}\{|\xi_\delta(t) - \zeta_\delta(t)|^2 \mid \mathfrak{F}_{t_k}\} \mid \mathfrak{F}_0\} \\ &= \mathsf{E}\{|\int_{t_k}^t A(\theta_s \zeta_\delta, ds)|^2 \mid \mathfrak{F}_0\} \\ &\leq C_0 \mathsf{E}\{\int_{t_k}^t (1 + \|\theta_s \zeta_\delta\|^2) \, ds \mid \mathfrak{F}_0\}, \end{aligned}$$

which yields the inequality

$$w(t) \leq C_0 K \int_{t_k}^t (1 + \sup_{-\infty < t' \leq s} \mathsf{E}\{|\xi_\delta(t')|^2 | \mathfrak{F}_0\}) \, ds.$$

We can bound the quantity $\mathsf{E}\{|\xi_\delta(t)|^2 \mid \mathfrak{F}_0\}$ in the same manner as the bound on $\mathsf{E}\{|\xi(t)|^2 \mid \mathfrak{F}_0\}$ was derived (cf. Theorem 7; cf. also Lemma 10, which implies inequality (74)). We have

$$\mathsf{E}\{|\xi_\delta(t)|^2 \mid \mathfrak{F}_0\} \leq C_0'(1 + \|\varphi\|^2), \tag{74}$$

where C_0' depends on C, K, and T only. We thus obtain the bound

$$z''(t) \leq C_0''(1 + \|\varphi\|^2)|\delta|,$$

where $|\delta| = \max \Delta t_k$. Hence

$$z_\delta(t) \leq C'K \int_0^t z_\delta(s)\,ds + C_0''(1+\|\varphi\|^2)|\delta|,$$

which implies that

$$z_\delta(t) \leq C_0''(1+\|\varphi\|^2)\,e^{C''t}|\delta|.$$

The following theorem is thus proved:

Theorem 12. *If $A(\varphi, t) \in S(C_0, C)$, then*

$$\mathsf{E}\{\sup_{0\leq t\leq T} |\xi(t)-\xi_\delta(t)|^2 \,|\, \mathfrak{F}_0\} \leq C_1(1+\|\varphi\|^2)|\delta|, \tag{75}$$

where C_1 depends on C_0, C, K, and T only.

Theorem 13. *If $A(\varphi, t) \in S(\lambda_0, \lambda_N)$, then*

$$\mathsf{P}\{\sup_{0\leq t\leq T} |\xi(t)-\xi_\delta(t)| > \varepsilon \,|\, \mathfrak{F}_0\} \to 0 \quad \text{as } |\delta| \to 0 \ (\text{mod } \mathsf{P}),$$

and the convergence is uniform in the class of all functions $A(\varphi, t)$ with fixed functions $\lambda_0(t)$ and $\lambda_N(t)$.

Proof. Set

$$\tau = \inf\{t\colon \lambda_0(t) > N\} \qquad (\inf \varnothing = T),$$
$$A^\tau(\varphi, t) = A(\varphi, t\wedge\tau),$$

and let $\xi_\delta^\tau(t)$ be a finite-difference approximation of the solution of equation

$$d\xi^\tau = A^\tau(\theta_t\xi^\tau, dt), \qquad t \in [0, T],$$
$$\xi^\tau(t) = \varphi(t), \qquad t \leq 0.$$

Then $\xi_\delta^\tau(t) = \xi_\delta(t)$ for $t < \tau$. Inequality (74) implies

$$\mathsf{P}\{\sup_{0\leq t\leq T} |\xi_\delta(t)| N_1 | \mathfrak{F}_0\} \leq \mathsf{P}\{\tau < T | \mathfrak{F}_0\} + \frac{C(N)(1+\|\varphi\|^2)|\delta|}{N_1^2}.$$

Thus $\mathsf{P}\{\sup_{0\leq t\leq T} |\xi_\delta(t)| > N_1 | \mathfrak{F}_0\} \to 0$ as $N_1 \to \infty$ uniformly in δ (with probability 1).
First choose N_1 such that (for a given ω)

$$\mathsf{P}\{\sup_{0\leq t\leq T} (|\xi_\delta(t) \vee |\xi(t)|) > N_1 \,|\, \mathfrak{F}_0\} < \frac{\varepsilon}{4},$$

for all δ, where ε is an arbitrary positive number. Introduce a new random time (retaining the same designation τ):

$$\tau=\inf\{t\colon (\lambda_0(t)>N)\wedge(|\xi_\delta(t)|>N_1)\wedge(|\xi(t)|>N_1)\wedge(\lambda_{N_1}(t)>N_2)\}.$$

Then $\xi(t)=\xi^\tau(t)$ and $\xi_\delta^\tau(t)=\xi_\delta(t)$ for $t<\tau$.

Consequently if $\|\varphi\|<N_1$, then

$$\mathsf{P}\{\sup_{0\le t\le T}|\xi(t)-\xi_\delta(t)|>\varepsilon\,|\,\mathfrak{F}_0\}\le\mathsf{P}(\tau<T)+\frac{C(N_1,N_2)(1+\|\varphi\|^2)|\delta|}{\varepsilon^2},$$

and, moreover, for N and N_2 sufficiently large

$$\mathsf{P}\{\tau<T\,|\,\mathfrak{F}_0\}\le\mathsf{P}\{\lambda_0(t)>N\,|\,\mathfrak{F}_0\}+\mathsf{P}\{\lambda_{N_1}(t)>N_2\,|\,\mathfrak{F}_0\}+\frac{\varepsilon}{4}<\frac{3\varepsilon}{4}.$$

Thus for $|\delta|<\varepsilon_0$ we have with probability 1

$$\mathsf{P}\{\sup_{0\le t\le T}|\xi(t)-\xi_\delta(t)|>\varepsilon\,|\,\mathfrak{F}_0\}<\varepsilon,$$

where the choice of ε_0 depends only on functions $\lambda_0(t)$, $\lambda_N(t)$, and ε. The theorem is proved. $\square$

Remark. The proof of Theorem 13 yields a somewhat stronger assertion. Actually, we have shown that the relation

$$\mathsf{P}\{\sup_{0\le t\le T}|\xi(t)-\xi_\delta(t)|>\varepsilon\}\to 0$$

is uniformly fulfilled in the class of H functions $A(\varphi,t)$ such that

$$\lim_{C\to\infty}\sup_{A\in H}\mathsf{P}\{\sup_{0\le t\le T}\lambda_0(t)>C\}=0,$$

$$\lim_{C\to\infty}\sup_{A\in H}\mathsf{P}\{\sup_{0\le t\le T}\lambda_N(t)>C\}=0\qquad\forall N>0.$$

§ 2. Stochastic Differential Equations without an After-Effect

Solutions of stochastic differential equations without an after-effect as a Markov process. An equation of the form (36) in Section 1 is called *a stochastic differential equation without an after-effect* provided $A(\varphi,t+h)-A(\varphi,t)$ does not depend on the σ-algebra $\mathfrak{F}_t$ and on the values of $\varphi(s)$ for $s<0$. Thus one can set $A(\varphi,t)=A(x,t)$ where $x=\varphi(0)$ and the process $A(x,t)$ for a fixed x is a process with independent increments.

Assume that $A(x, t)$ possesses finite moments of the second order and let

$$A(x, t) = \alpha(x, t) + \beta(x, t)$$

where $\beta(x, t)$ is a square integrable martingale with independent increments and $\alpha(x, t)$ is a nonrandom vector-valued function. In our case the condition $\alpha \in S_\alpha(\lambda_0, \lambda_N)$ implies that the function $\alpha(x, t)$ is a Borel function in arguments (x, t) and is differentiable with respect to t for almost all t. It is natural to assume that the function $\alpha_t'(x, t) \underset{\text{Def}}{=} a(x, t)$ exists for all t and to replace conditions (37) and (38) appearing in Section 1 by the following:

$$(1) \qquad |a(x, t)| \leq K(1+|x|) \qquad \forall x \in \mathscr{R}^m,$$

$$(2) \qquad |a(x, t) - a(y, t)| \leq C_N |x - y| \qquad \forall (x, y), |x| \leq N, |y| \leq N,$$

where K and C_N are constants. Thus in the case under consideration there is no point in distinguishing between classes $S_\alpha(\lambda_0, \lambda_N)$ and $S_\alpha(K, C_N)$. An analogous situation exists also in the case of the condition $\beta(x, t) \in S_\beta(\lambda_0, \lambda_N)$. We shall replace this condition by the following:

a) The function $\beta(x, t)$ is with probability 1 a Borel function in arguments (x, t) on each interval $t \in [0, s]$, is $\mathfrak{F}_t$-measurable as a function ω, and the sample functions belong for a fixed x with probability 1 to $\mathscr{D}^m[0, T]$.

b) $\mathsf{E}|\beta(x, \Delta)|^2 \leq K(1+|x|^2)\Delta t$ for each $x \in \mathscr{R}^m$, where $\Delta = (t, t+\Delta t]$.

c) For an arbitrary N there exists a constant C_N such that

$$\mathsf{E}|\beta(x, \Delta) - \beta(y, \Delta)|^2 \leq C_N |x - y|^2 \Delta t, \qquad \forall (x, y), |x| \leq N, |y| \leq N,$$

and, as above, the classes $S_\beta(\lambda_0, \lambda_N)$ and $S_\beta(K, C_N)$ coincide.

If $\beta(x, t) \in S_\beta^c(\lambda_0, \lambda_N)$, then $\beta(x, t)$ satisfies conditions a)–c) and, moreover, the sample functions of $\beta(x, t)$ are continuous functions for a fixed x. Thus $\beta(x, t)$ is, in the case under consideration, a Gaussian process with independent increments (for a fixed x).

We shall agree to write $A(x, t) \in \bar{S}(K, C_N)$ provided $A(x, t)$ is a process with independent increments,

$$A(x, t) = \int_0^t a(x, s)\, ds + \beta(x, t),$$

$a(x, t)$ satisfies conditions (1), (2), and $\beta(x, t)$ is a square integrable martingale (on $[0, T]$) with independent increments satisfying conditions a)–c) stipulated above. We shall use the notation $A(x, t) \in \bar{S}^c(K, C_N)$ if $\beta(x, t)$ is, in addition, a Gaussian process for a fixed x.

Set $\tilde{B}(x, t) = \mathsf{E}\beta(x, t)\beta^*(x, t)$. The function $\tilde{B}(x, t)$ is the matrix characteristic of the field $\beta(x, t)$. Since

$$\tilde{B}(x, \Delta) = \tilde{B}(x, t+\Delta t) - \tilde{B}(x, t) = \mathsf{E}\beta(x, \Delta)\beta^*(x, \Delta),$$

it is easy to verify that condition b) is equivalent to the requirement that function $\tilde{B}(x, t)$ be absolutely continuous in t,

$$(3) \qquad \tilde{B}(x, t) = \int_0^t \tilde{b}(x, s)\, ds,$$

and its derivative $\tilde{b}(x, t)$ satisfies inequality

$$|\tilde{b}(x, t)| \leq K(1 + |x|^2).$$

We introduce the joint characteristic $\tilde{B}(x, y, t)$ of the processes $\beta(x, t)$ and $\beta(y, t)$, i.e., $\tilde{B}(x, y, t) = \mathsf{E}\beta(x, t)\beta^*(y, t)$. It follows from equation (3) that

$$\tilde{B}(x, y, t) = \int_0^t \tilde{b}(x, y, s)\, ds;$$

moreover, $\tilde{b}(x, x, t) = \tilde{b}(x, t)$ and $\tilde{b}(x, y, t) = \tilde{b}(y, x, t)$. Condition c) is equivalent to the following:

$$(4) \quad |\tilde{b}(x, x, t) - 2\tilde{b}(x, y, t) + \tilde{b}(y, y, t)| \leq C_N |x - y|^2 \qquad \forall (x, y), |x| \leq N, |y| \leq N.$$

We now state a number of previously obtained results concerning stochastic differential equations suitably adopted to the case under consideration.

Consider the stochastic differential equation

$$(5) \qquad d\xi(t) = A(\xi(t), dt) = a(\xi(t), t)\, dt + \beta(\xi(t), dt), \qquad t \geq s,\ \xi(s) = x,$$

where $a(x, t)$ is a nonrandom function with values in $\mathcal{R}^m$, $(x, t) \in \mathcal{R}^m \times [0, T]$, $\beta(x, t)$ is a family of processes with independent increments taking on values in $\mathcal{R}^m$ and possessing finite moments of the second order.

Theorem 1. *Assume that $A(x, t) \in \bar{S}(\cdot, C_N)$ and the matrix function $\tilde{B}(x, y, t)$ is differentiable with respect to t.*

Then:

a) *A random time τ and a process $\xi(t)$ defined for $s \leq t < \tau$ exist such that* $\mathsf{P}(\tau > s) = 1$, *the process $\xi(t)$ satisfies equation* (5) *for $s \leq t < \tau$, and the sample functions of this process possess left-hand limits and are continuous on the right for all t, $s \leq t < \tau$. If $\xi'(t)$ is another solution of* (5) *with sample trajectories possessing the same property and defined for $t < \tau'$, then*

$$\mathsf{P}\{\exists t: \xi(t) \neq \xi'(t), s \leq t \leq \tau \wedge \tau'\} = 0.$$

b) *If $A(x, t) \in \bar{S}(K, C_N)$, equation* (5) *possesses a solution defined for $t \in [s, T]$ possessing finite moments of the second order with sample functions belonging to* $\mathcal{D}^m[s, T]$ (mod P).

c) *If $A(x, t) \in \bar{S}^c(K, C_N)$, equation* (5) *possesses a solution on the interval $[s, T]$ with sample functions belonging to* $\mathcal{C}^m$ (mod P). *This solution admits moments of all orders.*

Consider equation (5) and assume that for each $s \in [0, T]$ it possesses a unique solution on the interval $[s, T]$ satisfying the initial condition $\xi(s) = x$ with sample functions belonging to $\mathcal{D}^m[s, T]$. We denote this solution by $\xi_{xs}(t)$.

Let $\mathfrak{F}_t^s$ denote a completion of a σ-algebra generated by the random vectors $\beta(x, u) - \beta(x, s)$, $x \in \mathcal{R}^m$, $u \in (s, t]$, and let $\mathfrak{F}_t = \mathfrak{F}_t^0$. Clearly σ-algebras $\mathfrak{F}_{t_2}^{t_1}$ and $\mathfrak{F}_{t_3}^{t_2}$ are independent for $t_1 < t_2 < t_3$ and the variables $\xi_{xs}(t)$ are $\mathfrak{F}_t^s$-measurable.

We shall now derive several bounds which will be utilized below. These bounds are valid not only for equations without an after-effect and we shall therefore prove them for a more general case.

Denote by $\tilde{S}(\lambda_0, \lambda_N)$ ($\tilde{S}^c(\lambda_0, \lambda_N)$) the subclass of the class $S(\lambda_0, \lambda_N)$ ($S^c(\lambda_0, \lambda_N)$) consisting of random functions of the form

$$A(x, t) = \int_0^t \alpha(x, s)\, ds + \beta(x, t),$$

where $(x, t) \in \mathcal{R}^m \times [0, T]$. Equation (5) will be considered also in the case when $A(x, t) \in \tilde{S}(\lambda_0, \lambda_N)$. Results of Section 1, in particular the theorems concerning the existence and uniqueness of the solutions, are fully applicable in this case as well.

Lemma 1. *Let $A(x, t) \in \tilde{S}(K, C)$. Then for $0 \leq s \leq t \leq T$*

$$\mathsf{E}\{|\xi_{xs}(t) - \xi_{ys}(t)|^2 \mid \mathfrak{F}_s^0\} \leq \bar{C}|x - y|^2,$$

where the constant $\bar{C}$ depends on C and T only.

Proof. Since

$$\xi_{xs}(t) - \xi_{ys}(t) = x - y + \int_s^t [\alpha(\xi_{xs}(u), u) - \alpha(\xi_{ys}(u), u)]\, du$$
$$+ \int_s^t \beta(\xi_{xs}(u), du) - \beta(\xi_{ys}(u), du),$$

in view of Lemma 10 in Section 1 the function

$$v(t) = \mathsf{E}\{|\xi_{xs}(t) - \xi_{ys}(t)|^2 \mid \mathfrak{F}_s^0\}$$

satisfies the inequality

$$v(t) \leq 3\left(|x - y|^2 + C^2(T + 1) \int_s^t v(u)\, du\right).$$

It follows from Lemma 11 in Section 1 that $v(t) \leq \bar{C}|x - y|^2$ where $\bar{C}$ is a constant which depends on C and T only. $\square$

Lemma 2. *Assume that $A(x, t) \in \tilde{S}(K, C)$. Then*

$$\mathsf{E}|\xi_{x_1 s_1}(t) - \xi_{x_2 s_2}(t)|^2 \leq C'(|x_1 - x_2|^2 + (1 + |x_2|^2)(s_2 - s_1)), \tag{6}$$

where C' is a constant which depends on K, C, and T only.

Proof. We have

$$\mathsf{E}|\xi_{x_1s_1}(t)-\xi_{x_2s_2}(t)|^2 \leq 2\mathsf{E}|\xi_{x_1s_1}(t)-\xi_{x_2s_1}(t)|^2+2\mathsf{E}|\xi_{x_2s_1}(t)-\xi_{x_2s_2}(t)|^2.$$

Furthermore,

$$\begin{aligned}\mathsf{E}|\xi_{x_2s_1}(t)-\xi_{x_2s_2}(t)|^2 &= \mathsf{E}\{\mathsf{E}\{|\xi_{x_2s_2}(t)-\xi_{\xi_{x_2s_1}(s_2)s_2}(t)|^2\,|\,\mathfrak{F}^0_{s_2}\}\} \\ &= \mathsf{E}\{(\mathsf{E}|\xi_{x_2s_2}(t)-\xi_{ys_2}(t)|^2)_{y=\xi_{x_2s_1}(s_2)}|\,\mathfrak{F}^0_{s_2}\};\end{aligned}$$

the last expression, in view of Lemma 1, is bounded by the quantity

$$A\mathsf{E}|x_2-\xi_{x_2s_1}(s_2)|^2.$$

In turn, it follows from Theorem 3 in Section 1 that

$$\mathsf{E}|x_2-\xi_{x_2s_1}(s_2)|^2 \leq B(1+|x_2|^2)(s_2-s_1).$$

Utilizing Lemma 1 once again to bound the quantity $\mathsf{E}|\xi_{x_1s_1}(t)-\xi_{x_2s_1}(t)|^2$ we obtain inequality (6). □

Corollary. *If $f(x)$, $x\in\mathcal{R}^m$, is a bounded and continuous function and the conditions of the preceding lemma are satisfied, then the function*

$$v(t,x)=\mathsf{E}f(\xi_{xt}(T))$$

is bounded and is jointly continuous in the variables (x,t). Moreover, if $f(x)$ is continuous and $|f(x)|\leq C(1+|x|^\rho)$ and $\mathsf{E}|\xi_{xt}(T)|^p$ is uniformly bounded on an arbitrary compact set of values (x,t), with $p>\rho$, then the function $v(t,x)$ is also continuous in (x,t).

Indeed if the function $f(x)$ is continuous, then in view of the corollary to Lemma 2, $f(\xi_{xt}(T))$ is a function continuous in probability in (x,t). The stipulated assumptions assure the possibility of a limit transition under the sign of the mathematical expectation.

The following formula will be utilized below.

Let $f(x,\omega)$ be a bounded $\mathfrak{B}\times\mathfrak{Y}$-measurable function, $(x,\omega)\in\mathcal{X}\times\Omega$, $\mathfrak{B}$ be a σ-algebra of Borel sets in the metric space $\mathcal{X}$, $(\Omega,\mathfrak{Y},\mathsf{P})$ be a probability space. Assume that $\zeta=\zeta(\omega)$ is an $\mathfrak{F}$-measurable mapping of Ω into $\mathcal{X}$, where $\mathfrak{F}\subset\mathfrak{Y}$. Then

$$\mathsf{E}\{f(\zeta,\omega)\,|\,\mathfrak{F}\}=g(\zeta),\quad \text{where } g(x)=\mathsf{E}\{f(x,\omega)\,|\,\mathfrak{F}\}.$$

To prove this assertion we introduce the class K of functions $f(x,\omega)$ for which the stated formula is valid. Clearly this class is linear and monotone (i.e., given an arbitrary monotonically nondecreasing sequence of nonnegative functions converging to a finite limit belonging to K, the limit function also belongs to K). Furthermore, K contains functions of the form $f(x,\omega)=\sum_{k=1}^n h_k(x)l_k(\omega)$, where $h_k(x)$ are bounded $\mathfrak{B}$-measurable and $l_k(\omega)$ are bounded $\mathfrak{Y}$-measurable functions.

Indeed,

$$\mathsf{E}\left\{\sum_{k=1}^{n} h_k(\zeta)l_k(\omega)\,|\,\mathfrak{F}\right\}=\sum_{k=1}^{n} h_k(\zeta)\alpha_k(\omega),$$

where $\alpha_k(\omega)=\mathsf{E}\{l_k(\omega)\,|\,\mathfrak{F}\}$ and

$$\mathsf{E}\left\{\sum_{k=1}^{n} h_k(x)l_k(\omega)\,|\,\mathfrak{F}\right\}=\sum_{k=1}^{n} h_k(x)\alpha_k(\omega).$$

It follows from the above stated properties of class K that it contains arbitrary Borel functions measurable with respect to the minimal σ-algebra which contains all sets of the form $B\times A$ where $B\in\mathfrak{B}$, $A\in\mathfrak{Y}$, i.e. which are $\mathfrak{B}\times\mathfrak{Y}$-measurable.

We now return to the process $\xi_{xs}(t)$. Set

$$P(s, x, t, A)=\mathsf{P}(\xi_{xs}(t)\in A)$$

where A is an arbitrary Borel set in $\mathcal{R}^m$. The function $P(s, x, t, A)$ is a stochastic kernel. The equality

$$\mathsf{E}f(\xi_{xs}(t))=\int_{\mathcal{R}^m} f(y)P(s, x, t, dy)$$

is valid for an arbitrary nonnegative Borel function $f(x)$. (This follows from the general rule of change of variables in integral calculus.)

Theorem 2. *The family of stochastic kernels $P(s, x, t, A)$, $0\leqslant s<t\leqslant T$, is a Markov family.*

Proof. To prove the theorem it is required to verify that the kernels $P(s, x, t, A)$ satisfy the Chapman–Kolmogorov equation (Volume II, Chapter I, Section 1). Let $s<u<t$. Then

$$\begin{aligned}\mathsf{E}f(\xi_{xs}(t))&=\mathsf{E}(\mathsf{E}\{f(\xi_{xs}(t))\,|\,\mathfrak{F}_u^s\})\\&=\mathsf{E}\mathsf{E}\{f(\xi_{\xi_{xs}(u)u}(t))\,|\,\mathfrak{F}_u^s\}\\&=\mathsf{E}(\mathsf{E}f(\xi_{yu}(t)))|_{y=\xi_{xs}(u)}\end{aligned}$$

or

$$\int_{\mathcal{R}^m} f(z)P(s, x, t, dz)=\int_{\mathcal{R}^m} P(s, x, u, dy)\int_{\mathcal{R}^m} f(z)P(u, y, t, dz).$$

This implies that for an arbitrary Borel set A

$$P(s, x, t, A)=\int_{\mathcal{R}^m} P(u, y, t, A)P(s, x, u, dy),$$

i.e., the kernels $P(s, x, t, A)$ indeed satisfy the Chapman–Kolmogorov equation. $\square$

The proof of the theorem shows that the kernels $P(s, x, t, A)$ are transition probabilities of a certain Markov process (Volume II, Chapter I, Section 3).

We say that the stochastic differential equation under consideration generates a Markov process with transition probabilities $\mathsf{P}(\xi_{xs}(t) \in A)$. In this chapter we shall often identify this Markov process with the family of random processes $\xi_{xs}(t)$.

We now proceed to evaluate the generating operator of a Markov process $\xi_{xs}(t)$ generated by a stochastic differential equation without an after-effect.

Set

$$\xi'_{xs}(t) = x + \int_s^t a(x, u)\, du + \beta(x, t).$$

Lemma 3. *If* $A(x, t) \in \bar{S}(K, C)$, *then*

$$|\mathsf{E}[\xi_{xs}(t) - \xi'_{xs}(t)]| \leq C''(1 + |x|)(t - s)^{3/2}, \tag{7}$$

$$\mathsf{E}|\xi_{xs}(t) - \xi'_{xs}(t)|^2 \leq C'(1 + |x|^2)(t - s)^2. \tag{8}$$

Proof. Denote

$$v(t) = |\mathsf{E}[\xi_{xs}(t) - \xi'_{xs}(t)]|, \qquad z(t) = \mathsf{E}|\xi_{xs}(t) - \xi'_{xs}(t)|^2.$$

Then

$$v(t) = |\mathsf{E} \int_s^t [a(\xi_{xs}(u), u) - a(x, u)]\, du| \leq C \int_s^t \mathsf{E}|\xi_{xs}(u) - x|\, du.$$

Taking the corollary to Theorem 9 in Section 1 into account we obtain the inequality

$$v(t) \leq CC^{(t)\frac{1}{2}} \int_s^t \sqrt{u - s}\, du = C''(t - s)^{3/2}.$$

Furthermore,

$$z(t) \leq 2\,(\mathsf{E}|\int_s^t [a(\xi_{xs}(u), u) - a(x, u)]\, du|^2 + \mathsf{E}|\int_s^t [\beta(\xi_{xs}(u), du) - \beta(x, du)]|^2).$$

Using Lemma 9 in Section 1 we arrive at

$$z(t) \leq 2(TC^2 + C') \int_s^t \mathsf{E}|\xi_{xs}(u) - x|^2\, du,$$

which, together with the bound (69) in Section 1, yields inequality (8). $\square$

Let $f(x)$, $x \in \mathcal{R}^m$, be an arbitrary thrice continuously differentiable function with bounded partial derivatives of the first, second, and third orders. We show that relation

$$z(s, t) = \frac{1}{t - s} \mathsf{E}[f(\xi_{xs}(t)) - f(\xi'_{xs}(t))]$$

tends to zero uniformly on each compact set of the form $0 \leq s \leq t \leq T$, $|x| \leq N$, $N > 0$. Indeed, using Taylor's formula it is easy to establish the inequality of the form

$$(t-s)z(s, t) \leq K_1(|\mathsf{E}[\xi_{xs}(t) - \xi'_{xs}(t)]| \\ + \mathsf{E}|\xi_{xs}(t) - \xi'_{xs}(t)|\,|\xi'_{xs}(t) - x| + \mathsf{E}|\xi_{xs}(t) - \xi'_{xs}(t)|^2),$$

where the constant K_1 depends on the values of K, C only and the upper bounds on the derivatives of the first and second orders of the function $f(x)$. Clearly $\mathsf{E}|\xi'_{xs}(t) - x|^2 \leq C'(1 + |x|^2)(t-s)$. It follows from Lemma 3 that

$$(t-s)z(s, t) \leq K'(1 + |x|^2)(t-s)^{3/2}. \tag{9}$$

We now utilize the generalized Itô formula. Set

$$\beta(x, t) = \beta_c(x, t) + \zeta(x, t),$$

where $\beta_c(x, t)$ is a continuous component of the random function $\beta(x, t)$ and $\zeta(x, t)$ is its discontinuous martingale part, and let $\nu(x, t, A)$ be an integral-value measure constructed from the jumps of the process $\beta(x, t)$, $\mu(x, t, A)$ be the martingale measure associated with it, and $\pi(x, t, A)$ be its characteristic. Then

$$\zeta(x, t) = \int_{\mathcal{R}^m} u\mu(x, t, du).$$

Denote by $B(x, t)$ the matrix characteristic of the process $\beta_c(x, t)$. It follows from the orthogonality of $\beta_c(x, t)$ and $\zeta(x, t)$ that

$$\tilde{B}(x, t) = B(x, t) + \int_{\mathcal{R}_m} uu^* \pi(x, t, du).$$

Clearly the measure $\pi(x, t, A)$ is nonrandom. The condition $\beta(x, t) \in S_\beta(K, C_N)$ implies that $B(x, t)$ and the matrix function $\int_{\mathcal{R}^m} uu^*\pi(x, t, du)$ are absolutely continuous with respect to the Lebesgue measure. Set

$$B(x, t) = \int_0^t b(x, s)\, ds, \qquad \pi(x, t, A) = \int_0^t \Pi(x, s, A)\, ds,$$

where $\Pi(x, t, A)$ is a nonrandom function which is a measure on $\mathfrak{B}^m$ for a fixed (x, t). Moreover

$$\int_{\mathcal{R}^m} |u|^2 \Pi(x, t, du) < \infty, \qquad \forall (x, t) \in \mathcal{R}^m \times [0, T].$$

It follows from the generalized Itô formula (Chapter 1, Section 3, equation (45)) that:

$$f(\xi'_{xs}(t)) = f(x) + \int_s^t (L_c f(\xi'_{xs}(\theta)) + L_d f(\xi'_{xs}(\theta)))\, d\theta + \int_s^t (\nabla f(\xi'_{xs}(\theta)), \beta_c(x, d\theta)) \\ + \int_s^t \int_{\mathcal{R}^m} [f(\xi'_{xs}(\theta) + u) - f(\xi'_{xs}(\theta))]\mu(x, d\theta, du);$$

(observe that all the conditions for the applicability of this formula are fulfilled). Here

$$L_c f(\xi'_{xs}(\theta)) = (\nabla f(\xi'_{xs}(\theta)), a(x, \theta)) + \frac{1}{2} \sum_{k,j=1}^{m} \nabla^k \nabla^j f(\xi'_{xs}(\theta)) b^{kj}(x, \theta),$$
$$L_d f(\xi'_{xs}(\theta)) = \int_{\mathcal{R}^m} [f(\xi'_{xs}(\theta)+u) - f(\xi'_{xs}(\theta)) - (\nabla f(\xi'_{xs}(\theta)), u)] \Pi(x, \theta, du),$$

where $b^{kj}(x, t)$ are the entries of the matrix $b(x, t)$. From the assumptions on function $f(x)$ and the preceding bounds it is easy to obtain that for $t' \downarrow t$ and $s \uparrow t$

$$\lim_{\substack{t' \downarrow t \\ s \uparrow t}} \frac{\mathsf{E} f(\xi'_{xs}(t')) - f(x)}{t' - s} = (L_c + L_d) f(x)$$

uniformly in $(x, t) \in S_N \times [0, T]$ for any $N > 0$.

Finally, since

$$\lim_{\substack{t' \downarrow t \\ s \uparrow t}} \frac{\mathsf{E} f(\xi_{xs}(t')) - f(x)}{t' - s} = \lim_{\substack{t' \downarrow t \\ s \uparrow t}} \frac{z(s, t')}{t' - s} + \lim_{\substack{t' \downarrow t \\ s \uparrow t}} \frac{\mathsf{E} f(\xi_{xs}(t')) - f(x)}{t' - s},$$

we have

$$\lim_{\substack{t' \downarrow t \\ s \uparrow t}} \frac{\mathsf{E}(\xi_{xs}(t')) - f(x)}{t' - s} = (L_c + L_d) f(x)$$

(10)
$$= (\nabla f(x), a(x, t)) + \frac{1}{2} \sum_{k,j=1}^{m} \nabla^k \nabla^j f(x) b^{kj}(x, t)$$
$$+ \int_{\mathcal{R}^m} [f(x+u) - f(x) - (\nabla f(x), u)] \Pi(x, t, du).$$

The assumptions under which formula (10) was established can be somewhat weakened. Firstly it is sufficient to require only that $A(x, t) \in S(K, C_N)$.

Indeed, construct functions $a_N(x, t)$ and $\beta_N(x, t)$ such that they are linearly bounded, satisfy the uniform Lipschitz condition, and coincide with $a(x, t)$ and $\beta(x, t)$ on the sphere $S_N(x)$ of radius N with the center at point x. Let $\xi_N(t)$ be a solution of equation

$$d\xi_N(t) = A_N(\xi_N(t), dt), \qquad A_N(x, t) = \int_0^t a_N(x, s)\, ds + \beta_N(x, t).$$

Denote by τ_N the time of the first exit of function $\xi(t)$ from the sphere $S_N(x)$. Then $\xi_{xs}(t) = \xi_{N_{xs}}(t)$ for $t < \tau_N$. For an arbitrary bounded function $f(x)$ we have

$$\frac{1}{t' - s} |\mathsf{E}[f(\xi_{xs}(t')) - f(\xi_{N_{xs}}(t))]| \leq \frac{\mathsf{P}(\sup_{s \leq t \leq t'} |\xi_{N_{xs}}(t) - x| > N)}{t' - s}$$
$$\leq \frac{\mathsf{E} \sup_{s \leq t \leq t'} |\xi_{N_{xs}}(t) - x|^2}{(t' - s) N^2} \leq \frac{C}{N^2}.$$

Therefore applying relation (10) to process $\xi_{N_{xs}}(t)$ and then approaching the limit as $N\to\infty$, one verifies that this relation is preserved also for the classes of equations under consideration.

Analogously one can generalize equality (10) to arbitrary twice continuously differentiable and bounded functions $f(x)$ possessing bounded partial derivatives of the second order.

To show this, one constructs a sequence of functions $f_N(x)$ bounded and thrice continuously differentiable, with bounded partial derivatives up to the third order inclusive and such that these functions and their partial derivatives of the first and second order differ on the sphere $S_N(x)$ by at most $1/N$ from $f(x)$ and the corresponding partial derivatives of the function $f(x)$. Then

$$\frac{1}{t'-s}|\mathsf{E}f(\xi_{xs}(t'))-f(x)-[\mathsf{E}f_N(\xi_{xs}(t'))-f(x)]|$$

$$\leq\frac{1}{t'-s}|\mathsf{E}(f-f_N)(\xi_{xs}(t'))-(f-f_N)(x)|$$

$$\leq\frac{1}{t'-s}\left[\frac{C'}{N}\mathsf{E}|\xi_{xs}(t')-x|+C'\mathsf{P}(\tau_N<t)\right],$$

where C' is a constant independent of N and τ_N denotes as before the first exit time from the sphere $S_N(x)$. Since $\mathsf{P}(\tau_N<t)=N^{-2}\mathsf{E}\sup_{s\leq t\leq t'}|\xi_{xs}(t')-x|^2$, the quantity under consideration does not exceed

$$\frac{1}{t'-s}\frac{C''}{N^2}\mathsf{E}\sup_{s\leq t\leq t'}|\xi_{xs}(t')-x|^2\leq\frac{C^*}{N^2}.$$

It is also easy to verify that $L_c f-L_c f_N\to 0$ and $L_d f-L_d f_N\to 0$. Note that the boundedness of the second-order partial derivatives of function f is utilized only in the proof of relation $L_d f-L_d f_N\to 0$.

Theorem 3. *Equality* (10) *is valid for an arbitrary bounded and twice continuously differentiable function $f(x)$ with bounded partial derivatives of the second order and for a solution $\xi_{xs}(t)$ of the equation*

$$d\xi_{xs}(t)=A(\xi_{xs}(t),dt),\qquad \xi_{xs}(s)=x,$$

where $A(x,t)\in S(K,C_N)$.

If $A(x,t)$ is a continuous process, then

$$\lim_{\substack{t'\downarrow t\\ s\uparrow t}}\frac{1}{t'-s}[\mathsf{E}f(\xi_{xs}(t'))-f(x)]=(\nabla f(x),a(x,t))+\frac{1}{2}\sum_{k,j=1}^{m}b^{kj}(x,t)\frac{\partial^2 f(x)}{\partial x^k\,\partial x^j}$$

for an arbitrary twice continuously differentiable function which increases as $|x|\to\infty$ not faster than a power of $|x|$.

Proof. Only the second assertion requires a proof. Now let the function $f_N(x)$ coincide with the function $f(x)$ on $S_N(x)$ and be bounded together with its partial derivatives of the second order in $\mathcal{R}^m$. We apply relation (10) to this function. Note that

$$\begin{aligned}&\frac{1}{t'-s}|\mathsf{E}[f(\xi_{xs}(t'))-f(x)-(f_N(\xi_{xs}(t'))-f_N(x))]|\\ &\leqslant\frac{1}{t'-s}C\mathsf{E}\chi(\tau_N<t)(1+|\xi_{xs}(t')|^p)\\ &\leqslant\frac{C}{(t'-s)N^2}\mathsf{E}|\xi_{xs}(t')-x|^2(1+|\xi_{xs}(t')|^p).\end{aligned}$$

Using Itô's formula it is easy to verify, in the same manner as was done when the bounds on the moments of a solution of a stochastic differential equation were estimated, that for any $p>0$

$$\mathsf{E}|\xi_{xs}(t)-x|^2(1+|\xi_{xs}(t)|^p)\leqslant C(t-s),$$

where C is a constant. Thus

$$\frac{1}{t'-s}|\mathsf{E}f(\xi_{xs}(t'))-f_N(\xi_{xs}(t'))|\to 0$$

as $N\to\infty$. It is now obvious how to complete the proof of the theorem. □

Remark 1. In the case of a general equation belonging to the class $S(K, C_N)$ relation (10) can also be generalized to growing functions. One need only require the existence of moments of the process $\xi_{xs}(t)$ of a sufficiently high order.

Remark 2. Let function $f(t, x)$ and its partial derivatives with respect to x of the first and second order be uniformly bounded and continuous jointly in the variables (t, x). Then

$$\lim_{\substack{t'\downarrow t\\ s\uparrow t}}\frac{1}{t'-s}[\mathsf{E}f(t', \xi_{xs}(t'))-f(t', x)]=(L_c+L_d)f(t, x).$$

If $A(x, t)\in\bar{S}^c(K, C_N)$ it is sufficient to require, instead of the boundedness of $f(t, x)$ and its partial derivatives of the first and second orders with respect to x, that these derivatives increase as $x\to\infty$ not faster than a power of $|x|$.

This assertion is actually contained in the proof of Theorem 3.

Differentiability with respect to initial data of solutions of stochastic equations. Consider the problem of differentiability with respect to x of a solution $\xi_{xs}(t)$ of equation

$$
(11)\qquad \begin{aligned} d\xi_{xs}(t) &= A(\xi_{xs}(t), dt), \qquad t > s, \\ \xi_{xs}(s) &= x, \end{aligned}
$$

where $A(x, t) \in S(\tilde{C}, \lambda_N)$.

In what follows we shall interpret derivatives of random functions with respect to x in two different senses: as the ordinary derivatives existing with probability 1 and as mean square derivatives.

As far as the martingale field $\beta(x, t)$ is concerned, its derivative with respect to x would be interpreted in the sense of the mean square convergence.

Let δ_k be a vector with components $(\delta_{1k}, \delta_{2k}, \ldots, \delta_{mk})$, where $\delta_{kj} = 0$ for $k \neq j$ and $\delta_{kk} = 1$. Then

$$
\frac{\partial}{\partial x^k} \beta(x, t) \underset{\text{Def}}{=} \text{l.i.m.} \frac{\beta(x + h\delta_k, t) - \beta(x, t)}{h}.
$$

Several remarks connected with the differentiability of a square integrable martingale field $\beta(x, t)$ are in order. It easily follows from Theorem 16 in Section 1 of Chapter I that if the limit

$$
\underset{h \to 0}{\text{l.i.m.}} \frac{\beta(x + hy, t) - \beta(x, t)}{h}
$$

exists for $t = T$, then this limit also exists for any $t \in [0, T]$ and is a square integrable martingale.

Denote by $B(x, y, t)$ the joint matrix characteristic of martingales $\beta(x, t)$ and $\beta(y, t)$ and assume that it is absolutely continuous with respect to the Lebesgue measure:

$$
B(x, y, t) = \int_0^t b(x, y, s)\, ds.
$$

A necessary and sufficient condition for the existence of the mean square derivative $(\partial/\partial x^k)\beta^i(x, t)$ (see Volume I, Chapter IV, Section 1) is the existence of the limit

$$
\begin{aligned}
&\lim_{h_1 \to 0\, h_2 \to 0} \mathsf{E} \frac{\beta^i(x + h_1\delta_k, t) - \beta^i(x, t)}{h_1} \frac{\beta^i(x + h_2\delta_k, t) - \beta^i(x, t)}{h_2} \\
(12)\qquad &= \lim_{h_1 \to 0\, h_2 \to 0} \mathsf{E} \frac{1}{h_1 h_2} [B^{ii}(x + h_1\delta_k, x + h_2\delta_k, t) \\
&\qquad - B^{ii}(x + h_1\delta_k, x, t) - B^{ii}(x, x + h_2\delta_k, t) + B^{ii}(x, x, t)].
\end{aligned}
$$

We shall require somewhat more, namely, that with probability 1 there exists for each s a continuous (with respect to x), generalized mixed derivative.

$$\frac{\partial^2}{\partial x^k\,\partial y^k} b^{jj}(x, x, s)$$

$$= \lim_{h_1\to 0\, h_2\to 0} \frac{1}{h_1 h_2}[b^{jj}(x+h_1\delta_k, x+h_2\delta_k, s)$$

$$-b^{jj}(x+h_1\delta_k, x, s)-b^{jj}(x, x+h_2\delta_k, s)+b^{jj}(x, x, s)], \qquad j=1, \ldots, m,$$

and, moreover,

$$\frac{\partial^2}{\partial x^k\,\partial y^k} b^{jj}(x, x, s) \leq C, \qquad k, j=1, \ldots, m \pmod{d\mathsf{P}\times ds},$$

where C is a nonrandom constant.

Since $b^{jj}(x, y, s)$ is a nonnegative definite kernel the existence of the derivatives $(\partial^2/\partial x^k\,\partial y^k)b^{jj}(x, y, s)$ follows from the existence of the derivative $(\partial^2/\partial x^k\,\partial y^k)b^{jj}(x, x, s)$. The inequality

$$\left|\frac{\partial^2}{\partial x^k\,\partial y^k} b^{jj}(x, y, s)\right| \leq C \pmod{d\mathsf{P}\times ds},$$

as well as the uniform boundedness of the expression appearing under the expectation sign in the right-hand side of equality (12) also follow from the existence of this derivative. This implies that in this case the condition for the existence of a mean square derivative $(\partial/\partial x^k)\beta^j(x, t)$ is fulfilled.

It is easy to verify that the joint characteristic of martingales $(\partial/\partial x^k)\beta^j(x, t)$ and $(\partial/\partial y^r)\beta^i(y, t)$ satisfies the relation

$$\left\langle \frac{\partial}{\partial x^k}\beta^j(x, \cdot), \frac{\partial}{\partial y^r}\beta^i(y, \cdot)\right\rangle_t = \frac{\partial^2}{\partial x^k\,\partial y^r}\int_0^t b^{ji}(x, y, s)\, ds,$$

while the existence of the corresponding derivatives and their continuity in x and y (mod $d\mathsf{P}\times ds$) follows from the preceding assumptions. Moreover,

$$\left\langle \beta^j(x, \cdot), \frac{\partial}{\partial y^r}\beta^k(y, \cdot)\right\rangle_t = \int_0^t \frac{\partial}{\partial y^r} b^{jk}(x, y, s)\, ds,$$

and the characteristic of the martingale

$$\tilde{\beta}_h^j(t) = \frac{\beta^j(x+hy, t)-\beta^j(x, t)}{h} - \nabla\beta^j(x, t)\cdot y$$

satisfies the following relationship:

$$\langle \tilde\beta_h^i, \tilde\beta_h^i\rangle_t = \int_0^t \left\{ \frac{1}{h^2}[b^{ii}(x+hy, x+hy, t) - 2b^{ii}(x+hy, x, t) + b^{ii}(x, x, s)] - \frac{2}{h}[\nabla_y b^{ii}(x+hy, x, s)\cdot y - \nabla_y b^{ii}(x, x, s)\cdot y] + \sum_{k,r=1}^m \frac{\partial^2}{\partial x^k\, \partial y^r} b^{ii}(x, x, s) y^k y^r \right\} ds.$$

Utilizing Taylor's formula and the notation introduced we can rewrite the preceding relationship in the form

$$\langle \tilde\beta_h^i, \tilde\beta_h^i\rangle_t = \int_0^t [\nabla^2 b^{ii}(x, x, s) - \nabla^2 b^{ii}(x+\tilde h y, x+\tilde h y, s)]\cdot y \cdot y\, ds, \tag{13}$$

where $\tilde h$ is a number between 0 and h.

Theorem 4.

a) *Let the function $\alpha(x, t)$ be continuously differentiable with respect to x for a fixed t with probability* 1 *and $|\nabla\alpha(x, t)| \le C$.*

b) *Let the joint matrix characteristic $B(x, y, t)$ of the field $\beta(x, t)$ be differentiable with respect to t,*

$$B(x, y, t) = \int_0^t b(x, y, s)\, ds,$$

and the function $b(x, y, t)$, for a fixed t, possess with probability 1 *continuous and uniformly bounded derivatives $(\partial^2/\partial x^k\, \partial y^k) b(x, y, t)$, i.e.,*

$$\left| \frac{\partial^2}{\partial x^k\, \partial y^k} b(x, y, t) \right| \le C, \qquad k = 1, \ldots, m.$$

c) *Let the field $A(x, t) = \int_0^t \alpha(x, s)\, ds + \beta(x, t) \in \tilde S(C, C)$.*

Then $\xi_{xs}(t)$ is differentiable in the mean square with respect to x^k $(k = 1, \ldots, m)$ and $(\partial/\partial x^k)\xi_{xs}(t) = \eta_k(t)$ satisfies the linear stochastic differential equation

$$\eta_k(t) = \delta_k + \int_s^t \nabla A(\xi_{xs}(v), dv)\cdot \eta_k(v). \tag{14}$$

Proof. For simplicity of notation we set $s = 0$ and $\xi_{xs}(t) = \xi_x(t)$ and let

$$\eta_{h_k}(t) = \frac{1}{h}[\xi_{x+h_k}(t) - \xi_x(t)],$$

where $h_k = h\delta_k$ is a vector with components $h\delta_{kj}$, $j = 1, 2, \ldots, m$. The process $\eta_{h_k}(t)$ satisfies the equation

$$\eta_{h_k}(t) = \delta_k + \int_0^t A_h(\eta_{h_k}(s), ds),$$

where

$$A_h(y, t) = \int_0^t \frac{\alpha(\xi_x(x)+hy, s) - \alpha(\xi_x(s), s)}{h}\, ds$$
$$+ \int_0^t \frac{\beta(\xi_x(s)+h_y, ds) - \beta(\xi_x(s), ds)}{h}$$
$$= \int_0^t \alpha_h(y, s)\, ds + \beta_h(y, t).$$

Denote

$$A_0(y, t) = \int_0^t \nabla\alpha(\xi_x(s), s) \cdot y\, ds + \int_0^t \nabla\beta(\xi_x(s), ds) \cdot y$$
$$= \int_0^t \alpha_0(y, s)\, ds + \beta_0(y, t).$$

We shall verify that the conditions of Theorem 11 in Section 1 are satisfied for the fields $A_h(y, t)$ and $A_0(y, t)$. It follows from the assumptions of Theorem 4 that

$$|\alpha_h(y, t)| \leq C|y|, \qquad \mathsf{E}\{|\Delta\beta_h(y, t)|^2 \mid \mathfrak{F}_t\} \leq C^2|y|^2\, \Delta t.$$

Moreover, in view of Lagrange's formula

$$|\alpha_h(y, t) - \alpha_0(y, t)| = |\nabla\alpha(\xi_x(t)+\tilde{h}y, t) - \nabla\alpha(\xi_x(t), t) \cdot y|,$$

where $|\tilde{h}| \leq |h|$. Since the function $\nabla\alpha(y, t)$ is continuous in y with probability 1 for any $t \in [0, T]$ we have

$$\mathsf{P}\{\sup_{|y|\leq N} |\alpha_h(y, t) - \alpha_0(y, t)| > \varepsilon\} \to 0 \quad \text{as } h \to 0 \quad \forall \varepsilon > 0.$$

Furthermore

$$\mathsf{E}\{|\Delta\beta_h - \Delta\beta_0|^2 \mid \mathfrak{F}_t\} = \mathsf{E}\{\int_t^{t+\Delta t} \gamma_h(y, s)\, ds \mid \mathfrak{F}_t\},$$

where in view of formula (13)

$$\gamma_h(y, t) = \sum_{j=1}^{m} (\nabla^2 b^{ij}[\xi(t)+\tilde{h}y, \xi(t+\tilde{h}y), t] - \nabla^2 b^{ij}[\xi(t), \xi(t), t]) \cdot y \cdot y$$

and $|\tilde{h}| \leq |h|$. Since the functions $(\partial^2/\partial x^k\, \partial y^r) b(x, y, t)$ are continuous with probability 1 jointly in the variables x, y, t, and $\sup |\xi(t)| < \infty$ with probability 1 it is easy to verify that $\mathsf{P}\{\sup_{|y|\leq N} |\gamma_h(y, t)| > \varepsilon\} \to 0$ as $h \to 0$. Thus, the conditions of Theorem 11 in Section 1 are fulfilled. Taking into account the remarks following Theorem 10 in Section 1, we obtain

$$\mathsf{E} \sup_{0\leq t\leq T} |\eta_{hk}(t) - \eta_{0k}(t)|^2 \to 0 \quad \text{as } h \to 0,$$

where $\eta_{0k}(t)$ is a solution of equation (14). The theorem is proved. □

If we strengthen the assumptions about the field $A(x, t)$ we can obtain theorems on the existence of derivatives of the second order of the function $\xi_{xs}(t)$ according to the initial data.

A formal differentiation of equation (14) leads to the relation

$$\eta_{kr}(t)=\int_s^t \nabla^2 A(\xi_{xs}(v), dv)\cdot \eta_k(v)\cdot \eta_r(v)+\int_s^t \nabla A(\xi_{xs}(v), dv)\cdot \eta_{kr}(v), \tag{15}$$

where

$$\eta_{kr}(t)=\frac{\partial^2}{\partial x^k\,\partial x^r}\xi_{xs}(t).$$

In order that the derivative $\eta_{kr}(t)$ possess finite moments of the second order it is natural to require the existence of moments of the fourth order for the variables $\eta_k(t)$ and a uniform, in a certain sense, boundedness of the fields $\nabla^2 A(x, t)$ in x.

First we state the conditions for the existence of moments of the fourth order for a solution of equation (14).

We utilize the generalized Itô formula. For this purpose we decompose the field $\beta(x, t)$ into continuous and discontinuous components, i.e.,

$$\beta(x, t)=\beta_c(x, t)+\zeta(x, t),$$

and let

$$\langle \beta_c(x,\cdot), \beta_c(y,\cdot)\rangle_t=\int_s^t b_c(x, y, s)\,ds,$$
$$\langle \zeta(x,\cdot), \zeta(y,\cdot)\rangle_t=\int_s^t b_d(x, y, s)\,ds.$$

Assume that matrices $b_c(x, y, t)$ and $b_d(x, y, t)$ possess, with probability 1, continuous mixed derivatives $(\partial^2/\partial x^k\,\partial y^r)(\cdot)$. Then the fields $\beta_c(x, t)$ and $\zeta(x, t)$ are mean square differentiable with respect to x^k $(k=1,\dots,m)$. Set

$$A^\nabla(y, t)=\int_s^t \nabla A(\xi_{xs}(\theta), d\theta)\cdot y,$$

or, in more detail,

$$\begin{aligned} A^\nabla(y, t)&=\int_s^t \nabla\alpha(\xi_{xs}(\theta), \theta)\cdot y\,d\theta+\int_s^t \nabla\beta_c(\xi_{xs}(\theta), d\theta)\cdot y+\int_s^t \nabla\zeta(\xi_{xs}(\theta), d\theta)\cdot y\\ &=\int_s^t \alpha^\nabla(y, \theta)\,d\theta+\beta_c^\nabla(y, t)+\zeta^\nabla(y, t). \end{aligned}$$

The matrix characteristic of the process $\beta_c^\nabla(y, t)$ is equal to

$$\langle \beta_c^\nabla(y,\cdot), \beta_c^\nabla(y,\cdot)\rangle_t=\int_s^t \sum_{k,r=1}^m \frac{\partial^2}{\partial x^k\,\partial y^r} b_c(\xi_{xs}(\theta), \xi_{xs}(\theta), \theta) y^k y^r\,d\theta$$

(Lemma 8, Section 1) and an analogous expression is valid for the characteristic of the process $\zeta^\nabla(y, t)$.

It follows from the assumptions in Theorem 4 that

$$\left|\frac{\partial^2}{\partial x^k\,\partial y^r}b_c(x, y, t)\right| + \left|\frac{\partial^2}{\partial x^k\,\partial y^r}b_d(x, y, t)\right| \leq C,$$

$$|\alpha^\nabla(y, t)| \leq C(1+|y|),$$

and hence the field $A^\nabla(y, t) \in S(C, C)$.

Consider the equation

$$\eta(t) = z + \int_0^t A^\nabla(\eta(s), ds), \tag{16}$$

and for simplicity set $s = 0$. It follows from Theorem 9 in Section 1 that if $A^\nabla(y, t) \in S(C, C)$ and if, moreover,

$$\int_{\mathcal{R}^m} |u|^4 \pi^\nabla(y, T, du) \leq C(1+|y|^4), \tag{17}$$

where $\pi^\nabla(x, t, A)$ is a measure associated with the measure of jumps $\nu_y^\nabla(t, A)$ of the process $A^\nabla(y, t)$, then a solution of equation (16) possesses finite moments of the fourth order.

We now return to equation (15). We shall assume that the conditions of Theorem 4 and those given by (17) are satisfied. For simplicity we again set $s = 0$, $\xi_{xs}(t) = \xi_{x0}(t) = \xi_x(t)$. It is also necessary to assume the existence of the field $\nabla^2 A(x, t)$ and of the process

$$\varphi(t) = \int_0^t \nabla^2 A(\xi_x(v), dv) \cdot \eta_k(v) \cdot \eta_r(v).$$

Here the integral in the right-hand side of the equality is understood to be a line integral along the random curve $\xi_x(t)$ in the field

$$A_{kr}^{(2)}(x, t) = \int_0^t \nabla^2 A(x, dv) \cdot \eta_k(v) \cdot \eta_r(v).$$

We represent the random function $(\partial^2/\partial x^k\,\partial x^r)A(x, t)$ in the form

$$\frac{\partial^2}{\partial x^k\,\partial x^r}A(x, t) = \int_0^t \frac{\partial^2}{\partial x^k\,\partial x^r}\alpha(x, s)\,ds + \frac{\partial^2}{\partial x^k\,\partial x^r}\beta(x, t),$$

and assume the following:

a) $\alpha(x, t)$ is, with probability 1, twice continuously differentiable with respect to x for each $t \in [0, T]$ and

$$\left|\frac{\partial^2}{\partial x^k\,\partial x^r}\alpha(x, t)\right| \leq C, \qquad k, r = 1, \ldots, m,$$

where C is a nonrandom constant.

b) There exists, with probability 1, for each $t \in [0, T]$ the partial derivative

$$\frac{\partial^4}{\partial x^k\, \partial y^k\, \partial x^r\, \partial y^r} b^{ii}(x, y, t), \tag{18}$$

continuous in x and y and bounded for all x, y, and t by an absolute constant C.

The derivative (18) is interpreted here as the mixed derivative $\partial^2/\partial x^r\, \partial y^r$ of the derivative $(\partial^2/\partial x^k\, \partial y^k) b^{ii}(x, y, t)$ in the sense described above.

If condition b) is satisfied, then it follows from the above that the derivative in the mean square

$$\frac{\partial^2}{\partial x^k\, \partial x^r} \beta(x, t) = \frac{\partial}{\partial x^r}\left(\frac{\partial}{\partial x^k} \beta(x, t)\right)$$

exists. We shall now discuss the function

$$\begin{aligned} A^{(2)}(x, t) &= \int_0^t \nabla^2 \alpha(x, v) \cdot \eta_k(v) \cdot \eta_r(v)\, dv + \int_0^t \nabla^2 \beta(x, dv) \cdot \eta_k(v) \cdot \eta_r(v) \\ &= \int_0^t \alpha^{(2)}(x, v)\, dv + \beta^{(2)}(x, t). \end{aligned}$$

Clearly the first integral exists with probability 1 and possesses finite moments of the second order. The second integral is a square integrable martingale field. The joint characteristics of its components can be expressed as:

$$\begin{aligned} &\langle \beta^{(2)p}(x, \cdot), \beta^{(2)q}(y, \cdot) \rangle_t \\ &= \int_0^t \sum_{i,j,i',j'=1}^{m} \frac{\partial^4 b^{pq}(x, y, v)}{\partial x^i\, \partial y^j\, \partial x^{i'}\, \partial y^{j'}} \eta_k^i(v) \eta_r^j(v) \eta_k^{i'}(v) \eta_r^{j'}(v)\, dv. \end{aligned}$$

In view of these remarks the function $\varphi(t)$ exists and possesses finite moments of the second order. Also it follows from the available bounds that $\mathsf{E} \sup_{0 \le t \le T} |\varphi(t)|^2 < \infty$ as well. In that case, however, equation (15) possesses a unique solution and $\mathsf{E} \sup_{0 \le t \le T} |\eta_{kr}(t)|^2 < \infty$.

We now proceed to the problem of differentiability of the solution of equation (14) with respect to x. Denote the solution of equation (14) by $\eta(t, x)$ and set

$$\eta_h(t) = \frac{1}{h}[\eta(t, x + h\delta_r) - \eta(t, x)].$$

Function $\eta_h(t)$ satisfies the equation

$$\eta_h(t) = \varphi_h(t) + \int_0^t \nabla A(\xi_x(v), dv) \cdot \eta_h(v),$$

where

$$\varphi_h(t) = \frac{1}{h} \int_0^t [\nabla A(\xi_{x+h\delta_r}(v), dv) - \nabla A(\xi_x(v), dv)] \eta(v, x + h\delta_r).$$

Note that

$$\varphi_h(t)-\varphi(t)=\varphi'_h(t)+\varphi''_h(t),$$

$$\varphi'_h(t)=\int_0^t\left[\frac{1}{h}(\nabla A(\xi_{x+h\delta_r}(v),dv)-\nabla A(\xi_x(v),dv))\right.$$
$$\left.-\nabla^2A(\xi_x(v),dv)\eta_r(v)\right]\cdot\eta(v,x+h\delta_r),$$
$$\varphi''_h(t)=\int_0^t\nabla^2A(\xi_x(v),dv)\cdot\eta_r(v)\cdot[\eta(v,x+h\delta_r)-\eta(v,x)].$$

Utilizing the boundedness and continuity in x and y of function (18) as well as the expression for the joint characteristics of the field it is easy to arrive at relation

$$\mathsf{E}\sup_{0\le t\le T}|\varphi'_h(t)|^2\to 0.$$

Moreover,

$$\mathsf{E}\sup|\eta(t,x+\Delta x)-\eta(t,x)|^4=O(|\Delta x|^4).$$

This implies that

$$\mathsf{E}\sup_{0\le t\le T}|\varphi_h(t)-\varphi(t)|^2\to 0,$$

and in view of the remark following Theorem 11 in Section 1 we have

$$\mathsf{E}\sup_{0\le t\le T}|\eta_n(t)-\eta(t)|^2\to 0\quad\text{as } h\to 0.$$

Thus the following theorem is valid.

Theorem 5. *If the conditions of Theorem* 4 *are satisfied and inequality* (17) *and the above stated conditions* a) *and* b) *are fulfilled, then the derivatives* $\eta_{kr}(t)=(\partial^2/\partial x^k\,\partial x^r)\xi_{xs}(t)$ *exist in the mean square sense, satisfy equation* (15), *and are continuous in the mean square in* (x,s).

In the stated theorem only the continuity in the mean square of the second derivative of $\eta_{kr}(t)$ in (x,s) remains unproved. This can, however, be easily established analogously to Lemma 2 using the bounds obtained in Theorem 9 in Section 1.

Kolmogorov's equation. Let $\xi_{xs}(t)$ be a solution of equation (5) without an after-effect. It turns out that the function

$$F(t,x)=\mathsf{E}f(\xi_{xt}(T)),\qquad (t,x)\in[0,T)\times\mathcal{R}^m,$$

satisfies an important integro-differential equation of the form which does not depend on function $f(x)$. The dependence on $f(x)$ manifests itself in the boundary condition which should be adjoined to this equation.

Lemma 4. *Let the conditions of Theorem 4 be satisfied, $A(x, t) \in \bar{S}(C, C)$, and let the function $f(x)$ be twice continuously differentiable and its partial derivatives of the second order be uniformly bounded. Then the derivatives $\partial F(t, x)/\partial x^k$ exist, are continuous in (x, t), and*

$$\frac{\partial F(t, x)}{\partial x^k} = \mathsf{E}\left(\nabla f(\xi_{xt}(T)), \frac{\partial}{\partial x^k}\xi_{xt}(T)\right). \tag{19}$$

Proof. Indeed, let $h_k = h\delta_k$:

$$\begin{aligned}
&\left|\frac{F(t, x+h_k)-F(t, x)}{h} - \mathsf{E}\left(\nabla f(\xi_{xt}(T)), \frac{\partial}{\partial x^k}\xi_{xt}(T)\right)\right| \\
&\leqslant \left|\mathsf{E}\left(\nabla f(\xi_{xt}(T)), \frac{\Delta\xi_{xt}}{h} - \frac{\partial}{\partial x^k}\xi_{xt}(T)\right)\right| + \left|\mathsf{E}\nabla^2 f(\xi_{xt}(T)+\theta\Delta\xi_{xt})\cdot\frac{\Delta\xi_{xt}}{h}\cdot\Delta\xi_{xt}\right| \\
&\leqslant C\left[\mathsf{E}(1+|\xi_{xt}(T)|^2)\cdot\mathsf{E}\left|\frac{\Delta\xi_{xt}}{h} - \frac{\partial}{\partial x^k}\xi_{xt}(T)\right|^2\right]^{1/2} + C\left[\mathsf{E}|\Delta\xi_{xt}|^2\cdot\mathsf{E}\left|\frac{\Delta\xi_{xt}}{h}\right|^2\right]^{1/2},
\end{aligned}$$

where C is constant depending only on $\sup|\nabla^2 f(x)|$, $\Delta\xi_{xt} = \xi_{x+h_k t}(T) - \xi_{xt}(T)$. The inequality thus obtained yields the assertion of the lemma. □

Remark. The lemma is also valid for a twice continuously differentiable function $f(x)$ increasing as $|x| \to \infty$ not faster than a power of x provided we stipulate additionally the finiteness of moments of a suitable order for the variables $\xi_{xt}(T)$.

Lemma 5. *Let function $f(x)$ be twice continuously differentiable and its partial derivatives of the second order be uniformly bounded and let the function $\xi_{xt}(T)$ possess mean square partial derivatives with respect to x^k of the first and second orders which are continuous in the mean square in variables (x, t).*

Then the function $F(t, x)$ possesses partial derivatives with respect to x of the second order,

$$\begin{aligned}
\frac{\partial^2 F(t, x)}{\partial x^k\,\partial x^j} &= \left(\mathsf{E}\nabla^2 f(\xi_{xt}(T))\frac{\partial}{\partial x^k}\xi_{xt}(T), \frac{\partial}{\partial x^j}\xi_{xt}(T)\right) \\
&\quad + \left(\mathsf{E}\nabla f(\xi_{xt}(T)), \frac{\partial^2}{\partial x^k\,\partial x^j}\xi_{xt}(T)\right),
\end{aligned} \tag{20}$$

and these derivatives are continuous in (t, x).

Proof. Set

$$\frac{1}{h}\left(\frac{\partial F(t, x+h_k)}{\partial x^j}-\frac{\partial F(t, x)}{\partial x^j}\right)-\left(\mathsf{E}\nabla^2 f(\xi_{xt}(T))\frac{\partial}{\partial x^k}\xi_{xt}(T), \frac{\partial}{\partial x^j}\xi_{xt}(T)\right)$$

$$-\left(\mathsf{E}\nabla f(\xi_{xt}(T)), \frac{\partial^2}{\partial x^k\,\partial x^j}\xi_{xt}(T)\right)=z_1+z_2+z_3+z_4,$$

where

$$z_1=\left(\mathsf{E}\nabla f(\xi_{xt}(T)), \frac{1}{h}\left[\frac{\partial}{\partial x^j}\xi_{x+h_kt}(T)-\frac{\partial}{\partial x^j}\xi_{xt}(T)-\frac{\partial^2}{\partial x^j\,\partial x^k}\xi_{xt}(T)\right]\right),$$

$$z_2=\left(\mathsf{E}[\nabla^2 f(\tilde{\xi})-\nabla^2 f(\xi_{xt}(T))]\frac{\Delta\xi}{h}, \frac{\partial}{\partial x^j}\xi_{x+h_kt}(T)\right),$$

$$z_3=\left(\mathsf{E}\nabla^2 f(\xi_{xt}(T))\left[\frac{\Delta\xi}{h}-\frac{\partial}{\partial x^k}\xi_{xt}(T)\right], \frac{\partial}{\partial x^j}\xi_{x+h_kt}(T)\right),$$

$$z_4=\left(\mathsf{E}\nabla^2 f(\xi_{xt}(T))\frac{\partial}{\partial x^k}\xi_{xt}(T), \left[\frac{\partial}{\partial x^j}\xi_{x+h_kt}(T)-\frac{\partial}{\partial x^j}\xi_{xt}(T)\right]\right).$$

Here $\tilde{\xi}$ denotes a point situated on the interval joining the points $\xi_{xt}(T)$ and $\xi_{x+h_kt}(T)$.

Since the function $\nabla f(x)$ increases as $|x|\to\infty$ not faster than $|x|$ and the mean square derivative $(\partial^2/\partial x^j\,\partial x^k)\xi_{xt}(T)$ exists, the quantity $z_1\to 0$ as $h\to 0$. For z_2 we have the following bound:

$$|z_2|\leqslant\left[\mathsf{E}\left(\frac{\Delta\xi}{h}\right)^2\right]^{1/2}\left[\mathsf{E}|\nabla^2 f(\tilde{\xi})-\nabla^2 f(\xi_{xt}(T))|^2\left|\frac{\partial}{\partial x^j}\xi_{xt}(T)\right|^2\right]^{1/2}$$

$$+\left[\mathsf{E}|\nabla^2 f(\tilde{\xi})-\nabla^2 f(\xi_{xt}(T))|^2\left|\frac{\partial}{\partial x^j}\xi_{x+h_kt}(T)-\frac{\partial}{\partial x^j}\xi_{xt}(T)\right|^2\right]^{1/2}.$$

It is easily seen that this bound implies that $|z_2|\to 0$ as $h\to 0$. Analogously one can verify that $|z_3|\to 0$ and $|z_4|\to 0$ as $h\to 0$. We thus establish the existence of partial derivatives $(\partial^2/\partial x^k\,\partial x^j)F(t, x)$ and have shown the validity of formula (20).

Formula (20) shows that the continuity is mean square in (x, t) of the derivatives $(\partial/\partial x^k)\xi_{xt}(T)$ and $(\partial^2/\partial x^k\,\partial x^j)\xi_{xt}(T)$ implies the continuity in (t, x) of the derivatives $(\partial^2/\partial x^k\,\partial x^j)F(t, x)$. □

Let the conditions of the lemma be fulfilled and let $0\leqslant t'<t<t''<T$. Then

$$F(t', x)=\mathsf{E}f(\xi_{\xi_{xt'}(t'')}(t'')(T))$$

$$=\mathsf{E}\{[\mathsf{E}f(\xi_{yt''}(T))]_{y=\xi_{xt'}(t'')}\}=\mathsf{E}F(t'', \xi_{xt'}(t'')).$$

Consequently,

$$\frac{F(t', x)-F(t'', x)}{t''-t'}=\frac{1}{t''-t'}\mathsf{E}\{F(t'', \xi_{xt'}(t''))-F(t'', x)\}.$$

Since function $F(t, x)$ is twice continuously differentiable, Theorem 3 and Remark 2 are applicable to this theorem.

We thus obtain the following:

Theorem 6

a) *Let $A(x, t) \in \bar{S}(C, C_N)$ and the solution $\xi_{xs}(t)$ of the stochastic differential equation* (5) *possess mean square partial derivatives of the first and second order with respect to (x, s) and be continuous in the mean square in (x, s).*

b) *Let the function $f(x)$ be twice continuously differentiable and uniformly bounded and the partial derivatives of the first and second orders of $f(x)$ also be uniformly bounded.*

Then the function

$$F(t, x)=\mathsf{E}f(\xi_{xt}(T)) \tag{21}$$

is twice continuously differentiable with respect to x, differentiable with respect to t, and satisfies the equation

$$\begin{aligned}&\frac{\partial F(t, x)}{\partial t}+(\nabla F(t, x), a(x, t))+\frac{1}{2}\sum_{k,j=1}^{m} \nabla^k\nabla^j F(t, x)b^{kj}(x, t)\\&+\int_{\mathscr{R}^m}[F(t, x+u)-F(t, x)-(\nabla F(t, x), u)]\Pi(x, t, du)=0\end{aligned} \tag{22}$$

and the boundary condition

$$\lim_{t\to T} F(t, x)=f(x). \tag{23}$$

Corollary 1

a) *Let a nonrandom function $a(x, t)$ be continuous and twice continuously differentiable with respect to x and the partial derivatives with respect to x of the first and second orders be uniformly bounded.*

b) *Let a random function $\beta(x, t)$ for a fixed x be a process with independent increments with finite moments of the second order*

$$\mathsf{E}\beta(x, t)=0, \qquad \mathsf{E}\beta(x, t)\beta^*(y, t)=\int_0^t b(x, y, s)\, ds,$$

and let $b(x, y, t)$ possess uniformly bounded mixed partial derivatives of the second and fourth orders of the form

$$\frac{\partial^2}{\partial x^k\,\partial y^r}b^{ij}(x, y, t), \qquad \frac{\partial^4}{\partial x^k\,\partial y^k\,\partial x^r\,\partial y^r}b^{ij}(x, y, t).$$

c) *Let the discontinuous component of process $\nabla\beta(x, t)$ satisfy condition* (17).

d) *Let the function $f(x)$ be twice continuously differentiable with respect to x and uniformly bounded; also, let its partial derivatives of the first and second order be uniformly bounded.*

Then the function $F(t, x) = \mathsf{E}f(\xi_{xt}(T))$, where $\xi_{xt}(s)$ is a Markov process defined by the stochastic equation

$$d\xi(s) = a(\xi(s), s)\,ds + \beta(\xi(s), ds), \qquad s \geq t, \tag{24}$$
$$\xi(t) = x,$$

satisfies equation (22) *and the boundary condition* (23).

We now show how one can arrive at a very general equation of the form (22) starting from the simplest probabilistic objects such as standard Wiener processes and the Poisson measure.

Corollary 2. *Assume that $w_1(t), \ldots, w_q(t)$ are mutually independent Wiener processes and $\nu(A, \Delta)$ is the Poisson measure on $\mathcal{R}^q \times [0, T]$ independent of the Wiener processes $w_j(t)$, $j = 1, \ldots, q$,*

$$\mathsf{E}\nu(A, [0, T]) = \Pi(A)t, \qquad \tilde{\nu}(A, \Delta) = \nu(A, \nu) - \Pi(A)\,\Delta t.$$

Let $a(x, t)$, $\sigma_k^j(x, t)$, $g^j(x, t, u)$ be nonrandom functions, $j = 1, \ldots, m$, $k = 1, 2, \ldots, q$, $(x, t, u) \in \mathcal{R}^m \times [0, T] \times \mathcal{R}^q$, satisfying the following conditions:

a) *Functions $a^j(x, t)$, $\sigma_k^j(x, t)$, and $g^j(x, t, u)$, $j = 1, \ldots, m$, $k = 1, \ldots, q$, are continuous in x, t and twice continuously differentiable with respect to x.*

b) *Partial derivatives with respect to x of the first and second orders of functions $a^j(x, t)$ and $\sigma_k^j(x, t)$ are uniformly bounded.*

c)
$$\int (|\nabla g|^2 + |\nabla g|^4 + |\nabla^2 g|^2)\Pi(du) \leq C,$$

where C does not depend on (x, t).

Denote by $\xi_{xt}(s)$ the solution of the stochastic differential equation

$$d\xi(s) = a(\xi(s), s)\,ds + \sum_{k=1}^{q} \sigma_k(\xi(s), s)\,dw^k(s)$$
$$+ \int_{\mathcal{R}^q} g(\xi(s), s, u)\tilde{\nu}(du, ds),$$
$$\xi(t) = x.$$

Then the function $F(t, x) = \mathsf{E}f(\xi_{xt}(T))$, where $f(x)$ satisfies the conditions of Theorem 6, *is a solution of equation*

$$\frac{\partial F(t, x)}{\partial t} + (a(t, x), \nabla F(t, x)) + \frac{1}{2}\sum_{k,j=1}^{m} b^{kj}(x, t)\nabla^k\nabla^j F(t, x)$$
$$+ \int_{\mathcal{R}^q} [F(t, x + g(t, x, u)) - F(t, x) \tag{25}$$
$$- (g(t, x, u), \nabla F(t, x))]\Pi(du) = 0,$$

where $b^{kj}(x, t) = \sum_{r=1}^{q} \sigma_r^k(x, t)\sigma_r^j(x, t)$.

To prove this assertion we observe that the field

$$\beta_c(x, t) = \sum_{k=1}^{q} \int_0^t \sigma_k(x, s)\, dw(s)$$

possesses the joint characteristic defined by relation

$$\mathsf{E}\{\beta_c^j(x, \Delta)\beta_c^k(y, \Delta) \mid \mathfrak{F}_t\} = \mathsf{E}\beta_c^j(x, \Delta)\beta_c^k(y, \Delta)$$

$$= \int_t^{t+\Delta t} \sum_{r=1}^{q} \sigma_r^j(x, s)\sigma_r^k(y, s)\, ds,$$

where $\mathfrak{F}_t$ is the completion of the σ-algebra generated by the random variables $w^k(s)$, $\nu(A, s)$, $s \leq t$, $A \in \mathfrak{B}^m$, $k = 1, \ldots, q$.

Analogously for the field

$$\zeta(x, t) = \int_0^t \int_{\mathcal{R}^q} g(x, s, u)\tilde{\nu}(du, ds)$$

we have

$$\mathsf{E}\{\zeta^j(x, \Delta)\zeta^k(y, \Delta) \mid \mathfrak{F}_t\} = \int_t^{t+\Delta t} \int_{\mathcal{R}^q} g^j(x, s, u) g^k(y, s, u) \Pi(du)\, ds.$$

It follows easily from the stipulated assumptions that the field

$$A(x, t) = \int_0^t a(x, s)\, ds + \beta_c(x, t) + \zeta(x, t)$$

satisfies the conditions of Theorem 6. Moreover,

$$\Pi(x, t, B) = \Pi\{u \colon g(x, t, u) \in B\}.$$

If we replace the variable of integration $u \to g(t, x, u)$ in equation (22) for the function $F(t, x)$, then equation (22) becomes (25).

Formula (21) can be viewed as a probabilistic representation of the solution of Cauchy's problem for an integro-differential equation in partial derivatives (22). On the one hand, equation (22) can be used, for example, for defining the transition probabilities of a Markov process $\xi_{xt}(s)$ or for a study of the analytical properties of these probabilities. On the other hand, if it is required to obtain a numerical or an approximate solution of equation (22) (or that of (25)), then one can utilize expression (21) for a probabilistic modeling of this solution (the Monte Carlo method). Theorems proved above dealing with the convergence of finite-difference approximations of solutions of stochastic differential equations serve, in particular, as a basis and as a justification for a simple finite-difference approximation procedure for solving equation (22) (or (25)).

One can extend the class of integro-differential equations in partial derivatives associated with solutions of stochastic differential equations. For this purpose consider the problem of determining the distribution of the random vector

$$\int_t^T h(\xi_{xt}(s), s)\, ds,$$

where $h(x, t)$, $(x, t) \in \mathcal{R}^m \times [0, T]$ is a function with values in $\mathcal{R}^q$ continuous and twice continuously differentiable with respect to x with uniformly bounded partial derivatives of the first and second orders.

To solve this problem we proceed as follows. We adjoin relations

$$d\eta(s) = h(\xi_{xt}(s), s)\, ds, \qquad s \geq t,$$
$$\eta(t) = y$$

to equation (24) and interpret them as a single stochastic differential equation

$$(26) \qquad \begin{aligned} &d\zeta_{zt}(s) = B(\zeta_{zt}(s), ds); \\ &\zeta_{zt}(s) = (\xi_{xt}(s); \eta(s)), \qquad \zeta_{zt}(s) = z, \quad z = (x, y). \end{aligned}$$

Set

$$\bar{F}(t, x, y) = \bar{F}(t, z) = \mathsf{E}\bar{f}(\zeta_{zt}(T)),$$
$$\bar{f}(z) = \bar{f}(x, y) = f(x) \exp\{i(\lambda, x) + i(\mu, y)\},$$

where $f(x)$ is a twice continuously differentiable function with uniformly bounded partial derivatives of the first and second orders, λ is an m-dimensional vector and μ is a q-dimensional vector. Theorem 6 is applicable to equation (26); hence $F(z, t)$ satisfies the equation

$$(27) \qquad \begin{aligned} &\frac{\partial}{\partial t}\bar{F} + (a, \nabla_x \bar{F}) + (h, \nabla_y \bar{F}) + \frac{1}{2}\sum_{k,j=1}^{m} b^{kj}\nabla_x^k \nabla_x^j \bar{F} \\ &+ \int_{\mathcal{R}^m} [\bar{F}(t, x+u, y) - \bar{F}(t, x, y) - (u, \nabla_x \bar{F})]\Pi(x, t, du) = 0. \end{aligned}$$

Here ∇_x denotes the gradient with respect to x and ∇_y with respect to y. Since $\eta(T) = y + \int_t^T h(s, \xi_{xt}(s))\, ds$, it follows that $\nabla_y \bar{F} = i\mu\bar{F}$. Replacing $\nabla_y \bar{F}$ by $i\mu\bar{F}$ in formula (27) and setting $y = 0$ we obtain

$$\begin{aligned} F(t, x) &= \bar{F}(t, x, 0) \\ &= \mathsf{E}f(\xi_{xt}(T)) \exp\{i(\lambda, \xi_{xt}(T)) + i(\mu, \int_t^T h(s, \xi_{xt}(s))\, ds)\}. \end{aligned}$$

Function $F(t, x)$ satisfies the equation

$$(28) \qquad \begin{aligned} &\frac{\partial}{\partial t}F + (a, \nabla F) + \frac{1}{2}\sum_{k,j=1}^{m} b^{kj}\nabla^k \nabla^j F + i(\mu, h)F \\ &+ \int_{\mathcal{R}^m} [F(t, x+u) - F(t, x) - (u, \nabla F(t, x))]\Pi(x, t, du) = 0 \end{aligned}$$

and the boundary condition

$$F(T, x)=f(x)\, e^{i(\lambda, x)}.$$

If we set $f(x)=1$ then the joint characteristic function of the distribution of random vectors $(\xi_{xt}(T), \int_t^T h(\xi_{xt}(s), s)\, ds)$ will satisfy equation (28). If we set $F(T, x)=1$ we obtain an equation for the characteristic function of the distribution of the additive functional under consideration on the solution $\xi_{xt}(s)$ for the stochastic differential equation (24). Equation (28) differs from equation (22) by the presence of the additional term

$$i(\mu, h(x, t))F(t, x).$$

Example. Distribution of an additive functional on a Wiener process. We present several remarks concerning the evaluation of the distribution of homogeneous additive functionals (of an integral type) on a homogeneous Wiener process.

In the case under consideration $\xi_{xt}(s)=x+w(s)-w(t)$, $s \geqslant t$, and

$$\eta(T)=\int_t^T h(x+w(s)-w(t))\, ds.$$

The function

$$F(t, x)=\mathsf{E}\, e^{i\mu\eta(T)} f(x+w(T)-w(t))$$

satisfies the equation

$$\frac{\partial F(t, x)}{\partial t}+\frac{1}{2}\frac{\partial^2 F(t, x)}{\partial x^2}+i\mu h(x)F(t, x)=0, \qquad t<T,$$

and the boundary condition

$$F(T, x)=f(x).$$

Setting $v(T-t, x)=F(t, x)$ we obtain the following equation for the function $v(t, x)$

$$\frac{\partial v(t, x)}{\partial t}=\frac{1}{2}\frac{\partial^2 v(t, x)}{\partial x^2}+i\mu h(x) v(t, x), \tag{29}$$

with the initial condition $v(0, x)=f(x)$. Moreover, the function $v(t, x)$ can be represented in the form

$$v(t, x)=\mathsf{E} \exp\{i\mu \int_0^t h(w(s)+x)\, ds\} f(x+w(t)).$$

Since the process $w(s)$ is stochastically equivalent to the process $\sqrt{t}\, w(s/t)$ the last expression for $v(t, x)$ can be replaced by the following:

$$v(t, x) = \mathsf{E} \exp\{i\mu t \int_0^1 h(\sqrt{t}\, w(s)+x)\, ds\} f(x+\sqrt{t}\, w(1)). \tag{30}$$

Formula (30) yields a solution of the Cauchy problem for a parabolic equation (29) using the "quadrature" method (in the present case a "quadrature" is interpreted as an integral of a functional defined on $\mathscr{C}[0, 1]$ and the integration is carried over the standard Wiener measure, i.e., the measure generated in $\mathscr{C}[0, 1]$ by a Wiener process).

In what follows we shall set $f(x) \equiv 1$. In other words, we shall be dealing with the evaluation of the characteristic function of the distribution of variable $\eta(T)$. Equation (29) can be solved utilizing the Laplace transform with respect to t. Set

$$z(p, x) = \int_0^\infty e^{-pt} v(t, x)\, dt,$$

where p is a nonnegative number. Multiplying equation (29) by e^{-pt} and integrating with respect to t from 0 to ∞ we obtain

$$pz(p, x) - 1 = \frac{1}{2} \frac{\partial^2}{\partial x^2} z(p, x) + i\mu h(x) z(p, x). \tag{31}$$

We show that equation (31) is also valid in the case when $h(x)$ is a piecewise continuous bounded function. We choose a sequence of uniformly bounded functions $h_n(x)$ such that these functions converge for each x to $h(x)$ and each one of them is twice continuously differentiable and possesses bounded derivatives of the first and second orders. Let

$$z_n(p, x) = \int_0^\infty e^{-pt}\, \mathsf{E} \exp\{i\mu \int_0^t h_n(x+w(s))\, ds\}\, dt.$$

Then $|z_n(p, x)| \le 1/p$ and $z_n(p, x) \to z(p, x)$ as $n \to \infty$. Functions $z_n(p, x)$ satisfy equation (31). It can be seen from this equation that the derivatives $(\partial^2/\partial x^2) z_n(p, x)$ are uniformly bounded and converge to the limit $2(pz(p, x) - 1 - i\mu h(x) z(p, x))$ as $n \to \infty$. This implies the following.

Theorem 7. *If the function $h(x)$ is bounded and piecewise continuous, then the function $z(p, x)$ is continuously differentiable, possesses at all points of continuity of function $h(x)$ a second derivative, and satisfies equation* (31).

We shall utilize Theorem 7 to evaluate the distribution of the variable

$$\eta(t) = \int^t \operatorname{sgn} w(s)\, ds.$$

In the case under consideration equation (31) becomes

$$z''(p, x) + 2(i\mu \operatorname{sgn} x - p) = -2.$$

Solving this equation separately in the regions $x>0$ and $x<0$ we obtain

$$z(p,x)=\frac{1}{p-i\mu}+C_1\,e^{\sqrt{2p-2i\mu}\,x}+C_2\,e^{-\sqrt{2p-2i\mu}\,x},\qquad x>0,$$

$$z(p,x)=\frac{1}{p+i\mu}+C_3\,e^{\sqrt{2p+2i\mu}\,x}+C_4\,e^{-\sqrt{2p+2i\mu}\,x},\qquad x<0.$$

Since $z(p,x)$ is bounded as $x\to\pm\infty$ it follows that $C_1=C_4=0$. Utilizing the continuity at $x=0$ of functions $z(p,x)$ and $z'_x(p,x)$ we obtain the equalities

$$\frac{1}{p-i\mu}+C_2=\frac{1}{p+i\mu}+C_3,$$
$$-C_2\sqrt{2p-2i\mu}=C_3\sqrt{2p+2i\mu},$$

which imply that

$$C_3=\frac{1}{p+i\mu}\left[-1+\sqrt{\frac{p+i\mu}{p-i\mu}}\right].$$

To determine the distribution of the variable $\eta(T)$ it is sufficient to have the values of $z(p,0)$. For $|\mu|<p$

$$z(p,0)=\frac{1}{\sqrt{p^2+\mu^2}}=\frac{1}{p}\left(1+\frac{\mu^2}{p^2}\right)^{-1/2}=\frac{1}{p}\sum_{n=0}^{\infty}(-1)^n\frac{(2n-1)!!}{(2n)!!}\left(\frac{\mu}{p}\right)^{2n}.$$

Since

$$\int_0^\infty t^n\,e^{-pt}\,dt=\frac{n!}{p^{n+1}},$$

$$\int_{-\pi/2}^{\pi/2}\sin^k\varphi\,d\varphi=\begin{cases}0 & \text{for an odd } k,\\ \dfrac{(2n-1)!!}{(2n)!!}\pi, & k=2n,\end{cases}$$

we have

$$z(p,0)=\int_0^\infty\left(\sum_{n=0}^{\infty}(-1)^n\frac{(2n-1)!!}{(2n)!!}\,\frac{\mu^{2n}t^{2n}}{(2n)!}\right)e^{-pt}\,dt$$
$$=\int_0^\infty e^{-pt}\left(\sum_{k=0}^{\infty}\frac{1}{\pi}\int_{-\pi/2}^{\pi/2}\sin^k\varphi\frac{(i\mu t)^k}{k!}\,d\varphi\right)dt$$
$$=\frac{1}{\pi}\int_0^\infty e^{-pt}\int_{-\pi/2}^{\pi/2}e^{i\mu t\sin\varphi}\,d\varphi\,dt.$$

Thus

$$\mathsf{E}\exp\left\{i\mu\frac{1}{t}\int_0^t \operatorname{sgn} w(s)\,ds\right\}=\frac{1}{\pi}\int_{-\pi/2}^{\pi/2} e^{i\mu\sin\varphi}\,d\varphi=\int_{-\infty}^{\infty} e^{i\mu x}f(x)\,dx,$$

where

$$f(x)=\begin{cases} 0 & \text{for } |x|>1 \\ \dfrac{1}{\pi}\dfrac{1}{\sqrt{1-x^2}} & \text{for } |x|<1. \end{cases} \tag{32}$$

We thus obtain the following result: the variable $(1/t)\int_0^t \operatorname{sgn} w(s)\,ds$ possesses density (32).

The random variable

$$\tau_t=\int_0^t \frac{1+\operatorname{sgn} w(s)}{2}\,ds$$

has an intuitive meaning. It equals the time spent by the process $w(s)$ on the positive semiaxis during the (time) interval $(0, t)$. Using the density (32) one can find the distribution of the variable τ_t. Indeed,

$$\mathsf{P}\{\tau_t<xt\}=\mathsf{P}\left\{\frac{1}{t}\int_0^t \operatorname{sgn} w(s)\,ds<2x-1\right\}$$

$$=\frac{1}{\pi}\left(\arcsin(2x-1)+\frac{\pi}{2}\right).$$

Usually the formula obtained is written somewhat differently. Observe that

$$\arcsin(2x-1)+\frac{\pi}{2}=\arccos(1-2x).$$

If we set $\frac{1}{2}\arccos(1-2x)=z$, then

$$1-2x=\cos 2z,\qquad x=\frac{1-\cos 2z}{2}=\sin^2 z,\quad z=\arcsin\sqrt{x}.$$

Consequently

$$\mathsf{P}(\tau_t<x)=\frac{2}{\pi}\arcsin\sqrt{\frac{x}{t}},\qquad 0\leqslant x\leqslant t. \tag{33}$$

The result obtained is known as the *arcsine law*.

§ 3. Limit Theorems for Sequences of Random Variables and Stochastic Differential Equations

Let a sequence of series of random vectors

$$(1) \qquad \xi_{n0}, \xi_{n1}, \ldots, \xi_{nm_n}, \qquad n = 1, 2, \ldots$$

with the values in $\mathcal{R}^m$ be given. Assume that the increments $\Delta\xi_{nk} = \xi_{nk+1} - \xi_{nk}$ are small random variables. A classical problem in probability theory is to describe the class of possible limit distributions of the variable ξ_{nm_n} as $n \to \infty$ under varying assumptions on the variables $\Delta\xi_{nk}$. In the case when $\Delta\xi_{nk}$, $k = 0, 1, \ldots, m_n - 1$, are independent we are dealing with the thoroughly investigated problem of summation of the independent variables.

In this subsection we shall consider the general problem of investigating the limit distribution of a sequence of series of random variables (1) from the aspect of the theory of random processes or, more precisely, from its relation to the theory of stochastic differential equations.

A sequence of a series of random vectors (1) will correspond to a sequence of random processes $\xi_n(t)$, to be called processes associated with or generated by the sequence of series (1). To define these processes we must also specify a sequence of real numbers

$$0 = t_{n0} < t_{n1} < \cdots < t_{nm_n-1} < t_{nm_n} = T, \qquad n = 1, 2, \ldots,$$

and then we set

$$\xi_n(t) = \xi_{nk} \quad \text{if } t \in [t_{nk}, t_{nk+1}).$$

If for $n \to \infty$ $\max_{1 \le k \le m_n} \Delta t_{nk} \to 0$ and all the variables $\{\Delta\xi_{nk},\ k = 0, \ldots, m_n - 1\}$ are close in a certain sense—which will be described below—to the variables $\{\Delta\xi(t_{nk}),\ k = 0, 1, \ldots, m_n - 1\}$, where $\xi(t)$, $t \in [0, T]$, is a random process and $\Delta\xi(t_{nk}) = \xi(t_{nk+1}) - \xi(t_{nk})$, then one might expect that the distribution of the variable ξ_{nm_n} converges to the distribution of the variable $\xi(T)$ and, moreover, for continuous functionals $f[x(\cdot)]$ defined on $\mathcal{D}^m[0, T]$ the distribution of variables $f[\xi_n(\cdot)]$ will be close to the distribution of the variable $f[\xi(\cdot)]$.

Thus we would like to incorporate the problem under investigation into the general framework of limit theorems for random processes discussed in Chapter VI of Volume I.

In accordance with the results obtained in Volume I, when studying limit theorems for random processes, two problems can be distinguished: a) the investigation of conditions for the weak convergence of marginal distributions of random processes and the characterization of limiting distributions; and b) the determination of criteria of weak compactness of a sequence of measures corresponding to random processes in an appropriate functional space. General criteria for weak compactness of measures in functional spaces were established in Chapter VI of Volume I. In this section based on results established above we present some sufficient conditions for weak compactness of measures which, for the problems under consideration, are more convenient for verification. Next, we consider the weak convergence of marginal distributions of processes, which are either constructed from the sequence of series (1) or which are solutions of stochastic equations, to the marginal distributions of solutions of stochastic differential equations. In conclusion we present examples of the application of general theorems to more particular models and specific problems.

On a weak compactness of measure in $\mathscr{D}$ associated with sequences of series of random variables. In this section we shall use the letter $\mathscr{D}$ to denote the space $\mathscr{D}^m[0, T]$ and the measure defined on the σ-algebra generated by cylindrical sets in $\mathscr{D}$ will be referred to as the measure in $\mathscr{D}$.

Let $\xi_n(t)$, $n = 1, 2, \ldots, t \in [0, T]$, be a sequence of random processes with values in $\mathscr{R}^m$ and with sample functions belonging to $\mathscr{D}$ with probability 1. The process $\xi_n(t)$ generates on $\mathscr{D}$ a measure q_n, to be called the *measure associated in $\mathscr{D}$ with the process $\xi_n(t)$*, defined on cylindrical sets of the space $\mathscr{D}$ by relations

$$q_n(C_{t_1 t_2 \cdots t_r}(A^r)) = \mathsf{P}\{(\xi_n(t_1), \ldots, \xi_n(t_r)) \in A^r\}.$$

Here A^r is a Borel set in the space $\mathscr{R}^m \times \cdots \times \mathscr{R}^m = \mathscr{R}^{mr}$ and $C_{t_1 \cdots t_r}(A^r) = \{x(\cdot): x(\cdot) \in \mathscr{D},\ (x(t_1), \ldots, x(t_r)) \in A^r\}$ is a cylindrical set with basis A^r over coordinates $(t_1, t_2, \ldots, t_r)$. We shall now be concerned with the conditions under which the sequence of measures $q_n(\cdot)$ converges weakly to a limit. The importance of this problem was clarified in Volume I, Chapter VI. Recall, for instance, that if a sequence $q_n(\cdot)$ is weakly convergent to $q(\cdot)$, where $q(\cdot)$ is the measure associated in $\mathscr{D}$ with a certain process $\xi(t)$, then for any bounded functional $f[x(\cdot)]$ q-almost everywhere continuous (in the metric of the space $\mathscr{D}$) the distribution of the random variable $\zeta_n = f[\xi_n(\cdot)]$ converges weakly to the distribution of the variable $\zeta = f[\xi(\cdot)]$.

In what follows we shall confine ourselves to the derivation of conditions for the weak convergence of measures in $\mathscr{D}$. Results related to the weak convergence of measures in $\mathscr{C} = \mathscr{C}^m[0, T]$ can be obtained as particular cases. The weak convergence of a sequence of measures $q_n(\cdot)$ associated with random processes $\xi_n(\cdot)$ is equivalent to the weak compactness of measures and weak convergence of all the marginal distributions of processes $\xi_n(t)$. Therefore in this subsection conditions for weak compactness of a sequence of measures are studied.

Recall the basic limit theorem for processes without discontinuities of the second kind (Volume I, Chapter VI, Section 5, Theorem 2).

Let $\xi_n(t)$, $t \in [0, T]$, $n = 0, 1, 2, \ldots$ be a sequence of processes with sample functions belonging to $\mathscr{D}$ and let the finite-dimensional distributions of $\xi_n(t)$ converge weakly to finite-dimensional distributions of the process $\xi_0(t)$. Then for weak convergence of measures $q_n(\cdot)$ in $\mathscr{D}$ associated with random processes $\xi_n(t)$ to measure $q_0(\cdot)$, it is necessary and sufficient that

$$\lim_{c \to 0} \varlimsup_{n \to \infty} \mathsf{P}\{\Delta_c(\xi_n(\cdot)) > \varepsilon\} = 0, \tag{2}$$

where

$$\Delta_c(x(\cdot)) = \sup_{t-c\leqslant t'\leqslant t\leqslant t''\leqslant t+c} \{|x(t')-x(t)| \wedge |x(t)-x(t'')|\}$$
$$+ \sup_{0\leqslant t\leqslant c} |x(t)-x(0)| + \sup_{T-c\leqslant t\leqslant T} |x(T)-x(t)|.$$

In view of Volume I, Chapter VI, Section 5, Theorem 3, condition (2) is satisfied provided for some $\beta > 0$ and for $0 \leqslant t_1 < t_2 < t_3 \leqslant T$,

$$\mathsf{E}|\xi_n(t_2)-\xi_n(t_1)|^\beta \, |\xi_n(t_3)-\xi_n(t_2)|^\beta \leqslant H(t_3-t_1)^{1+\alpha},$$

where $\alpha > 0$ and the constant H does not depend on n. We shall require the following refinement of this result.

Assume that

$$\lim_{N\to\infty} \overline{\lim_{n\to\infty}} \, \mathsf{P}\{\sup_{0\leqslant t\leqslant T} |\xi_n(t)| > N\} = 0. \tag{3}$$

Let

$$\tau_n = \inf\{t\colon \sup_{0\leqslant t\leqslant T} |\xi_n(t)| > N\} \qquad (\inf \varnothing = T),$$

and set

$$\xi_n^N(t) = \xi_n(t) \quad \text{for } t < \tau_n \quad \text{and} \quad \xi_n^N(t) = \xi(\tau_n -) \quad \text{for } t \geqslant \tau_n.$$

Then

$$\mathsf{P}\{\Delta_c(\xi_n(\cdot)) > \varepsilon\} \leqslant \mathsf{P}\{\tau_n < T\} + \mathsf{P}\{\Delta_c(\xi_n^N(\cdot)) > \varepsilon\}.$$

Thus the following assertion is valid.

Theorem 1. *If*

a) *a sequence of random processes $\xi_n(t)$ with sample functions in $\mathscr{D}$ satisfies condition* (3),

b) *for some $\beta > 0$ and any $N > 0$*

$$\mathsf{E}|\xi_n^N(t_2)-\xi_n^N(t_1)|^\beta \, |\xi_n^N(t_3)-\xi_n^N(t_2)|^\beta \leqslant H_N(t_3-t_1)^{1+\alpha}, \tag{4}$$

c) *finite-dimensional distributions of processes $\xi_n(t)$ converge weakly to the corresponding distributions of the process $\xi_0(t)$,*

then a sequence of measures $q_n(\cdot)$ in $\mathscr{D}$ associated with the random processes $\xi_n(t)$, $n = 0, 1, \ldots$ is weakly convergent to $q_0(\cdot)$.

Remark. Conditions (2) and (3) are necessary and sufficient for the weak compactness of sequences of measures $q_n(\cdot)$ in $\mathscr{D}$ associated with processes $\xi_n(t)$.

The proof of this assertion is actually contained in theorems presented in Volume I, Chapter VI, Section 5.

We now proceed to a discussion of processes $\xi_n(t)$ constructed from a sequence of series (1). We correspond to them the current of σ-algebras $\{\mathfrak{F}_{nk}, k = 0, 1, \ldots, m_n\}$, $n = 1, 2, \ldots$, where $\mathfrak{F}_{nk}$ is the σ-algebra generated by the random vectors $\xi_{n0}, \xi_{n1}, \ldots, \xi_{nk}$. It is understood here that the variables ξ_{nk} appearing in a given single series are defined on the same probability space, while distinct series are, in general, defined on different probability spaces.

Assume that variables ξ_{nk} possess finite moments of the second order. Set

$$\mathsf{E}\{\Delta\xi_{nk} \mid \mathfrak{F}_{nk}\} = \alpha_{nk}\,\Delta t_{nk},$$
$$\mathsf{E}\{(\Delta\xi_{nk} - \alpha_{nk}\,\Delta t_{nk})(\Delta\xi_{nk} - \alpha_{nk}\,\Delta t_{nk})^* \mid \mathfrak{F}_{nk}\} = \beta_{nk}^2\,\Delta t_{nk}.$$

Here the quantities Δt_{nk} are chosen arbitrarily, subject only to the following restrictions: $\Delta t_{nk} \to 0$; $\sum_{k=1}^{m_n-1} \Delta t_{nk} = T$ (T is fixed and is not a random quantity, $\max_k \Delta t_{nk} \to 0$ as $n \to \infty$). As far as random vectors α_{nk} and matrices β_{nk}^2 are concerned, these are uniquely determined for a chosen sequence of Δt_{nk} by the preceding equalities. Evidently, the matrix β_{nk}^2 is symmetric and nonnegative-definite. Denote by β_{nk} "the nonnegative-definite square root" of matrix β_{nk}^2. This quantity is also a symmetric and nonnegative-definite matrix. In what follows we shall assume that matrices β_{nk}^2 are nonsingular (with probability 1) so that their inverse β_{nk}^{-1} exists.

We represent the variable $\Delta\xi_{nk}$ in the form

$$\Delta\xi_{nk} = \alpha_{nk}\,\Delta t_{nk} + \beta_{nk}\,\Delta\psi_{nk},$$

where

$$\Delta\psi_{nk} = \beta_{nk}^{-1}(\Delta\xi_{nk} - \alpha_{nk}\,\Delta t_{nk})$$

and

$$\psi_{n0} = 0, \qquad \psi_{nk} = \sum_{j=0}^{k-1} \Delta\psi_{nj} = \sum_{j=0}^{k-1} \beta_{nj}^{-1}(\Delta\xi_{nj} - \alpha_{nj}\,\Delta t_{nj}), \qquad k = 1, \ldots, m_n.$$

Set

$$\varphi_{n0} = 0, \qquad \varphi_{nk} = \sum_{j=0}^{k-1} \beta_{nj}\,\Delta\psi_{nj} = \sum_{j=0}^{k-1} (\Delta\xi_{nj} - \alpha_{nj}\,\Delta t_{nj}), \qquad k = 1, \ldots, m_n.$$

Sequences $\{\psi_{nk}, k = 0, 1, \ldots, m_n\}$, $\{\varphi_{nk}, k = 0, 1, \ldots, m_n\}$ are $\mathfrak{F}_{nk}$-martingales. Moreover,

$$\mathsf{E}\{\Delta\psi_{nk}\,\Delta\psi_{nk}^* \mid \mathfrak{F}_{nk}\} = \boldsymbol{I}\,\Delta t_{nk}, \qquad \mathsf{E}\{\Delta\varphi_{nk}\,\Delta\varphi_{nk}^* \mid \mathfrak{F}_{nk}\} = \beta_{nk}^2\,\Delta t_{nk}.$$

Since α_{nk} and β_{nk} are $\mathfrak{F}_{nk}=\sigma(\xi_{n0}, \xi_{n1}, \dots, \xi_{nk})$-measurable, there exist nonrandom Borel functions $a_{nk}(x_0, x_1, \dots, x_k)$, $b_{nk}(x_0, x_1, \dots, x_k)$, $x_j \in \mathcal{R}^m$, $j=0, 1, \dots, k$, $k=1, \dots, m_n$, such that

$$\alpha_{nk}=a_{nk}(\xi_{n0}, \xi_{n1}, \dots, \xi_{nk}), \qquad \beta_{nk}=b_{nk}(\xi_{n0}, \xi_{n1}, \dots, \xi_{nk}).$$

Here the functions $a_{nk}(x_0, \dots, x_k)$ take on values in $\mathcal{R}^m$, while $b_{nk}(x_0, \dots, x_k)$ are matrix-valued functions.

Lemma 1. *Assume that functions $a_{nk}(x_0, \dots, x_k)$ and $b_{nk}(x_0, \dots, x_k)$ satisfy the condition*

$$|a_{nk}(x_0, \dots, x_k)|+|b_{nk}(x_0, \dots, x_k)| \leqslant C\big(1+\sup_{0\leqslant j\leqslant k} |x_j|\big), \tag{5}$$

where C is a constant independent of n. Then there are constants C_1 and C_2 which also do not depend on n such that

$$\mathsf{E}\{\sup_{0\leqslant j\leqslant k} |\xi_{nj}|^2 \,|\, \mathfrak{F}_{n0}\} \leqslant C_1(1+|\xi_{n0}|^2), \tag{6}$$

$$\mathsf{E}\{\sup_{s\leqslant j\leqslant r} |\xi_{nj}-\xi_{ns}|^2 \,|\, \mathfrak{F}_{ns}\} \leqslant C_2(1+|\xi_{ns}|^2)(t_{nr}-t_{ns}). \tag{7}$$

Proof. Since

$$\xi_{nk+1}=\xi_{n0}+\sum_{j=0}^{k} \alpha_{nj}\,\Delta t_{nj}+\sum_{j=0}^{k} \beta_{nj}\,\Delta\psi_{nj},\dagger$$

it follows that

$$\begin{aligned}
&\sup_{0\leqslant j\leqslant k+1} |\xi_{nj}|^2\\
&\leqslant 3\left[|\xi_{n0}|^2+\sup_{0\leqslant j\leqslant k}\left|\sum_{r=0}^{j} \alpha_{nr}\,\Delta t_{nr}\right|^2+\sup_{0\leqslant j\leqslant k}\left|\sum_{r=0}^{j} \beta_{nr}\,\Delta\psi_{nr}\right|^2\right]\\
&\leqslant 3\left[|\xi_{n0}|^2+t_{nk}\sum_{r=0}^{k} |\alpha_{nr}|^2\,\Delta t_{nr}+\sup_{j\leqslant k}\left|\sum_{r=0}^{j} \beta_{nr}\,\Delta\psi_{nr}\right|^2\right].
\end{aligned}$$

Set $v_{nk}=\mathsf{E}\{\sup_{0\leqslant j\leqslant k} |\xi_{nj}|^2 \,|\, \mathfrak{F}_{n0}\}$. The preceding inequality implies that

$$v_{nk+1}\leqslant 3\left[|\xi_{n0}|^2+2TC^2\sum_{r=0}^{k}(1+v_{nr})\,\Delta t_{nr}+\mathsf{E}\left\{\sup_{j\leqslant k}\left|\sum_{r=0}^{j} \beta_{nr}\,\Delta\psi_{nr}\right|^2 \,\middle|\, \mathfrak{F}_{n0}\right\}\right].$$

† In this equation (and in a number of succeeding expressions throughout this chapter) notation $\xi_{nk+1}(v_{nk+1}, z_{nz+1}, t_{nk+1}$ etc. should be interpreted as $\xi_{n\,k+1}(v_{n\,k+1}, z_{n\,z+1}, t_{n\,k+1}$ etc. respectively).

Noting that the sums $\sum_{r=0}^{j} \beta_{nr} \Delta\psi_{nr}$ form a martingale and utilizing Doob's inequality we obtain

$$\mathsf{E}\left\{ \sup_{j\leq k} \left| \sum_{r=0}^{j} \beta_{nr} \Delta\psi_{nr} \right|^2 \,\middle|\, \mathfrak{F}_{n0} \right\} \leq 4\mathsf{E}\left\{ \left| \sum_{r=0}^{k} \beta_{nr} \Delta\psi_{nr} \right|^2 \,\middle|\, \mathfrak{F}_{n0} \right\}$$

$$= 4\mathsf{E}\left\{ \sum_{r=0}^{k} |\beta_{nr} \Delta\psi_{nr}|^2 \,|\, \mathfrak{F}_{n0} \right\}$$

$$= 4\mathsf{E}\left\{ \sum_{r=0}^{k} \operatorname{sp} \beta_{nr}^2 \Delta t_{nr} \,|\, \mathfrak{F}_{n0} \right\}.$$

Thus

$$v_{nk+1} \leq 3|\xi_{n0}|^2 + C'\left(t_{nk} + \sum_{r=0}^{k} v_{nr} \Delta t_{nr} \right),$$

where C' is a constant which depends on T only.

We now introduce a piecewise constant function $v_n(t)$ by setting $v_n(t) = v_{nk}$ for $t \in [t_{nk}, t_{nk+1})$. The last inequality implies that

$$v_n(t) \leq 3|\xi_{n0}|^2 + C' \int_0^t (1 + v_n(s))\, ds.$$

Utilizing Lemma 1 in Section 1 we obtain

$$v_n(t) \leq (3|\xi_{n0}|^2 + C't)\, e^{C't}.$$

From here relation (6) follows.

To prove inequality (7) we proceed analogously. Inequality

$$\xi_{nk+1} - \xi_{ns} = \sum_{j=s}^{k} \alpha_{nj} \Delta t_{nj} + \sum_{j=s}^{k} \beta_{nj} \Delta\psi_{nj}$$

implies

$$\sup_{s\leq j\leq k+1} |\xi_{nj} - \xi_{ns}|^2 \leq 2\left[\sup_{s\leq j\leq k} \left| \sum_{r=s}^{j} \alpha_{nr} \Delta t_{nr} \right|^2 + \sup_{s\leq j\leq k} \left| \sum_{r=s}^{j} \beta_{nr} \Delta\psi_{nr} \right|^2 \right].$$

Set

$$z_{nr} = \mathsf{E}\{ \sup_{s\leq j\leq r} |\xi_{nj} - \xi_{ns}|^2 \,|\, \mathfrak{F}_{ns} \}.$$

From the preceding relation one easily obtains

$$z_{nr+1} \leq 4TC^2 \sum_{j=s}^{r} \mathsf{E}\{1 + \sup_{s\leq k\leq j} |\xi_{nk}|^2 \,|\, \mathfrak{F}_{ns}\} \Delta t_{nj} + 8\mathsf{E}\left\{ \sum_{j=s}^{r} |\beta_{nj} \Delta\psi_{nj}|^2 \,|\, \mathfrak{F}_{ns} \right\}.$$

This yields the following inequality for the expectations z_{nr},

$$z_{nr+1} \leqslant C'' \sum_{k=s}^{r} (1+v'_{nk})\,\Delta t_{nk};$$

here C'' is a constant which depends on C only and $v'_{nk} = \mathsf{E}\{\sup_{s\leqslant j\leqslant k} |\xi_{nj}|^2 \,|\, \mathfrak{F}_{ns}\}$. Variables v'_{nr} can be bounded using inequality (6), which implies that $v_{nk} \leqslant C_1(1+|\xi_{ns}|^2)$. The second assertion of the lemma thus follows from the bounds obtained. □

Theorem 2. *If a sequence of series* (1) *satisfies the condition*

$$(8)\qquad \begin{gathered} |a_{nk}(x_0, x_1, \dots, x_k)| + |b_{nk}(x_0, x_1, \dots, x_k)| \leqslant C(1+\sup_{0\leqslant j\leqslant k} |x_j|), \\ n = 1, 2, \dots, \qquad k = 0, 1, \dots, m_n, \end{gathered}$$

where C is a constant independent of n and $\sup_n \mathsf{E}|\xi_{n0}|^2 < \infty$, *then the sequence of measures $q_n(\,\cdot\,)$ in $\mathscr{D}$ is weakly compact.*

Proof. Theorem 2 is a corollary to the remark following Theorem 1 and Lemma 1. Indeed, Lemma 1 and Chebyshev's inequality imply that

$$\mathsf{P}\{\sup_{0\leqslant t\leqslant T} |\xi_n(t)| > N\} \leqslant \frac{C\mathsf{E}(1+|\xi_{n0}|^2)}{N^2};$$

thus condition (3) of Theorem 1 is fulfilled in our case. Next let $\mathfrak{F}_n(t) = \mathfrak{F}_{nk}$ for $t \in [t_{nk}, t_{nk+1})$. Then

$$\mathsf{E}\{|\xi_n^N(t_3) - \xi_n^N(t_2)|^2 | \mathfrak{F}_n(t_2)\} \leqslant \chi_N(t_2)\mathsf{E}\{\sup_{t_2\leqslant t\leqslant t_3} |\xi_n(t) - \xi_n(t_2)|^2 | \mathfrak{F}_n(t_2)\},$$

where $\chi_N(t)$ is the indicator of the event $\{\tau_n > t\}$. Utilizing Lemma 1 once again we obtain

$$\mathsf{E}\{\sup_{t_2\leqslant t\leqslant t_3} |\xi_n^N(t) - \xi_n^N(t_2)|^2 | \mathfrak{F}_n(t_2)\} \leqslant \chi_N(t_2) C_2 (1+|\xi_n^N(t_2)|^2)(t_3 - t_2).$$

Finally, we have

$$\begin{aligned} &\mathsf{E}|\xi_n^N(t_3) - \xi_n^N(t_2)|^2 \, |\xi_n^N(t_2) - \xi_n^N(t_1)|^2 \\ &\quad\leqslant \mathsf{E}(\mathsf{E}\{|\xi_n^N(t_3) - \xi_n^N(t_2)|^2 \,|\, \mathfrak{F}_n(t_2)\} |\xi_n^N(t_2) - \xi_n^N(t_1)|^2) \\ &\quad\leqslant C_2^2 (1+N^2) \mathsf{E}(1+|\xi_{n0}|^2)(t_3 - t_1)^2. \end{aligned}$$

Thus, the conditions of Theorem 1 are fulfilled and therefore Theorem 2 is proved. □

We note yet the following application of Theorem 1 to a sequence of series (1) which are square integrable martingales.

Let $\mathsf{E}\{\Delta\xi_{nk}|\mathfrak{F}_{nk}\}=0$. Set

$$\mathsf{E}\{|\Delta\xi_{nk}|^2\,|\,\mathfrak{F}_{nk}\}=\gamma_{nk}\,\Delta t_{nk},\qquad k=0,\dots,m_n-1, \tag{9}$$

and let $\rho_n=\inf\{r;\,\gamma_{nr}\geq N\}$ $(\inf\varnothing=m_n)$. Then ρ_n is a random time on $\{\mathfrak{F}_{nr},\,r=0,\dots,m_n\}$. Let $\xi_n^N(t)=\xi_n(t\wedge t_{n\rho_n})$. The process $\xi_n^N(t)$ is also a martingale. Moreover,

$$\mathsf{P}\{\Delta_c(\xi_n(\cdot))>\varepsilon\}\leq\mathsf{P}\{\rho_n<m_n\}+\mathsf{P}\{\Delta_c(\xi_n^N(\cdot))>\varepsilon\}.$$

Furthermore, if the variables t_2 and t_3 are of the form $t_2=t_{nj}$, $t_3=t_{nr}$, then

$$\mathsf{E}\{|\xi_n^N(t_3)-\xi_n^N(t_2)|^2\,|\,\mathfrak{F}_n(t_2)\}\leq\sum_{k=j\wedge\rho_n}^{n\wedge\rho_n}\gamma_{nk}\,\Delta t_{nk}\leq N(t_3-t_2).$$

Thus (for $t_1<t_2<t_3$)

$$\mathsf{E}|\xi_n^N(t_3)-\xi_n^N(t_2)|^2|\xi_n^N(t_2)-\xi_n^N(t_1)|^2\leq N^2(t_3-t_1)^2.$$

The additional assumption that t_1, t_2 and t_3 are of the form t_{ni}, t_{nj}, and t_{nr} respectively is inessential and we arrive at the following theorem.

Theorem 3. *If each series in the sequence of the series is a martingale and if*

$$\lim_{N\to\infty}\overline{\lim_{n\to\infty}}\,\mathsf{P}\{\sup_{0\leq r\leq m_n-1}\gamma_{nr}>N\}=0,$$

where the variables γ_{nr} are defined by relation (9), *then the sequences of measures in $\mathcal{D}$ associated with processes $\xi_n(\cdot)$ is weakly compact.*

Corollary. *A sequence of measures in $\mathcal{D}$ associated with processes $\psi_n(t)$ such that $\mathsf{E}\{\Delta\psi_{nk}\,|\,\mathfrak{F}_{nk}\}=0$ and $\mathsf{E}\{\Delta\psi_{nk}\cdot\Delta\psi_{nk}^*\,|\,\mathfrak{F}_{nk}\}=\boldsymbol{I}\,\Delta t_{nk}$ is weakly compact.*

Theorem 2 can easily be generalized to the case of sequences of series of random vectors without finite moments of the second order. For this purpose we introduce on the current of σ-algebras $\{\mathfrak{F}_{nk},\,k=1,\dots,m_n\}$ a random time j_n by setting $j_n=\min\{k\colon|\xi_{nk}|>N\}$ (or $j_n=m_n+1$ if the set $\{k\colon|\xi_{nk}|>N\}$ is void). Consider now for each $N>0$ a sequence of the series $\{\xi_{nk}^N,\,k=0,\dots,m_n\}$, $n=1,2,\dots$, where $\xi_{nk}^N=\xi_{nk}$ for $k<j_n$ and $\xi_{nk}^N=\xi_{nj_n-1}^N$ for $k\geq j_n$. Observe that vectors ξ_{nk}^N possess moments of all orders. Let $a_{nk}^N(x_0,x_1,\dots,x_k)$ and $b_{nk}^N(x_0,x_1,\dots,x_k)$ be constructed from $\{\xi_{nk}^N,\ k=0,1,\dots,m_n\}$ in the same manner as $a_{nk}(x_0,\dots,x_k)$ and $b_{nk}(x_0,\dots,x_k)$ were constructed from the sequence $\{\xi_{nk},\,k=0,1,\dots,m_n\}$.

Theorem 4. *If*

$$\lim_{N\to\infty}\overline{\lim_{n\to\infty}}\,\mathsf{P}\{\max_{0\le k\le m_n}|\xi_{nk}|>N\}=0$$

and

$$|a_{nk}^N(x_0,x_1,\dots,x_k)|+|b_{nk}^N(x_0,\dots,x_k)|\le C^N(1+\max_{0\le j\le k}|x_j|),$$

where C^N is a constant dependent possibly on N but independent of n and k, then the sequence of measures $q_n(\cdot)$ associated with processes $\xi_n(t)$ is weakly compact in $\mathcal{D}$.

An assertion analogous to Theorem 2 holds also for measures associated with solutions of stochastic differential equations. Consider the family of equations

$$\begin{aligned} d\xi_\alpha&=A_\alpha(\theta_t\xi_\alpha,dt), && t\ge 0,\\ \xi_\alpha(t)&=\varphi(t), && t\le 0, \end{aligned}\tag{10}$$

depending on the parameter α.

Theorem 5. *Let $A_\alpha(\varphi,t)\in S(\lambda_0^\alpha,\lambda_N^\alpha)$ (or $A_\alpha(\varphi,t)\in S^c(\lambda_0^\alpha,\lambda_N^\alpha)$ and*

$$\lim_{N\to 0}\sup_\alpha\mathsf{P}\{\sup_{0\le t\le T}|\lambda_0^\alpha(t)|>N\}=0.$$

Then the family of measures $q_\alpha(\cdot)$ in $\mathcal{D}[0,T]$ associated with solutions $\xi_\alpha(t)$ of equations (10) *is weakly compact.*

The proof is analogous to the proof of preceding theorems. It is based on Theorem 1 and utilizes Theorem 7 (and Theorem 4) in Section 2 instead of Lemma 1.

Conditions for convergence to a Wiener process. We now proceed to study conditions for convergence of a sequence of processes $\{\xi_n(t),t\in[0,T]\}$, $n=1,2,\dots$, constructed from a sequence of series of random vectors (1) to a Wiener process.

For an arbitrary $\varepsilon>0$ set

$$\mathsf{P}\{|\Delta\xi_{nk}|\ge\varepsilon\mid\mathfrak{F}_{nk}\}=\rho'_{nk}\,\Delta t_{nk},\tag{11}$$

$$\mathsf{E}\{\chi_{nk},\Delta\xi_{nk}^*\mid\mathfrak{F}_{nk}\}=\rho''_{nk}\,\Delta t_{nk},\tag{12}$$

$$\mathsf{E}\{\chi_{nk}\,\Delta\xi_{nk}\,\Delta\xi_{nk}^*\mid\mathfrak{F}_{nk}\}=(I+\rho'''_{nk})\,\Delta t_{nk}.\tag{13}$$

Here

$$\Delta t_{nk}=t_{nk+1}-t_{nk},\qquad 0=t_{n0}<t_{n1}<\cdots<t_{nm_0}=T,$$

and the numbers T and t_{nk} are arbitrary subject only to the condition

$$\max_{0\leq k\leq m_n-1} \Delta t_{nk}\to 0 \quad \text{as } n\to\infty$$

and the fulfillment of the subsequent assumptions: ρ'_{nk} are scalar, ρ''_{nk} are vector, and ρ'''_{nk} are matrix random variables defined by the corresponding equations (11)–(13). Finally, $\chi_{nk}=\chi_\varepsilon(\Delta\xi_{nk})=1$ if $|\Delta\xi_{nk}|<\varepsilon$ and $\chi_{nk}=0$ if $|\Delta\xi_{nk}|\geq\varepsilon$.

For an m-dimensional Wiener process $\{w(t), \mathfrak{F}_t, t\geq 0\}$ conditional probabilities and expectations (11)–(13) coincide with the unconditional ones and are of order

$$\mathsf{P}\{|\Delta w|\geq\varepsilon\}=O\left[\left(\frac{\varepsilon^2}{\Delta t}\right)^{(m-2)/2} e^{-\varepsilon^2/2\Delta t}\right],$$

$$|\mathsf{E}\{\chi_\varepsilon(\Delta w)\,\Delta w\}|=O\left(\left(\frac{\varepsilon^2}{\Delta t}\right)^{(m-1)/2} e^{-\varepsilon^2/2\Delta t}\right),$$

$$\mathsf{E}\{\chi_\varepsilon(\Delta w)|\Delta w|^2\}=\boldsymbol{I}\Delta t+O\left(\left(\frac{\varepsilon^2}{\Delta t}\right)^{m/2} e^{-\varepsilon^2/2\Delta t}\right).$$

One would expect that if ρ'_{nk}, ρ''_{nk}, and ρ'''_{nk} for an arbitrary fixed ε are "sufficiently" small then the marginal distributions of the process $\xi_n(t)-\xi_n(0)$ should weakly converge to a Wiener process as $n\to\infty$.

First we shall establish conditions for convergence of the distribution of the variable $\xi_n(T)-\xi_n(0)=\xi_{nm_n}-\xi_{n0}$ to the distribution of $w(T)$. For this purpose, consider the difference between conditional characteristic functions

$$\sigma_n=\mathsf{E}\{\exp\{i(\xi_n(T)-\xi_n(0), z)\}|\mathfrak{F}_{n0}\}-\mathsf{E}\{\exp\{i(w(T), z)\}|\mathfrak{F}_0\}$$

$$=\mathsf{E}\left\{\exp\{i(\xi_n(T)-\xi_n(0), z)\}-\exp\left\{-\frac{|z|^2T}{2}\right\}\,\Big|\,\mathfrak{F}_{n0}\right\}.$$

Set

$$\chi^{n0}=1, \qquad \chi^{nk}=\prod_{j=0}^{k-1}\chi_{nj}.$$

Represent σ_n in the form:

$$\sigma_n=\sum_{k=0}^{m_n-1}\mathsf{E}\{\sigma_{nk}|\mathfrak{F}_{n0}\}+\mathsf{E}\{(1-\chi^{nm_n})\exp\{i(\xi_n(T)-\xi_n(0), z)\}|\mathfrak{F}_{n0}\},$$

where

$$\sigma_{nk}=\chi^{nk+1}\exp\{i(\xi_n(t_{nk+1})-\xi_n(0),z)\}\exp\left\{-\frac{|z|^2(T-t_{nk+1})}{2}\right\}$$
$$-\chi^{nk}\exp\{i(\xi_n(t_{nk})-\xi_n(0),z)\}\exp\left\{-\frac{|z|^2(T-t_{nk})}{2}\right\}$$
$$=\chi^{nk}\exp\{i(\xi_n(t_{nk})-\xi_n(0),z)\}\exp\left\{-\frac{|z|^2(T-t_{nk+1})}{2}\right\}\tilde{\sigma}_{nk},$$
$$\tilde{\sigma}_{nk}=\chi_{nk}\exp\{i(\Delta\xi_{nk},z)\}-\exp\left\{-\frac{|z|^2\,\Delta t_{nk}}{2}\right\}.$$

We now bound the quantity $\gamma_{nk}=\mathsf{E}\{\tilde{\sigma}_{nk}\,|\,\mathfrak{F}_{nk}\}$. Utilizing Taylor's formula we represent γ_{nk} as

$$\gamma_{nk}=\delta_{nk}^{(1)}+\delta_{nk}^{(2)}+\delta_{nk}^{(3)}+\delta_{nk}^{(4)}+\delta_{nk}^{(5)},$$

where

$$\delta_{nk}^{(1)}=\mathsf{E}\{\chi_{nk}\,|\,\mathfrak{F}_{nk}\}-1,$$
$$\delta_{nk}^{(2)}=i\mathsf{E}\{\chi_{nk}(\Delta\xi_{nk},z)\,|\,\mathfrak{F}_{nk}\},$$
$$\delta_{nk}^{(3)}=\left(\frac{|z|^2}{2}\Delta t_{nk}-\tfrac{1}{2}\mathsf{E}\{\chi_{nk}(\Delta\xi_{nk},z)^2\,|\,\mathfrak{F}_{nk}\}\right),$$
$$\delta_{nk}^{(4)}=\frac{|z|^2\,\Delta t_{nk}}{2}\left(\exp\left\{-\frac{|z|^2\,\Delta t_{nk}}{2}\right\}-1\right),$$
$$\delta_{nk}^{(5)}=\tfrac{1}{2}\mathsf{E}\{\chi_{nk}(\Delta\xi_{nk},z)^2(\exp\{i\theta(\Delta\xi_{nk},z)\}-1)\,|\,\mathfrak{F}_{nk}\}.$$

Clearly,

$$|\delta_{nk}^{(1)}|=\rho'_{nk}\,\Delta t_{nk},\qquad |\delta_{nk}^{(2)}|\leqslant|z|\,|\rho''_{nk}|\,\Delta t_{nk},$$
$$|\delta_{nk}^{(3)}|\leqslant\tfrac{1}{2}|z|^2|\rho''_{nk}|\,\Delta t_{nk},\qquad |\delta_{nk}^{(4)}|\leqslant|\Delta t_{nk}|^2|z|^2,$$
$$|\delta_{nk}^{(5)}|\leqslant\frac{|z|^3}{2}|I+\rho'''_{nk}|\varepsilon\,\Delta t_{nk}.$$

Here $|A|$ denotes the operator norm of matrix A. Observe that

$$|\mathsf{E}\{(1-\chi^{nm_n})\exp\{i(\xi_n(T)-\xi_n(0),z)\}\,|\,\mathfrak{F}_{n0}\}$$
$$\leqslant\mathsf{P}\{\chi^{nm_n}=0\,|\,\mathfrak{F}_{n0}\}\leqslant\mathsf{E}\left\{\sum_{k=1}^{m_n}\mathsf{P}(|\Delta\xi_{nk}|\geqslant\varepsilon\,|\,\mathfrak{F}_{nk})\,|\,\mathfrak{F}_{n0}\right\}$$
$$=\mathsf{E}\left\{\sum_{k=1}^{m_n}\rho'_{nk}\,\Delta t_{nk}\,|\,\mathfrak{F}_{n0}\right\}.$$

Thus for an arbitrary $\varepsilon > 0$ and $\max_k \Delta t_{nk}$ sufficiently small,

$$\sigma_n \leqslant C(z)\left[\mathsf{E}\left\{\sum_{k=0}^{m_n} (\rho'_{nk} + |\rho''_{nk}| + |\rho'''_{nk}|)\,\Delta t_{nk}\,|\,\mathfrak{F}_{n0}\right\} + \varepsilon T\right]. \tag{14}$$

We arrive at the following result:

Theorem 6. *If a sequence of series* (1) *is such that*

$$\mathsf{E}\left\{\sum_{k=0}^{m_n} (\rho'_{nk} + |\rho''_{nk}| + |\rho'''_{nk}|\,\Delta t_{nk})\right\} \to 0 \tag{15}$$

and $\max_k \Delta t_{nk} \to 0$, *where* ρ'_{nk}, ρ''_{nk}, *and* ρ'''_{nk} *are defined by equations* (11)–(13), *then the conditional distribution of the variable* $\xi_{nm_n} - \xi_{n0}$ *converges to the Gaussian distribution with mean* $\mathbf{0}$ *and covariance matrix* TI.

Remark 1. Modify the conditions of Theorem 6 as follows: set $t_{nm_n} = T_n$ and

$$\mathsf{E}\{\chi_{nk}\,\Delta\xi_{nk}\,\Delta\xi^*_{nk}\,|\,\mathfrak{F}_{nk}\} = (B + \rho'''_{nk})\,\Delta t_{nk},$$

where B is a constant matrix and $T_n \to T$ as $n \to \infty$. If we retain all the other conditions, then the distribution of the difference $\xi_{nm_n} - \xi_{n0}$ converges to the Gaussian distribution with mean $\mathbf{0}$ and covariance matrix TB.

Remark 2. If the assumptions of Theorem 6 are fulfilled and the distribution of the variable ξ_{n0} is weakly convergent to a measure $F(\cdot)$ on $\mathfrak{B}^m$, then the distribution of the variable ξ_{nm_n} is weakly convergent to the distribution with the density

$$\int_{\mathfrak{R}^m} \frac{1}{\sqrt{2\pi}\,T}\, e^{-|x-y|^2/2T} F(dy).$$

Theorem 6 can be easily generalized.

Theorem 7. *Let the conditions of Theorem 6 be satisfied and let* $t_{nk_j} \to t_j$ *as* $n \to \infty$, $j = 1, 2, \ldots, r$, $0 \leqslant t_1 < t_2 < \cdots < t_r < T$. *Then the joint distribution of the random vectors*

$$\xi_{nk_1} - \xi_{n0},\ \xi_{nk_2} - \xi_{nk_1}, \ldots, \xi_{nk_r} - \xi_{nk_{r-1}}$$

converges weakly to the joint distribution of the sequence

$$w(t_1) - w(0),\ w(t_2) - w(t_1), \ldots, w(t_r) - w(t_{r-1}),$$

where $w(t)$ *is an m-dimensional Wiener process.*

Proof. To prove this assertion consider the difference

$$\sigma_n = \mathsf{E}\left\{\exp\left\{i\sum_{j=0}^{r-1}(\xi_n(t_{nk_{j+1}})-\xi_n(t_{nk}),\, z_j)\right\}\right.$$
$$\left.-\exp\left\{-\frac{1}{2}\sum_{j=0}^{r-1}|z_j|^2(t_{j+1}-t_j)\right\}\,\middle|\,\mathfrak{F}_{n0}\right\},$$

where z_j, $j=0,\ldots,r$ are arbitrary vectors in $\mathscr{R}^m$, and express it as

$$\sigma_n = \sum_{k=0}^{r-1}\mathsf{E}\left\{\exp\left\{i\sum_{j=0}^{k}(\xi_n(t_{nk_{j+1}})-\xi_n(t_{nk_j}),\, z_j)-\frac{1}{2}\sum_{j=k+1}^{r-1}|z_j|^2(t_{j+1}-t_j)\right\}\right.$$
$$\left.-\exp\left\{i\sum_{j=0}^{k-1}(\xi_n(t_{nk_{j+1}})-\xi_n(t_{nk_j}),\, z)-\frac{1}{2}\sum_{j=k}^{r-1}|z_j|^2(t_{j+1}-t_j)\right\}\,\middle|\,\mathfrak{F}_{n0}\right\}$$
$$=\sum_{k=0}^{r-1}\mathsf{E}\{\mathsf{E}(\sigma_{nk}\,|\,\mathfrak{F}_{nk})\,|\,\mathfrak{F}_{n0}\}.$$

The quantity $\mathsf{E}(\sigma_{nk}\,|\,\mathfrak{F}_{nk})$ appearing in the last equality can be bounded as above using inequality (14) (with obvious modifications). □

Corollary. *If the sequence $\{\xi_{nk}, \mathfrak{F}_{nk}, k=1,\ldots,m_n\}$ is a martingale possessing finite moments of the second order and if, moreover,*

$$\mathsf{E}\{\Delta\xi_{nk}\,\Delta\xi_{nk}^*\,|\,\mathfrak{F}_{nk}\} = \boldsymbol{I}\Delta t_{nk} \tag{16}$$

and

$$\mathsf{E}\sum_{k=1}^{m_n-1}(1-\chi_{nk})|\,\Delta\xi_{nk}|^2 \to 0 \quad \text{as } n\to\infty, \tag{17}$$

then the sequence of measures $q(\cdot)$ in $\mathscr{D}$ associated with the processes $\xi_n(t)-\xi_n(0)$ converges weakly to a Wiener measure.

Condition (17) is the classical Lindeberg condition in the central limit theorem for a sum of independent random variables. It is easy to verify using Chebyshev's inequality that (17) implies (15).

On the other hand, in the case under consideration one can apply the corollary to Theorem 3 which implies that a family of measures associated with processes $\xi_n(t)$ constructed from the sequence of series $\{\xi_{nk}, k=1,\ldots,m_k\}$, $n=1,2,\ldots$, is weakly compact.

It is of interest to generalize the preceding theorem by allowing the times t_{nk} to be chosen in a random manner. In this case, one should firstly require that the choice of variables $\Delta t_{nk}=t_{nk+1}-t_{nk}$ could not anticipate "the future," i.e., that the variables Δt_{nk} be $\mathfrak{F}_{nk}$-measurable. Under certain additional assumptions to be presented immediately below, the calculations and bounds utilized in the proof of Theorem 6 will be only slightly altered.

In this manner we obtain the following assertion.

Theorem 8. *Assume that*

a) *times Δt_{nk} are $\mathfrak{F}_{nk}$-measurable random variables, $k=1,2,\dots,m_n-1$,*
b) $\mathsf{E}|t_{nm_n}-T|\to 0$ *as $n\to\infty$, where T is not a random quantity,*
c) *conditions* (16) *and* (17) *are satisfied.*

Then the distribution of the vector $\xi_{nm_n}-\xi_{n0}$ converges weakly to the distribution of $w(T)$ where $w(t)$ is an m-dimensional Wiener process. If, moreover, $t_{nm_n}=T$, i.e., t_{nm_n} is not a random quantity, then the joint distribution of the variables

$$\xi_n(t_{nk_1})-\xi_n(t_{n0}),\dots,\xi_n(t_{nk_r})-\xi_n(t_{nk_{r-1}})$$

is weakly convergent to the distribution of

$$w(t_j)-w(0),\dots,w(t_r)-w(t_{r-1})$$

as $t_{nk_j}\to t_j$, $j=1,2,\dots,r$, in probability (where t_j is not a random quantity). Moreover, the measures in $\mathcal{D}$ associated with processes $\xi_n(t)$ converge weakly to a Wiener measure.

Proof. To prove this assertion we return to Theorem 6 and examine the modifications which are needed in its proof in order that it will be applicable for the case under consideration. Since now $t_{nm_n}\neq T$ in general, the expression for σ_N will contain an additional summand of the form

$$\mathsf{E}\Big\{\chi^{nm_n}\exp\{i(\xi_n(t_{nm_n})-\xi_n(0),z)\}\Big(\exp\Big\{-\frac{|z|^2}{2}(T-t_{nm_n})\Big\}-1\Big)\Big|\,\mathfrak{F}_{n0}\Big\},$$

which tends to zero in probability.

A bound on the sum $\sum_{k=1}^{m}\delta_{nk}^{(4)}$ will also require a different approach. In the present case it can be bounded by means of the inequality

$$\mathsf{E}\Big\{\sum_{k=1}^{m_n}\delta_{nk}^{(4)}\Big|\,\mathfrak{F}_{n0}\Big\}\leqslant\frac{|z|^4}{4}\varepsilon\mathsf{E}\{t_{nm_n}\,|\,\mathfrak{F}_{n0}\}+|z|^2\mathsf{E}\Big\{\sum_{k=1}^{m_n}\Delta t_{nk}(1-\chi_\varepsilon(\Delta t_{nk}))\Big|\,\mathfrak{F}_{n0}\Big\}.$$

Since the function $|t|(1-\chi_\varepsilon(|t|))$ is convex downward and $\Delta t_{nk}=(1/m)\mathsf{E}\{|\Delta\xi_{nk}|^2\,|\,\mathfrak{F}_{nk}\}$, it follows that

$$\Delta t_{nk}(1-\chi_\varepsilon(\Delta t_{nk}))\leqslant\mathsf{E}\,\frac{|\Delta\xi_{nk}|^2}{m}\Big(1-\chi_\varepsilon\Big(\frac{|\Delta\xi_{nk}|^2}{m}\Big)\Big).$$

Therefore condition (17) implies that for any $\varepsilon>0$

$$\lim_{n\to\infty}\mathsf{E}\sum_{k=1}^{m_n}\Delta t_{nk}(1-\chi_\varepsilon(\Delta t_{nk}))=0.$$

Taking into account condition a) of the theorem it is not difficult to verify that all the other transformations and inequalities used in the proof of Theorem 6 are applicable in the case under consideration. □

Let $\{\xi_n, \mathfrak{F}_n, n=0, 1, \ldots\}$ be a martingale with finite second-order moments. Set

$$\gamma_n^2 = \mathsf{E}\{(\xi_n - \xi_{n-1})^2 \mid \mathfrak{F}_{n-1}\}$$

and assume that there exists an $\mathfrak{F}_0$-measurable function $\varphi(n)$ such that

$$\frac{1}{\varphi(n)} \sum_{k=1}^{n} \gamma_k^2 \to 1 \tag{18}$$

in the sense of convergence in L_1.

Set

$$\xi_{nk} = \frac{1}{\sqrt{\varphi(n)}} \xi_k, \qquad k = 0, 1, \ldots, n, \qquad \Delta t_{nk} = \frac{\gamma_{k+1}^2}{\varphi(n)}.$$

Then the variables $t_{nk} = \Delta t_{n0} + \cdots + \Delta t_{nk-1}$ are $\mathfrak{F}_{k-1}$-measurable and

$$\mathsf{E}\{\xi_{nk+1} - \xi_{nk} \mid \mathfrak{F}_k\} = 0,$$

$$\mathsf{E}\{(\xi_{nk+1} - \xi_k)^2 \mid \mathfrak{F}_k\} = \frac{\gamma_{k+1}^2}{\varphi(n)} = \Delta t_{nk}.$$

Theorem 8 is thus applicable. Hence we arrive at the following assertion.

Theorem 9. *If the martingale* $\{\xi_n, \mathfrak{F}_n, n = 1, 2, \ldots\}$ *satisfies condition* (18) (*in the sense of convergence in* L_1) *and if for any* $\varepsilon > 0$

$$\frac{1}{\varphi(n)} \sum_{k=1}^{n} \int_{\{(\xi_k - \xi_{k-1})^2 > \varepsilon^2 \varphi(n)\}} (\xi_k - \xi_{k-1})^2 \, d\mathsf{P} \to 0$$

with probability 1 *as* $n \to \infty$, *then the conditional distribution of the variable* $(1/\sqrt{\varphi(n)})(\xi_n - \xi_0)$ *is asymptotically normal* (0, 1).

Remark. The random process $\xi_n(t)$ constructed for the sequence $\xi_{n0}, \ldots, \xi_{nk}$ considered in Theorem 9 cuts off at the random time t_{nn}. Since $\varphi(n) \to \infty$ with probability 1 we can extend the construction of the process $\xi_n(t)$ by means of variables ξ_{nk} for $k > n$ in order that it be defined on a fixed time interval, e.g., [0, 1]. It then follows from Theorem 8 that measures in $\mathscr{D}[0, 1]$ associated with processes $\xi_n(t)$ converge weakly to the Wiener measure.

Conditions for convergence to an arbitrary process with independent increments. Recall first that if ζ_h is a family of random vectors, $\zeta_h \to 0$ as $h \to 0$ and if the limit of

$(1/\Delta(h))\mathsf{E}(e^{i(\zeta_h, z)}-1)$ exists then it is of the form (cf. Volume I, Chapter III)

$$\lim_{h\to 0}\frac{1}{\Delta(h)}\mathsf{E}(e^{i(\zeta_h,z)}-1)$$
$$=i(a, z)-\tfrac{1}{2}(bz, z)+\int_{\mathscr{R}^m}\left(e^{i(z,u)}-1-\frac{i(u, z)}{1+|z|^2}\right)\frac{1+|u|^2}{|u|^2}\Pi(du),$$

where $\Pi(\cdot)$ is a finite measure continuous at point 0. The parameter $\Delta(h)$ can be interpreted as a natural local time corresponding to the random vector ζ_h.

In this connection we shall assume that one can correspond a positive nonrandom quantity Δt_{nk} to each vector $\Delta\xi_{nk}=\xi_{nk+1}-\xi_{nk}$ such that

$$\frac{1}{\Delta t_{nk}}\mathsf{E}\{e^{i(\Delta\xi_{nk},z)}-1\,|\,\mathfrak{F}_{nk}\}=L(t_{nk}, z)+\rho_{nk},$$

where

$$L(t, z)=i(a(t), z)-\tfrac{1}{2}(b(t)z, z)+\int_{\mathscr{R}^m}\left(e^{i(z,u)}-1-\frac{i(u, z)}{1+|z|^2}\right)\frac{1+|u|^2}{|u|^2}\Pi(t, du),$$

$t_{nk}=\Delta t_{n0}+\cdots+\Delta t_{nk}$, $a(t)$, $b(t)$, and $\Pi(t, A)$ are nonrandom quantities, $a(t)$ is a vector function, and $b(t)$ is a nonnegatively definite matrix; also, $\Pi(t, A)$ is a finite measure on $\mathfrak{B}^m$ with $\Pi(t, \{0\})=0$.

Furthermore, we assume that $t_{nm_n}=T$, $\max_k \Delta t_{nk}\to 0$ as $n\to\infty$, and the function $L(t, z)$ is Riemann integrable on the segment $[0, T]$.

Theorem 10. *If the preceding assumptions are fulfilled and if*

$$\mathsf{E}\left(\sum_{k=0}^{m_n-1}|\rho_{nk}|\,\Delta t_{nk}\right)\to 0 \quad \text{as } n\to\infty, \tag{19}$$

then the distribution of the vector $\xi_{nm_n}-\xi_{n0}$ is weakly convergent (as $n\to\infty$) to the distribution with the characteristic function

$$J(z)=\exp\left\{\int_0^T L(t, z)\,dt\right\}.$$

Proof. The proof of this theorem is analogous to the proof of Theorem 6. We introduce the quantity

$$\sigma_n=\mathsf{E}\{\exp\{i(\xi_{nm_n}-\xi_{n0}, z)\}-\exp\{\textstyle\int_0^T L(t, z)\,dt\}\,|\,\mathfrak{F}_{n0}\}$$

and represent it in the form

$$\sigma_n=\mathsf{E}\left\{\sum_{k=1}^{m_n-1}\exp\{i(\xi_{nk}-\xi_{n0}, z)\}\exp\{\textstyle\int_{t_{nk+1}}^T L(t, z)\,dt\}\mathsf{E}\{\tilde\sigma_{nk}\,|\,\mathfrak{F}_{nk}\}\,\middle|\,\mathfrak{F}_{n0}\right\},$$

where

$$\tilde{\sigma}_{nk} = \exp\{i(\Delta\xi_{nk}, z)\} - \exp\{\int_{t_{nk}}^{t_{nk+1}} L(t, z)\,dt\}.$$

Observe that

$$|\mathsf{E}\{\tilde{\sigma}_{nk} \mid \mathfrak{F}_{nk}\}| \le |\rho_{nk}|\,\Delta t_{nk} + |\int_{t_{nk}}^{t_{nk+1}} [L(t, z) - L(t_{nk}, z)]\,dt$$
$$+ \int_{t_{nk}}^{t_{nk+1}} L(t, z)\,dt\,(\exp\{\theta \int_{t_{nk}}^{t_{nk+1}} L(t, z)\,dt\} - 1)|.$$

Taking into account the fact that $\exp\{\int_a^T L(t, z)\,dt\}$ is a characteristic function of a distribution we obtain the following bound for σ_n:

$$\sigma_n \le \mathsf{E} \sum_{k=0}^{m_n-1} |\rho_{nk}|\,\Delta t_{nk} + \sum_{k=0}^{m_n-1} (\delta_{nk}\,\Delta t_{nk} + C(z)\,\Delta t_{nk}^2);$$

here δ_{nk} is the oscillation of the function $L(t, z)$ on the interval $[t_{nk}, t_{nk+1}]$, $C(z)$ is a constant dependent on T and $\sup_t |L(t, z)|$ only. The inequality obtained proves the theorem. □

In the same manner as in the case of convergence to a Wiener process one can easily deduce the following result from the theorem just proved.

Theorem 11. *If the conditions of Theorem* 10 *are satisfied and if* $t_{nk_j} \to t_j$ $(j = 1, \ldots, r)$, *then the joint distribution of the differences*

$$\xi_{nk_1} - \xi_{n0},\ \xi_{nk_2} - \xi_{nk_1},\ \ldots,\ \xi_{nk_r} - \xi_{nk_{r-1}}$$

converges weakly to the joint distribution of the vectors

$$\xi(t_1) - \xi(0),\ \xi(t_2) - \xi(t_1),\ \ldots,\ \xi(t_r) - \xi(t_{r-1}),$$

where $\xi(t)$ *is an* m*-dimensional process with independent increments such that the distribution of the random variable* $\xi(s+h) - \xi(s)$ *possesses the characteristic function*

$$J(s, s+h, z) = \exp\{\int_s^{s+h} L(t, z)\,dt\}.$$

Limit theorems for sequences of series of random vectors with finite moments of the second order. We now investigate the conditions for convergence of a sequence of series (1) to processes more general than the processes with independent increments. In accordance with the constructions above we shall assume that a sequence of nonrandom times $0 < t_{n0} < t_{n1} < \cdots < t_{nm_n} = T$ corresponds to the sequence $\xi_{n0}, \xi_{n1}, \ldots, \xi_{nm_n}$ and we shall introduce the following representation for the variables

$$\Delta\xi_{nk} = a_{nk}(\xi_{n0}, \xi_{n1}, \ldots, \xi_{nk})\,\Delta t_{nk} + b_{nk}(\xi_{n0}, \xi_{n1}, \ldots, \xi_{nk})\,\Delta\psi_{nk},$$

where $\{\psi_{nk}, k=0, 1, \ldots, m_n\}$ is a martingale and

$$\mathsf{E}\{\Delta\psi_{nk}\,\Delta\psi_{nk}^* \mid \mathfrak{F}_{nk}\} = \boldsymbol{I}\,\Delta t_{nk}.$$

Define on $[0, T]\times\mathscr{D}[0, T]$ functions

$$a_n(t, x(\cdot)), \qquad b_n(t, x(\cdot)), \qquad t\in[0, T], \quad x(\cdot)\in\mathscr{D}[0, T],$$

by setting

$$a_n(t, x(\cdot)) = a_{nk}(x(0), x(t_{n1}), \ldots, x(t_{nk})) \quad \text{for } t\in[t_{nk}, t_{nk+1}),$$
$$k=0, 1, \ldots, m_n-1,$$
$$a_n(T, x(\cdot)) = a_{nm_n-1}(x(0), x(t_{n1}), \ldots, x(t_{nm_n-1})).$$

The quantities $b_n(t, x(\cdot))$ are defined analogously. It follows from the definition that if $x(t)=y(t)$ for $t\in[0, s]$, then $a_n(t, x(\cdot))=a_n(t, y(\cdot))$ and $b_n(t, x(\cdot))= b_n(t, y(\cdot))$ for all $t\in[0, s]$. Our basic assumption is now as follows: the functions $a_n(t, x(\cdot))$ and $b_n(t, x(\cdot))$ converge in $[0, T]\times\mathscr{D}[0, T]$ as $n\to\infty$ to the functions $a(t, x(\cdot))$ and $b(t, x(\cdot))$, respectively. More precisely, we shall assume that the following condition is satisfied:

$$\lim_{n\to\infty} \sup_{\substack{t\in[0, T]\\ x(\cdot)\in\mathscr{D}[0, T]}} \{[1+\|x(\cdot)\|]^{-1}[|a_n(t, x(\cdot))-a(t, x(\cdot))| + |b_n(t, x(\cdot))-b(t, x(\cdot))|]\}=0, \tag{20}$$

where $\|x(\cdot)\| = \sup_{0\le t\le T}|x(t)|$.

In accordance with our general aim we would like now to establish that the process $\xi_n(t)$ constructed from the sequence of series of random vectors (1) converges to the process $\xi(t)$ which is a solution of the stochastic differential equation

$$d\xi(t) = a(t, \xi(\cdot)) + b(t, \xi(\cdot))d\psi(t), \tag{21}$$

where $\psi(t)$ is the limit process for $\psi_n(t)$ constructed from the martingale ψ_{nk}, $k= 1, \ldots, m_n$. To achieve this, several bounds are required.

Together with the system of random vectors (1) consider the sequence of series $\{\eta_{nk}, k=0, 1, \ldots, m_n\}$, $n=1, 2, \ldots$, defined by the recurrent sequence of relations

$$\begin{aligned} &\eta_{n0} = \xi_{n0}, \\ &\Delta\eta_{nk} = \eta_{nk+1}-\eta_{nk} = a(t_{nk}, \eta_n(\cdot))\,\Delta t_{nk} + b(t_{nk}, \eta_n(\cdot))\,\Delta\psi_{nk}, \end{aligned} \tag{22}$$

where $\eta_n(t)=\eta_{nk}$ for $t\in[t_{nk}, t_{nk+1})$, $k=0, 1, \ldots, m_n$. Such a definition makes sense since to evaluate the values $a(t_{nk}, \eta_n(\cdot))$ and $b(t_{nk}, \eta_n(\cdot))$ only the values of $\eta_{n0}, \eta_{n1}, \ldots, \eta_{nk}$ are needed.

Lemma 2. *Assume that conditions* (5) *and* (20) *are satisfied and, moreover, let*

$$(23)\qquad |a(t, x(\cdot))-a(t, y(\cdot))|+|b(t, x(\cdot))-b(t, y(\cdot))|\leqslant C\|x(\cdot)-y(\cdot)\|.$$

Then

$$\mathsf{E}\{\sup_{0\leqslant k\leqslant r} |\eta_{nk}-\xi_{nk}|^2 \mid \mathfrak{F}_{n0}\}\leqslant \varepsilon_n(1+|\xi_{n0}|^2)t_{nr},$$

where ε_n is a nonrandom quantity which tends to 0 *as $n\to\infty$.*

Proof. We represent the difference $\eta_{nk+1}-\xi_{nk+1}$ in the form

$$\begin{aligned}\eta_{nk+1}-\xi_{nk+1} &= \sum_{j=0}^{k} [a(t_{nj}, \eta_n(\cdot))-a(t_{nj}, \xi_n(\cdot))]\,\Delta t_{nj}\\ &\quad+\sum_{j=0}^{k} [b(t_{nj}, \eta_n(\cdot))-b(t_{nj}, \xi_n(\cdot))]\,\Delta\psi_{nj}\\ &\quad+\sum_{j=0}^{k} [a(t_{nj}, \xi_n(\cdot))-a_n(t_{nj}, \xi_n(\cdot))]\,\Delta t_{nj}\\ &\quad+\sum_{j=0}^{k} [b(t_{nj}, \xi_n(\cdot))-b_n(t_{nj}, \xi_n(\cdot))]\,\Delta\psi_{nj}\\ &=\textstyle\sum_k'+\sum_k''+\sum_k'''+\sum_k^{\mathrm{IV}}.\end{aligned}$$

Set

$$v_{nk}=\mathsf{E}\{\sup_{0\leqslant j\leqslant k} |\eta_{nj}-\xi_{nj}|^2 \mid \mathfrak{F}_{n0}\}.$$

We now bound the sums $\sum_k', \ldots, \sum_k^{\mathrm{IV}}$ using methods analogous to those applied in the proof of Lemma 1. For instance, utilizing the fact that $\sum_k''$ is a martingale we obtain

$$\begin{aligned}\mathsf{E}\{\sup_{j\leqslant k} |\textstyle\sum_j''|^2 \mid \mathfrak{F}_{no}\} &\leqslant 4\mathsf{E}\{|\textstyle\sum_k''|^2 \mid \mathfrak{F}_{n0}\}\\ &\leqslant 4\sum_{j=0}^{k} \mathsf{E}\{|b(t_{nj}, \eta_n(\cdot))-b(t_{nj}, \xi_n(\cdot))|^2\,\Delta t_{nj} \mid \mathfrak{F}_{n0}\}.\end{aligned}$$

Applying inequality (23) we observe that the quantity to be bound does not exceed

$$4C\sum_{j=0}^{k} v_{nj}\,\Delta t_{nj}.$$

Using (20) one easily obtains the inequality

$$\mathsf{E}\{\sup_{j\le k} |\Sigma_j^{\mathrm{IV}}|^2 \,|\, \mathfrak{F}_{n0}\} \le \varepsilon_n \sum_{j=0}^{k} \mathsf{E}\{(1+\sup_{r\le j} |\xi_{nr}|^2)\, \Delta t_{nj} \,|\, \mathfrak{F}_{n0}\},$$

where $\varepsilon_n \to 0$ as $n \to \infty$. The quantities $\sup_{j\le k} |\Sigma_j'|^2$ and $\sup_{j\le k} |\Sigma_j'''|^2$ are estimated analogously.

Utilizing Lemma 1 we obtain the relation

$$v_{nk+1} \le C' \sum_{j=0}^{k} v_{nj}\, \Delta t_{nj} + \varepsilon_n t_{nk+1}(1+|\xi_{n0}|^2),$$

where C' is a constant which depends on C and T only. This implies that

$$v_{nk+1} \le \varepsilon_n (1+|\xi_{n0}|^2)(e^{C't}-1).$$

The lemma is proved. □

It follows from Lemma 2 that the marginal distributions of processes $\xi_n(t)$ and $\eta_n(t)$ can weakly converge only simultaneously and that the corresponding limits coincide. Now it would be more convenient to study the limiting behavior of processes $\eta_n(t)$.

Let η'_{nk} and η''_{nk}, $k = 0, 1, \ldots, m_n$, be sequences constructed from formulas (22) under the distinct initial conditions $\eta'_{n0} = \xi'$ and $\eta''_{n0} = \xi''$. Analogously to Lemma 2 the following lemma can be proven:

Lemma 3. *If the conditions of Lemma 2 are satisfied, then*

$$\mathsf{E}\{\sup_{0\le j\le k} |\eta'_{nj} - \eta''_{nj}|^2 \,|\, \mathfrak{F}_{n0}\} \le e^{c't_{nk}} |\xi' - \xi''|^2,$$

where c' is a constant.

Above we introduced finite-difference approximations for stochastic differential equations and showed that these converge to solutions of stochastic differential equations (Section 1, Theorems 12 and 13). We shall now verify analogous assertions for the processes $\eta_n(t)$. The role of finite-difference approximations for processes $\eta_n(t)$ is taken here by processes $\zeta_n(t)$ defined as follows. Choose some values $t_{nk_1}, t_{nk_2}, \ldots, t_{nk_r}$ where r is a fixed integer. For brevity of notation set $t_{nk_j} = s_j$, $j = 1, 2, \ldots, r$, $s = 0$, $s_{r+1} = T$, and for $t \in (s_j, s_{j+1}]$, $j = 0, 1, \ldots, r-1$, let

$$\zeta_n(0) = \xi_{n0},$$
$$\zeta_n(t) = \zeta_n(s_j) + a(s_j, \zeta_n(\cdot))(t - s_j) + b(s_j, \zeta_n(\cdot))[\psi_n(t) - \psi_n(s_j)].$$

We now bound the quantity

$$v_n(t) = \mathsf{E}\{\sup_{0\le s = t_{nj} < t} |\eta_n(s) - \zeta_n(s)|^2\}.$$

We have

$$v_n(T)\le 2\mathsf{E}\sup|\Sigma_t'|^2+2\mathsf{E}\sup|\Sigma_t''|^2, \tag{24}$$

where

$$\Sigma_t'=\sum_{k=0}^{j-1}[a(s_{nk},\zeta_n(\cdot))-a(t_{nk},\eta_n(\cdot))]\,\Delta t_{nk}+\sigma_n'(t),$$

$$\Sigma_t''=\sum_{k=0}^{j-1}[b(s_{nk},\zeta_n(\cdot))-b(t_{nk},\eta_n(\cdot))]\,\Delta\psi(t_{nk})+\sigma_n''(t),$$

and

$$\sigma_n'(t)=a(s_{nj},\zeta_n(\cdot))(t-t_{nj}),$$
$$\sigma_n''(t)=b(s_{nj},\zeta_n(\cdot))[\psi_n(t)-\psi_n(t_{nj})],$$

for $t\in[t_{nj},t_{nj+1})$ and $s_{nk}=s_i$ if $t_{nk}\in[s_i,s_{i+1})$.

Observe that if $t_{nk}\in[s_i,s_{i+1})$, then

$$\begin{aligned}&|a(s_{nk},\zeta_n(\cdot))-a(t_{nk},\eta_n(\cdot))|\\ &\le|a(s_i,\zeta_n(\cdot))-a(s_i,\eta_n(\cdot))|+|a(s_i,\eta_n(\cdot))-a(t_{nk},\eta_n(\cdot))|\\ &\le C\sup_{s\le s_i}|\eta_n(s)-\zeta_n(s)|+\rho(t_{nk}-s_i)(\sup_{s\le t_{nk}}|\eta_n(s)|+1).\end{aligned} \tag{25}$$

Here we introduce the following condition: for $t>s$

$$|a(s,x(\cdot))-a(t,x(\cdot))|\le\rho(t-s)(1+\sup_{0\le t'\le t}|x(t')|), \tag{26}$$

where $\rho(t)$, $t>0$, is a nonnegative monotonically nondecreasing function and $\rho(0+)=0$.

Assume that the same inequality is valid for the matrix-valued function $b(t,x(\cdot))$:

$$|b(s,x(\cdot))-b(t,x(\cdot))|\le\rho(t-s)(1+\sup_{0\le t'\le t}|x(t')|). \tag{27}$$

Then an inequality analogous to (25) will hold also for the differences $|b(s_{nk},\zeta_n(\cdot))-b(t_{nk},\eta_n(\cdot))|$.

Also set $w_n(t)=\mathsf{E}\sup_{0\le s\le t}|\eta_n(s)|^2$. It is easy to verify that

$$\begin{aligned}\mathsf{E}\sup_t|\Sigma_t'|^2&\le 2T\sum_k C^2\mathsf{E}\sup_{s<s_i}|\eta_n(s)-\zeta_n(s)|^2\,\Delta t_{nk}\\ &\quad+2T\sum_k\rho(t_{nk}-s_i)(1+w_n(t_{nk}))\,\Delta t_{nk}\end{aligned}$$

or

$$\mathsf{E}\sup_t |\textstyle\sum_t'|^2 \leq 2TC^2 \sum_{i=0}^{r} v_n(s_i)(s_{i+1}-s_i) + 2T^2(1+w_n(T))\rho(|\delta|),$$

where $|\delta| = \max_{0\leq i<r}(s_{i+1}-s_i)$.

We now bound the second summand in the right-hand side of inequality (24). For this purpose observe that the sum $\sum_t''$ as a function of t is a square integrable martingale. Therefore

$$\mathsf{E}\sup_t |\textstyle\sum_t''|^2 \leq 4\mathsf{E}|\sum_T''|^2 = 4\mathsf{E}\sum_k |b(s_{nk}, \zeta_n(\cdot)) - b(t_{nk}, \eta_n(\cdot))|^2 \Delta t_{nk}.$$

From this relation we obtain analogously to the above that

$$\mathsf{E}\sup_t |\textstyle\sum_t''|^2 \leq 8C^2 \sum_{i=0}^{r} v_n(s_i)(s_{i+1}-s_i) + 8T\rho(|\delta|)(1+w_n(T)).$$

Thus

$$v_n(T) \leq 2C^2(T+4)\sum_{i=0}^{r} v_n(s_i)(s_{i+1}-s_i) + \rho(|\delta|)C'(w_n(T)+1). \tag{28}$$

The bound on $w_n(T)$ can be deduced from Lemma 1:

$$w_n(T) \leq C_1(1+\mathsf{E}|\xi_{n0}|^2).$$

Observe that the function $v_n(t)$ is monotonically nondecreasing.

Let $\bar{v}_n(t) = v_n(s_i)$ for $t \in [s_i, s_{i+1})$. One may replace the function $v_n(s)$ in inequality (28) by $\bar{v}_n(s)$ and T by any $t \in [0, T]$. This yields the following integral inequality,

$$\bar{v}_n(t) \leq C_1 \int_0^t \bar{v}_n(s)\, ds + C_2\rho(|\delta|),$$

which implies that

$$\bar{v}_n(t) \leq C_2\rho(|\delta|)\, e^{C_1 T} \leq C_3\rho(|\delta|).$$

Here C_3 is a constant of the form $C_3 = C'(1+\mathsf{E}|\xi_{n0}|^2)$ and C' depends on C and T only. We have thus proved the following lemma:

Lemma 4. *If the conditions of Lemma* 2 *and inequalities* (26) *and* (27) *are satisfied, then*

$$\mathsf{E}\{\sup_{0\leq t\leq T} |\eta_n(t) - \zeta_n(t)|^2 \,|\, \mathfrak{F}_{n0}\} \leq C'\rho(|\delta|)(1+\mathsf{E}|\xi_{n0}|^2).$$

Up until now no assumptions have been imposed concerning the convergence of processes $\psi_n(t)$. Recall that processes $\psi_n(t)$, $t\in[0, T]$, are $\mathfrak{F}_{nt}$-martingales, where $\mathfrak{F}_{nt}=\mathfrak{F}_{nk}$ for $t\in[t_{nk}, t_{nk+1})$ and, moreover,

$$\mathsf{E}\{\Delta\psi_{nk}\,\Delta\psi^*_{nk}\,|\,\mathfrak{F}_{nk}\}=I\,\Delta t_{nk}.$$

It is natural to suppose that the limiting process $\psi(t)$ is also a square integrable martingale with respect to a current of σ-algebras $\{\mathfrak{F}_t, t\in[0, T]\}$ and that

$$\mathsf{E}\{\Delta\psi\,\Delta\psi^*\,|\,\mathfrak{F}_s\}=I(t-s), \tag{29}$$

where $\Delta\psi=\psi(t)-\psi(s)$, $s<t$. The conditions for convergence to martingales with independent increments were discussed above.

We thus assume that the following condition is fulfilled:

Ψ_1: Measures $Q_n(\cdot,\cdot)$ in the space $\mathcal{R}^m\times\mathcal{D}$ generated by random vectors ξ_{n0} and random processes $\psi_n(t)$, $t\in[0, T]$, converge weakly to the measure $Q(\cdot,\cdot)$ associated with the random vector ξ_0 and a square integrable martingale $\psi(t)$.

We now proceed to the proof of the weak convergence of marginal distributions of the process $\xi_n(t)$ to the marginal distributions of a solution of stochastic equation (21). Note that if the conditions of Lemma 2 are satisfied, equation (21) has then a unique solution with finite second-order moments (Theorem 3 in Section 1).

Let $\xi_\delta(t)$ denote a finite-difference approximation of the solution of equation (21) constructed from the subdivision

$$\delta=\{0=s_0, s_1, \dots, s_r, s_{r+1}=I\}\quad\text{of the interval } [0, T].$$

Here we shall modify slightly the definition of the process $\xi_\delta(t)$. Namely, for $t\in(s_i, s_{i+1}]$ we set

$$\xi_\delta(t)=\xi_\delta(s_i)+a(s_i, \xi_\delta(\cdot))(t-s_i)+b(s_i, \xi_\delta(\cdot))(\psi(t)-\psi(s_i)),$$

and $\xi_\delta(0)=\xi_0$. It is easy to verify that if inequalities (26) and (27) are fulfilled, then the assertion of Theorem 12 in Section 1 remains valid under this modification as well, so that

$$\mathsf{E}\{\sup_{0\le t\le T}|\xi_\delta(t)-\xi(t)|^2\}\to 0.$$

Using induction it is easy to show that $\xi_\delta(t)$ is a continuous function in the argument ξ_0 for each t and is a continuous functional on $\psi(s)$, $0\le s\le t$, $\xi_\delta(t)=g_{(t,\delta)}(\xi_0, \psi(\cdot))$. Also, $\zeta_n(t)$ is expressible in terms of ξ_{n0} and $\psi_n(s)$ in the same manner: $\zeta_n(t)=g_{(t,\delta)}(\xi_{n0}, \psi_n(\cdot))$.

Now choose an arbitrary sequence $\{t_k, k=1, \dots, p\}$, $t_k\in[0, T]$, and an arbitrary continuous bounded function $f(x_0, \dots, x_p)$, $x_k\in\mathcal{R}^m$, and bound the difference

$$r_n = \mathsf{E}f(\xi_0, \xi(t_1), \ldots, \xi(t_p)) - \mathsf{E}f(\xi_{n0}, \eta_n(t_1), \ldots, \eta_n(t_p)).$$

We have

$$\begin{aligned} |r_n| &\leq \mathsf{E}|f(\xi_0, \xi(t_1), \ldots, \xi(t_p)) - f(\xi_0, \xi_\delta(t_1), \ldots, \xi_\delta(t_p))| \\ &\quad + |\mathsf{E}f(\xi_0, \xi_\delta(t_1), \ldots, \xi_\delta(t_p)) - \mathsf{E}f(\xi_{n0}, \zeta_n(t_1), \ldots, \zeta_n(t_p))| \\ &\quad + \mathsf{E}|f(\xi_{n0}, \zeta_n(t_1), \ldots, \zeta_n(t_p)) - f(\xi_{n0}, \eta_n(t_1), \ldots, \eta_n(t_p))| \\ &= r' + r_n'' + r_n'''. \end{aligned}$$

Given an arbitrary $\varepsilon > 0$ we first choose a δ such that $r' < \varepsilon/3$ and $r_n''' < \varepsilon/3$ for sufficiently large n. We need now only note that in view of the preceding remarks

$$f(\xi_0, \xi_\delta(t_1), \ldots, \xi_\delta(t_p)) = F_{(t,\delta)}(\xi_0, \psi(\cdot)), \qquad f(\xi_{n0}, \zeta_n(t_1), \ldots, \zeta_n(t_p)) = F_{(t,\delta)}(\xi_{n0}, \psi_n(\cdot)),$$

where $F_{(t,\delta)}(x, \psi(\cdot))$ is a bounded continuous function in x and ψ. Thus if the conditions Ψ_1 are satisfied $r_n'' \to 0$ as $n \to \infty$. Hence we have proved the following theorem:

Theorem 12. *Assume that a sequence of series of random vectors* (1) *satisfies conditions* (5), (20), *and* Ψ_1 *and the functionals* $a(t, x(\cdot)), b(t, x(\cdot))$ *satisfy conditions* (23), (26), *and* (27). *Then the marginal distributions of the processes* $\xi_n(t)$ *constructed from the sequence* (1) *converge weakly to the corresponding marginal distributions of a solution of the stochastic equation* (21).

An important particular case of the model just discussed is one in which the coefficients $a(t, x(\cdot))$ and $b(t, x(\cdot))$ in equation (21) are independent of the "past", i.e.,

$$a(t, x(\cdot)) = a(t, x(t)), \qquad b(t, x(\cdot)) = b(t, x(t)). \tag{30}$$

Equation (21) then becomes a stochastic differential equation without delay,

$$d\xi = a(t, \xi(t))\, dt + b(t, \xi(t))\, d\psi(t), \qquad \xi(0) = \xi_0, \tag{31}$$

and the functions $a(t, x)$ and $b(t, x)$, where $(t, x) \in [0, T] \times \mathcal{R}^m$, satisfy the conditions

$$|a(t, x)| + |b(t, x)| \leq C(1 + |x|), \tag{32}$$

$$|a(t, x) - a(t, y)| + |b(t, x) - b(t, y)| \leq C|x - y|, \tag{33}$$

$$|a(s, x) - a(t, x)| + |b(s, x) - b(t, x)| \leq \rho(t - s)(1 + |x|), \tag{34}$$

where $\rho(t)$ is as defined above.

If, moreover, the martingale $\psi(t)$ turns out to be a process with independent increments, then equation (31) becomes an equation without an after-effect and

$$A(x, t)=\int_0^t a(t, x)\,dt+\int_0^t b(t, x)\,d\psi(t)\in \bar{S}(C, C).$$

Theorem 12 can be somewhat strengthened in the case under consideration by noting that $\xi_\delta(t)$ is a continuous function in arguments $\xi_0, \Delta\psi(s_k)=\psi(s_{k+1})-\psi(s_k)$, $k=0, 1, \ldots, i-1$, and $\psi(t)-\psi(s_i)$ for $t\in(s_i, s_{i+1}]$. Thus the functionals $F_{(t,\delta)}(x, x(\cdot))$ introduced above become continuous functions in x and in a finite number of differences of the form $x(t_{k+1})-x(t_k)$. Hence we can weaken condition Ψ_1 and replace it by the following:

Ψ_2: For any r and $t_k, t_k\in[0, T]$, $k=1, \ldots, r$, the joint distribution of the random vectors

$$\xi_n, \psi_n(t_1), \ldots, \psi_n(t_r)$$

converges weakly to the joint distribution of random vectors

$$\xi_0, \psi(t_1), \ldots, \psi(t_r)$$

defined on a certain probability space, where $\psi(t)$ is a square integrable martingale satisfying condition (29).

Theorem 13. *Assume that a sequence of series* (1) *satisfies condition* (20) *and let the functionals $a(t, x(\cdot))$ and $b(t, x(\cdot))$ be functions of the form* (30) *satisfying conditions* (26), (27), (32), *and* (33), *and martingales $\psi_n(t)$ satisfy condition Ψ_2. Then the marginal distributions of processes $\xi_n(t)$ converge weakly to the corresponding marginal distributions of a solution $\xi(t)$ of the stochastic differential equation* (31).

If Lindeberg's condition is satisfied, i.e., for any $\varepsilon>0$

$$\sum_{k=1}^{m_n-1} \mathsf{E}\bar{\chi}_{nk}(\varepsilon)|\Delta\psi_{nk}|^2\to 0 \quad \text{as } n\to\infty,$$

where $\bar{\chi}_{nk}(\varepsilon)=1$ if $|\Delta\xi_{nk}|\geqslant\varepsilon$ and $\bar{\chi}_{nk}(\varepsilon)=0$ otherwise, then condition Ψ_2 is satisfied, $\psi(t)$ is a Wiener process, and the measures $q_n(\cdot)$ associated with the random processes $\xi_n(t)$ converge weakly in $\mathfrak{D}$ to the measure associated with a solution of the stochastic equation (31).

Limit theorems for stochastic differential equations. Consider stochastic differential equations

$$d\xi_u=A_u(\xi_u, dt), \qquad \xi_u(0)=\xi_u^0, \qquad t\in[0, T], \tag{35}$$

dependent on a parameter $u \in [0, u_0]$. Here

$$A_u(x, t) = \int_0^t \alpha_u(x, s)\, ds + \beta_u(x, t),$$

and $A_u \in S(\lambda_0^u, \lambda_N^u)$. One limit theorem for these equations was considered above (Theorem 11 in Section 1). In this subsection we shall discuss conditions for weak convergence of measures $q_u(\cdot)$ generated by solutions of equations (35) in $\mathscr{D}$ ($\mathscr{D} = \mathscr{D}^m[0, T]$). We shall prove the following theorem.

Theorem 14. *Let the following conditions be satisfied:*

(36) a) $$\lim_{N\to\infty} \sup_{u\in[0, u_0]} \mathsf{P}\{\sup_{0\le t\le T} |\lambda_0^u(t)| > N\} = 0;$$

b) *For any* $N_1 > 0$

(37) $$\lim_{N\to\infty} \sup_{u\in[0, u_0]} \mathsf{P}\{\sup_{0\le t\le T} \lambda_{N_1}^u(t) > N\} = 0;$$

c) *Marginal distributions of random functions*

$$\int_0^t \alpha_u(x, s)\, ds, \qquad \beta_u(x, t),$$

converge weakly as $u \to 0$ *to the corresponding distributions of random functions*

$$\int_0^t \alpha_0(x, s)\, ds \qquad \beta_0(x, t).$$

d) *The distribution of vector* ξ_u^0 *as* $u \to 0$ *converges weakly to the distribution of the vector* ξ^0.

Then the measures $q_u(\cdot)$ *converge weakly to* $q(\cdot)$.

Note that it is not assumed in Theorem 14 that functions $A_u(t, x)$ are defined on the same probability space.

In the course of the proof of this theorem the following lemma on small perturbations of stochastic differential equations will be used (cf. Theorem 10, Section 1).

Lemma 5. *Let*

$$d\tilde{\xi}_u = A_u(\tilde{\xi}_u(t), dt) + A_{\delta u}(\tilde{\xi}_u(t), dt), \qquad \delta > 0,$$
$$\tilde{\xi}_u(0) = \xi_u^0, \qquad u \in [0, u_0],$$

and let the following conditions be satisfied:

a) $A_u(x, t) \in S(\lambda_0^u, \lambda_N^u)$, *where the functions* λ_0^u *and* λ_N^u *satisfy conditions* (36) *and* (37).

b) $A_{\delta u}(x, t) \in S(\lambda_0^u, \tilde{\lambda}_N^{\delta u})$, *where* $\lambda_0^u(t)$ *is as in condition* a) *and, moreover,*

$$|\alpha_{\delta u}(x, t)| \leqslant \gamma_{\delta u}(x, t),$$
$$\mathsf{E}\{|\Delta \beta_{\delta u}(x, t)|^2 \,|\, \mathfrak{F}_t\} \leqslant \mathsf{E}\{\textstyle\int_t^{t+\Delta t} \gamma_{\delta u}^2(x, s)\, ds \,|\, \mathfrak{F}_t\},$$
$$\lim_{\delta \to 0} \sup_{u \in (0, u_0]} \mathsf{P}\{\sup_{\substack{t \in [0,T] \\ |x| \leqslant N}} \gamma_{\delta u}^2(x, t) > \varepsilon\} = 0$$

for any $\varepsilon > 0, N > 0$.

c) *The distribution of the initial vector* ξ_u^0 *as* $u \to 0$ *converges weakly to a limit. Then*

$$\mathsf{P}\{\sup_{0 \leqslant t \leqslant T} |\tilde{\xi}_u(t) - \xi_u(t)| > \varepsilon\} \to 0 \quad \text{as } \delta \to 0$$

uniformly in u.

Proof. Let ε' be an arbitrary given positive number,

$$\tau = \inf\{t\colon \lambda_0^u(t) \geqslant N, \quad \lambda_{N_1}^u(t) \geqslant N_2, \quad \sup\nolimits_{|x| \leqslant N_1} \gamma_{\delta u}(x, t) \geqslant \varepsilon'\},$$

if the set of values of t in the braces is nonvoid and $\tau = T$ otherwise. Here N, N_1, and N_2 are positive numbers to be specified below. For the time being we note only that inequality

$$\mathsf{P}\{\tau < T\} \leqslant \mathsf{P}\{\sup_{0 \leqslant t \leqslant T} \lambda_0^u(t) \geqslant N, \sup_{0 \leqslant t \leqslant T} \lambda_{N_1}^u(t) \geqslant N_2\} + \mathsf{P}\{\sup_{\substack{0 \leqslant t \leqslant T \\ |x| \leqslant N_1}} \gamma_{\delta u}(x, t) > \varepsilon'\}$$

and the assumptions of the lemma imply that for any N_1 and $\varepsilon > 0$ there exists a δ_0 sufficiently small independent of u, N, and N_2, and N^0, N_2^0 sufficiently large, independent of u, ε', and δ, such that

$$\mathsf{P}\{\tau < T\} < \varepsilon$$

for $N \geqslant N^0$, $N_2 \geqslant N_2^0$ and $\delta < \delta_0$.

We now construct functions $\alpha'_u(x, t)$ and $\beta'_u(x, t)$ in such a manner that they coincide with $\alpha_u(x, t)$ and $\beta_u(x, t)$, respectively, for $|x| \leqslant N_1$ and belong to the class $S(\lambda_0^u, \lambda^{u'})$, where $\lambda^{u'} = \lambda^{u'}(t)$ does not depend on N and $\lambda^{u'}(t) \leqslant 1 + \lambda_{N_1}^u(t)$. Set $\bar{\alpha}_u(x, t) = \alpha'_u(x, t)$ for $t \leqslant \tau$, $\bar{\alpha}_u(x, t) = 0$ for $t > \tau$, $\bar{\beta}_u(x, t) = \beta'_u(x, t \wedge \tau)$, $\bar{\alpha}_{\delta u}(x, t) = \alpha_{\delta u}(x, t)$ *for* $t \leqslant \tau$, $\bar{\alpha}_{\delta u}(x, t) = 0$ *for* $t > \tau$, $\bar{\beta}_{\delta u}(x, t) = \beta_{\delta u}(x, t \wedge \tau)$. Moreover, $\bar{A}_u(x, t) \in S(N, N_1 + 1)$, $\bar{A}_{\delta u}(x, t) \in S(N, \tilde{\lambda}_N^{\delta u})$, where

$$\bar{A}_u(x, t) = \textstyle\int_0^t \bar{\alpha}_u(x, s)\, ds + \bar{\beta}(x, t)$$

and $\bar{A}_{\delta u}(x, t)$ is defined analogously via $\bar{\alpha}_{\delta u}(x, t)$.

Consider the equations

$$d\eta(t)=\bar{A}_u(\eta(t), dt),$$
$$d\tilde{\eta}(t)=\bar{A}_u(\tilde{\eta}(t), dt)+\bar{A}_{\delta u}(\tilde{\eta}(t), dt),$$
$$\eta(0)=\tilde{\eta}(0)=\xi_u^0.$$

Solutions for these equations do exist and, furthermore, $\eta(t)=\xi_u(t)$ as long as $t<\tau$ and $|\eta(t)|<N_1$ analogously $\tilde{\eta}(t)=\tilde{\xi}_u(t)$ as long as $t<\tau$ and $\sup_{s<t}|\tilde{\eta}(s)|<N_1$.

We now bound the difference $\eta(t)-\tilde{\eta}(t)$.

Note that

$$\mathsf{P}\{\sup_{0\le t\le T}|\eta(t)-\tilde{\eta}(t)|>\varepsilon\}\le\mathsf{P}\{\sup_{0\le t\le T}|\tilde{\eta}(t)|\ge N_1\}$$
$$+\mathsf{P}\{\sup_{0\le t\le T}|\eta(t\wedge\tau_1)-\tilde{\eta}(t\wedge\tau_1)|>\varepsilon\},$$

where $\tau_1=\inf\{t\colon|\tilde{\eta}(t)|\ge N_1\}$ ($\inf\varnothing=T$).

It follows from the previous results (in Theorem 7, Section 1) that

$$\mathsf{E}\{\sup_{0\le t\le T}|\tilde{\eta}(t)|^2\,|\,\mathfrak{F}_0\}\le(1+|\xi_u^0|^2)C(N).$$

Consequently for any $C_0>0$

$$\mathsf{P}\{\sup_{0\le t\le T}|\tilde{\eta}(t)|\ge N_1\}\le\frac{(1+C_0^2)C(N)}{N_1^2}+\mathsf{P}\{|\xi_u^0|>C_0\}.$$

Set

$$v_u(t)=\mathsf{E}\sup_{0\le s\le t}|\eta(s\wedge\tau_1)-\tilde{\eta}(s\wedge\tau_1)|^2.$$

Since

$$|\eta(t)-\tilde{\eta}(t)|^2\le3(|\int_0^t[\bar{\alpha}_u(\eta(s), s)-\bar{\alpha}_u(\tilde{\eta}(s), s)]\,ds|^2$$
$$+|\int_0^t\bar{\beta}_u(\eta(s), ds)-\bar{\beta}_u(\tilde{\eta}(s), ds)|^2+|\int_0^t\bar{A}_{\delta u}(\tilde{\eta}(s), ds)|^2),$$

applying the methods utilized above we obtain

$$v_u(t)\le12\mathsf{E}\sup_{0\le s\le t}|\int_0^{s\wedge\tau_1}\bar{A}_{\delta u}(\tilde{\eta}(s), ds)|^2+C(N_2)\int_0^t v_u(s)\,ds,$$

where $C(N_2)$ depends on T and N_2 only. It follows from the last integral inequality that

$$v_u(t)\le C'(N_2)\mathsf{E}\sup_{0\le s\le t}|\int_0^{s\wedge\tau_1}\bar{A}_{\delta u}(\tilde{\eta}(s), ds)|^2.$$

Furthermore, we have

$$\begin{aligned}
&\mathsf{E}\sup_{0\le s\le t}\left|\int_0^{s\wedge\tau_1}\bar{A}_{\delta u}(\tilde{\eta}(s),ds)\right|^2\\
&\le 2\Big(t\mathsf{E}\int_0^{t\wedge\tau_1}|\bar{\alpha}_{\delta u}(\tilde{\eta}(s),s)|^2\,ds+\mathsf{E}\sup_{0\le s\le t}\left|\int_0^{s\wedge\tau_1}\tilde{\beta}_{\delta u}(\tilde{\eta}(s),ds)\right|^2\Big)\\
&\le (2t+8)\mathsf{E}\int_0^{t\wedge\tau_1\wedge\tau}\gamma_{\delta u}^2(\tilde{\eta}(s),s)\,ds.
\end{aligned}$$

Taking the definition of random times τ and τ_1 into account we obtain

$$v_u(t)\le\varepsilon' C''(N_2).$$

This implies that

$$\mathsf{P}\{\sup_{0\le t\le T}|\eta(t\wedge\tau_1)-\tilde{\eta}(t\wedge\tau_1)|>\varepsilon\}<\frac{\varepsilon' C''(N_2)}{\varepsilon^2}$$

and

$$\mathsf{P}\{\sup_{0\le t\le T}|\eta(t)-\tilde{\eta}(t)|>\varepsilon\}\le\mathsf{P}\{|\xi_u^0|>C_0\}+\frac{(1+C_0^2)C(N)}{N_1^2}+\frac{\varepsilon' C''(N_2)}{\varepsilon^2}.$$

Furthermore,

$$\begin{aligned}
\mathsf{P}\{\sup_{0\le t\le T}|\xi_u(t)-\tilde{\xi}_u(t)|>\varepsilon\}\le{}&\mathsf{P}\{\sup_{0\le t\le T}|\xi_u(t)-\eta(t)|>0\}\\
&+\mathsf{P}\{\sup_{0\le t\le T}|\eta(t)-\tilde{\eta}(t)|>\varepsilon\}\\
&+\mathsf{P}\{\sup_{0\le t\le T}|\tilde{\eta}(t)-\tilde{\xi}_u(t)|>0\};
\end{aligned}$$

moreover,

$$\mathsf{P}\{\sup_{0\le t\le T}|\xi_u(t)-\eta(t)|>0\}\le\mathsf{P}(\tau<T)+\mathsf{P}\{\sup_{0\le t\le T}|\eta(t)|\ge N_1\},$$

with an analogous inequality valid for the probability $\mathsf{P}\{\sup_{0\le t\le T}|\tilde{\xi}_u(t)-\tilde{\eta}(t)|>0\}$. Hence

$$\begin{aligned}
\mathsf{P}\{\sup_{0\le t\le T}|\xi_u(t)-\tilde{\xi}(t)|>\varepsilon\}\le{}&2\mathsf{P}(\tau<T)+3\mathsf{P}\{|\xi_u^0|>C_0\}\\
&+\frac{3(1+C_0^2)C(N)}{N_1^2}+\frac{\varepsilon' C''(N_2)}{\varepsilon^2}.
\end{aligned}$$

Given an ε one can now choose constants C_0, N, N_1, N_2, and ε' independent of u

in the following manner. Choose C_0 first from the condition $3\mathsf{P}\{|\xi_u^0|>C_0\}<\varepsilon/4$ and then N such that

$$2\mathsf{P}\{\sup_{0\le t\le T}\lambda_0^u(t)>N\}<\frac{\varepsilon}{12};$$

next, determine N_1 so that the inequality

$$3(1+C_0^2)C(N)/N_1^2<\varepsilon/4$$

will be fulfilled. Furthermore, let N_2 be such that

$$2\mathsf{P}\{\sup_{0\le t\le T}\lambda_{N_1}^u(t)>N_2\}<\frac{\varepsilon}{12};$$

now determine ε' from the condition $\varepsilon' C''(N)/\varepsilon^2<\varepsilon/4$ and, finally, choose δ_0 in such a manner that for $\delta<\delta_0$

$$2\mathsf{P}\{\sup_{\substack{0\le t\le T\\ |x|\le N_1}}\gamma_{\delta u}(x,t)>\varepsilon'\}<\frac{\varepsilon}{12}.$$

We then obtain that $\mathsf{P}\{\sup_{0\le t\le T}|\xi_u(t)-\tilde{\xi}_u(t)|>\varepsilon\}<\varepsilon$ for all $\delta<\delta_0$ and for any $u\in[0,u_0]$. The lemma is thus proved. □

Proof of Theorem 14. In view of the fact that condition a) of the theorem assures weak compactness of measures in $\mathscr{D}$ which are associated with solutions of equations (35), to prove the theorem it is sufficient to verify the convergence of marginal distributions of processes $\xi_u(t)$ to the corresponding marginal distributions of the process $\xi_0(t)$. We shall first prove this for the fields $\alpha_u(x,t)$ and $\beta_u(x,t)$ of a particular form and then proceed to the general case.

Let

$$\alpha_u(x,t)=\sum_{k=1}^r a_k(x)\alpha_u^k(t),\qquad \beta_u(x,t)=\sum_{k=1}^r b_k(x)\beta_u^k(t),$$

where $a_k(x)$ and $b_k(x)$, $k=1,\dots,r$, are nonrandom scalar differentiable functions with uniformly bounded derivatives, let $\beta_u^k(t)$ be square integrable martingales, and let

$$|\alpha_u^k(t)|\le\lambda_0^u(t),\qquad k=1,\dots,r,$$
$$\mathsf{E}\{|\Delta\beta_u^k(t)|^2\mid\mathfrak{F}_t\}\le\mathsf{E}\{\textstyle\int_t^{t+\Delta t}(\lambda_0^u(s))^2\,ds\mid\mathfrak{F}_t\}.$$

Assume that the function $\lambda_0^u(t)$ satisfies condition (36), that marginal distributions of the compound process

$$\left(\int_0^t\alpha_u^1(s)\,ds,\dots,\int_0^t\alpha_u^r(s)\,ds,\qquad \beta_u^1(t),\dots,\beta_u^r(t)\right)$$

are weakly convergent as $u \to 0$ to the corresponding distributions of a process

$$\left(\int_0^t \alpha_u^1(s)\,ds, \ldots, \int_0^t \alpha_0^r(s)\,ds, \qquad \beta_0^1(t), \ldots, \beta_0^r(t)\right)$$

and that the random vector ξ_u^0 converges in distribution to the vector ξ_0.

We shall prove Theorem 14 under these additional assumptions.

Let t_k, $t_k \in [0, T]$, $k = 1, \ldots, s$, be a given sequence of numbers. Introduce finite-difference approximations $\xi_{\delta u}(t)$ of solutions of equations (35) and assume that the points t_k, $k = 1, \ldots, s$, are included in the subdivision δ. Since the conditions stipulated in the Remark following Theorem 13 in Section 1 are fulfilled in our case, given an arbitrary $\varepsilon > 0$ one can find δ_0 such that for $|\delta| < \delta_0$

$$\mathsf{P}\{\sup_{0 \le t \le T} |\xi_u(t) - \xi_{\delta u}(t)| > \varepsilon\} \le \varepsilon$$

for all $u \in [0, u_0]$. Let $f(x_1, \ldots, x_s)$ be an arbitrary continuous and bounded function ($x_k \in \mathcal{R}^m$) and let its partial derivatives of the first order be bounded also. We have

$$\begin{aligned}
&|\mathsf{E}f[\xi_u(t_1), \ldots, \xi_u(t_s)] - \mathsf{E}f[\xi_0(t_1), \ldots, \xi_0(t_s)]| \\
&\le |\mathsf{E}(f[\xi_u(t_1), \ldots, \xi_u(t_s)] - f[\xi_{\delta u}(t_1), \ldots, \xi_{\delta u}(t_s)])| \\
&\quad + |\mathsf{E}f[\xi_{\delta u}(t_1), \ldots, \xi_{\delta u}(t_s)] - \mathsf{E}f[\xi_{\delta 0}(t_1), \ldots, \xi_{\delta 0}(t_s)]| \\
&\quad + |\mathsf{E}(f[\xi_{\delta 0}(t_1), \ldots, \xi_{\delta 0}(t_s)] - f[\xi_0(t_1), \ldots, \xi_0(t_s)])| \\
&= I_1 + I_2 + I_3.
\end{aligned}$$

Moreover,

$$I_1 \le C[\varepsilon + \mathsf{P}\{\sup_{0 \le t \le T} |\xi_{\delta u}(t) - \xi_u(t)| > \varepsilon\}],$$

$$I_3 \le C[\varepsilon + \mathsf{P}\{\sup_{0 \le t \le T} |\xi_{\delta 0}(t) - \xi_0(t)| > \varepsilon\}],$$

where C is a constant. Thus we have for $|\delta| < \delta_0$, $I_1 + I_3 \le 4C\varepsilon$ independently of the values of u. Furthermore, it is easy to verify that $f[\xi_{\delta u}(t_1), \ldots, \xi_{\delta u}(t_s)]$ is a continuous and bounded function of the quantities

$$\int_{s_j}^{s_{j+1}} \alpha_u(s)\,ds, \qquad \beta^k(s_{j+1}) - \beta^k(s_j), \qquad j = 0, 1, \ldots, l, \qquad k = 1, \ldots, r,$$

where s_j are the points which form the subdivision δ. Therefore for a δ chosen as above, we have

$$I_2 = \mathsf{E}f[\xi_{\delta u}(t_1), \ldots, \xi_{\delta u}(t_s)] - \mathsf{E}f[\xi_{\delta 0}(t_1), \ldots, \xi_{\delta 0}(t_s)] \to 0$$

as $u \to 0$. Thus the weak convergence as $u \to 0$ of the marginal distributions of processes $\xi_u(t)$ is proved for the particular case under consideration.

We now turn to the proof of Theorem 14 in the general case.

We introduce fields $\tilde{A}_{\delta u}(x, t)$ approximating $A_u(x, t)$, $u \in [0, u_0]$. For this purpose we construct for each $\delta > 0$ in the sphere $\{x : |x| \leq 1/\delta\}$ a δ-net $x_1, x_2, \ldots, x_{n_\delta}$ and a system of functions $g_j(x)$, $j = 1, \ldots, n_\delta$, satisfying the following conditions: $g_j(x) \geq 0$ and $g_j(x) = 0$ for $|x - x_j| \geq \delta$, $\sum_{j=1}^{n_\delta} g_j(x) = 1$, $|x| \leq 1/\delta$, and the functions $g_j(x)$ are continuously differentiable.

Set

$$\tilde{\alpha}_{\delta u}(x, t) = \sum_{j=1}^{n_\delta} g_j(x) \alpha_u(x_j, t),$$

$$\hat{\beta}_{\delta u}(x, t) = \sum_{j=1}^{n_\delta} g_j(x) \beta_u(x_j, t),$$

$$\tilde{A}_{\delta u}(x, t) = \int_0^t \tilde{\alpha}_{\delta u}(x, s)\, ds + \tilde{\beta}_{\delta u}(x, t),$$

$$A_{\delta u}(x, t) = \tilde{A}_{\delta u}(x, t) - A_u(x, t) = \int_0^t \alpha_{\delta u}(x, s)\, ds + \beta_{\delta u}(x, t).$$

Introduce stochastic differential equations

$$d\eta_u(t) = \tilde{A}_{\delta u}(\eta_u(t), dt), \qquad \eta_u(0) = \xi_u^0, \qquad t \in [0, T]. \tag{38}$$

Note that if the conditions of Theorem 14 are satisfied, then equation (38) satisfies the conditions stipulated in the particular case above for any fixed δ.

As before, let $f(x_1, \ldots, x_s)$ denote an arbitrary continuous and continuously differentiable function which is bounded and possesses bounded partial derivatives. Set

$$\mathsf{E} f[\xi_u(t_1), \ldots, \xi_u(t_s)] - \mathsf{E} f[\xi_0(t_1), \ldots, \xi_0(t_s)] = J_1 + J_2 + J_3,$$

where

$$J_1 = \mathsf{E}(f[\xi_u(t_1), \ldots, \xi_u(t_s)] - f[\eta_u(t_1), \ldots, \eta_u(t_s)]),$$

$$J_2 = \mathsf{E} f[\eta_u(t_1), \ldots, \eta_u(t_s)] - \mathsf{E} f[\eta_0(t_1), \ldots, \eta_0(t_s)],$$

$$J_3 = \mathsf{E}(f[\eta_0(t_1), \ldots, \eta_0(t_s)] - f[\xi_0(t_1), \ldots, \xi_0(t_s)]).$$

It follows from the particular case of Theorem 14 discussed above that $J_2 \to 0$ for any fixed δ as $u \to 0$. Thus to prove the theorem in the general case it is sufficient to show that $J_1 + J_3 \to 0$ as $\delta \to 0$ uniformly in u. Now for any $\varepsilon > 0$

$$|J_1| + |J_3| \leq C(2\varepsilon + \mathsf{P}\{\sup_{0 \leq t \leq T} |\xi_u(t) - \eta_u(t)| > \varepsilon\} + \mathsf{P}\{\sup_{0 \leq t \leq T} |\xi_0(t) - \eta_0(t)| > \varepsilon\}),$$

where C is a constant which depends on the function $f(x_1, \ldots, x_s)$ only.

We show that Lemma 5 is applicable to equations (35) and (38). To do this we observe that for $|x|\leq N_0<1/\delta$

$$|\alpha_{\delta u}(x,t)|=|\tilde{\alpha}_{\delta u}(x,t)-\alpha_u(x,t)|$$

$$\leq \sum_{j=1}^{n_\delta} g_j(x)|\alpha_u(x_j,t)-\alpha_u(x,t)|$$

$$\leq \sum_{j:|x_j-x|<\delta} g_j(x)|\alpha_u(x_j,t)-\alpha_u(x,t)|\leq \delta\lambda_{N_0}(t).$$

Analogously,

$$\mathsf{E}\{|\Delta\beta_{\delta u}(x,t)|^2|\mathfrak{F}_t\}\leq \sum_{j=1}^{n_\delta} g_j(x)\sum_{j=1}^{n_\delta}\mathsf{E}\{g_j(x)|\,\Delta\beta_u(x_j,t)-\Delta\beta_u(x,t)|^2|\mathfrak{F}_t\}$$

$$\leq \delta^2\mathsf{E}\{\textstyle\int_t^{t+\Delta t}\lambda_{N_0}^2(s)\,ds\,|\mathfrak{F}_t\}.$$

Moreover, it is easy to verify that

$$|\tilde{\alpha}_{\delta u}(x,t)|\leq 2(1+|x|)\lambda_0(t),$$

$$\mathsf{E}\{|\Delta\tilde{\beta}_{\delta u}(x,t)|^2|\mathfrak{F}_t\}\leq 2(1+|x|^2)\mathsf{E}\{\textstyle\int_t^{t+\Delta t}\lambda_0^2(s)\,ds\,|\mathfrak{F}_t\}.$$

The bounds obtained do not depend on u and we observe that the conditions of Lemma 5 are indeed satisfied; moreover,

$$\sup_{\substack{t\in[0,T]\\|x|\leq N}}\gamma_{\delta u}^2(x,t)\leq\delta^2\sup_{t\in[0,T]}\lambda_N^2(t).$$

Hence there exists δ_0 independent of u such that $|J_1|+|J_3|<4C\varepsilon$ for $\delta<\delta_0$. □

Example. *Oscillations with a small nonlinearity.* Consider an oscillation equation with small nonlinear terms

$$\frac{d^2x}{dt^2}+\omega^2x=\varepsilon f_1\left(x,\frac{dx}{dt}\right)+\sqrt{\varepsilon}f_2\left(x,\frac{dx}{dt}\right)\dot{w}(t), \tag{39}$$

where $x=x(t)$, $f_1(x,y)$ and $f_2(x,y)$ are scalar functions, ε is a small parameter, and $w(t)$ is a Wiener process.

Equation (39) should be interpreted as a system of two stochastic differential equations of the form

$$\begin{aligned}d\dot{x}&=(-\omega^2x+\varepsilon f_1(x,\dot{x}))\,dt+\sqrt{\varepsilon}f_2(x,\dot{x})\,dw(t),\\ dx&=\dot{x}\,dt.\end{aligned} \tag{40}$$

We shall apply the standard methodology for investigating nonlinear oscillations to system (40). Introduce the change of variables

$$x = a\cos\psi,\qquad \dot{x} = -a\omega\sin\psi,\qquad \psi = \omega t + \theta$$

or

$$a = \sqrt{x^2 + \frac{1}{\omega^2}\dot{x}^2},\qquad \psi = -\arctan\frac{\dot{x}}{\omega x}.$$

Now we utilize Itô's formula to obtain an equation for quantities a and θ. The following relations will result:

$$da = \varepsilon\left(-\frac{\sin\psi}{\omega}\tilde{f}_1(a,\theta) + \frac{\cos^2\psi}{2\omega^2 a}\tilde{f}_2^2(a,\theta)\right)dt - \frac{\sqrt{\varepsilon}\sin\psi}{\omega}\tilde{f}_2(a,\theta)\,dw,$$

$$d\theta = \varepsilon\left(-\frac{\cos\psi}{\omega a}\tilde{f}_1(a,\theta) + \frac{\sin 2\psi}{2\omega^3 a^2}\tilde{f}_2^2(a,\theta)\right)dt - \frac{\sqrt{\varepsilon}\cos\psi}{2\omega a}\tilde{f}_2(a,\theta)\,dw.$$

These equations differ from the equations which would have been obtained if the process $w(t)$ were differentiable by the presence of the additional terms $\varepsilon(\cos^2\psi/2\omega a)\tilde{f}_2^2(a,\theta)$ in the first equations and $\varepsilon(\sin 2\psi/2\omega^3 a^2)\tilde{f}_2^2(a,\theta)$ in the second. Here

$$\tilde{f}_i(a,\theta) = f_i(a\cos\psi, -\omega a\sin\psi),\qquad i = 1, 2.$$

Introduce a new time by replacing $t \to t/\varepsilon$ and set

$$a_\varepsilon(t) = a\left(\frac{t}{\varepsilon}\right),\qquad \theta_\varepsilon(t) = \theta\left(\frac{t}{\varepsilon}\right),\qquad \xi_\varepsilon(t) = (a_\varepsilon(t), \theta_\varepsilon(t)).$$

Then

$$d\xi_\varepsilon(t) = A_\varepsilon(\xi_\varepsilon(t), dt),$$
$$A_\varepsilon^i(a,\theta,t) = \int_0^t \alpha_\varepsilon^i(a,\theta,s)\,ds + \beta_\varepsilon^i(a,\theta,t),\qquad i = 1, 2,$$

where

$$\alpha_\varepsilon^1(a,\theta,t) = -\frac{\sin\left(\frac{\omega}{\varepsilon}t+\theta\right)}{\omega}\tilde{f}_1\left(a, \frac{\omega}{\varepsilon}t+\theta\right) + \frac{\cos^2\left(\frac{\omega}{\varepsilon}t+\theta\right)}{2\omega^2 a}\tilde{f}_2^2\left(a, \frac{\omega}{\varepsilon}t+\theta\right),$$

$$\alpha_\varepsilon^2(a,\theta,t) = -\frac{\cos\left(\frac{\omega}{\varepsilon}t+\theta\right)}{\omega a}\tilde{f}_1\left(a, \frac{\omega}{\varepsilon}t+\theta\right) + \frac{\sin^2\left(\frac{\omega}{\varepsilon}t+\theta\right)}{2\omega^3 a^2}\tilde{f}_2^2\left(a, \frac{\omega}{\varepsilon}t+\theta\right),$$

$$\beta_\varepsilon^1(a,\theta,t) = -\frac{\sqrt{\varepsilon}}{\omega}\int_0^{t/\varepsilon}\tilde{f}_2(a, \omega s+\theta)\sin(\omega s+\theta)\,dw(s),$$

$$\beta_\varepsilon^2(a,\theta,t) = -\frac{\sqrt{\varepsilon}}{\omega a}\int_0^{t/\varepsilon}\tilde{f}_2(a, \omega s+\theta)\cos(\omega s+\theta)\,dw(s).$$

It is easy to verify that as $\varepsilon \to 0$ under quite general assumptions on functions $f_1(x, \dot{x})$ and $f_2(x, \dot{x})$,

$$\int_0^t \alpha_\varepsilon^1(a, \theta, s)\, ds \to t\bar{f}_1(a),$$
$$\int_0^t \alpha_\varepsilon^2(a, \theta, s)\, ds \to t\bar{f}_2(a),$$

where

$$\bar{f}_1(a) = \frac{1}{2\pi} \int_0^{2\pi} \left[-f_1(a \cos \psi, -a\omega \sin \psi) \frac{\sin \psi}{\omega} + \tilde{f}_2^2(a \cos \psi, -a\omega \sin \psi) \frac{\cos^2 \psi}{2\omega^2 a} \right] d\psi,$$

$$\bar{f}_2(a) = \frac{1}{2\pi} \int_0^{2\pi} \left[-f_1(a \cos \psi, -a\omega \sin \psi) \frac{\cos \psi}{\omega a} + \tilde{f}_2^2(a \cos \psi, -a\omega \sin \psi) \frac{\sin 2\psi}{2\omega^3 a^2} \right] d\psi.$$

On the other hand $\beta_\varepsilon(a, \theta, t)$ is a Gaussian field with independent (time) increments and, moreover,

$$\mathsf{E}(\beta_\varepsilon^1(a, \theta, t))^2 = \frac{\varepsilon}{\omega} \int_0^{t/\varepsilon} \tilde{f}_2^2(a, \omega s + \theta) \sin^2(\omega s + \theta)\, ds,$$

$$\mathsf{E}\beta_\varepsilon^1(a, \theta, t)\beta_\varepsilon^2(a, \theta, t) = \frac{\varepsilon}{\omega a} \int_0^{t/\varepsilon} \tilde{f}_2^2(a, \omega s + \theta) \sin(\omega s + \theta) \cos(\omega s + \theta)\, ds,$$

$$\mathsf{E}(\beta_\varepsilon^2(a, \theta, t))^2 = \frac{\varepsilon}{\omega^2 a^2} \int_0^{t/\varepsilon} \tilde{f}_2^2(a, \omega s + \theta) \cos^2(\omega s + \theta)\, ds.$$

The correlation matrix of the field $\beta_\varepsilon(a, \theta, t)$ converges to the following limit as $\varepsilon \to 0$ $(h > 0)$:

$$\lim R_\varepsilon(t, t+h) = \lim R_\varepsilon(t, t) = t \begin{pmatrix} b_{11}(a) & b_{12}(a) \\ b_{12}(a) & b_{22}(a) \end{pmatrix},$$

where

$$b_{11}(a) = \frac{1}{2\pi\omega} \int_0^{2\pi} \tilde{f}_2^2(a \cos \psi, -\omega a \sin \psi) \sin^2 \psi\, d\psi,$$

$$b_{12}(a) = \frac{1}{2\pi\omega a} \int_0^{2\pi} \tilde{f}_2^2(a \cos \psi, -\omega a \sin \psi) \sin \psi \cos \psi\, d\psi,$$

$$b_{22}(a) = \frac{1}{2\pi\omega^2 a^2} \int_0^{2\pi} \tilde{f}_2^2(a \cos \psi, -\omega a \sin \psi) \cos^2 \psi\, d\psi.$$

The matrix

$$B(a)=\{b_{ik}(a)\}$$

is symmetric and nonnegative-definite. We now construct a nonnegative-definite and symmetric matrix

$$\sigma(a)=\{\sigma_{ik}\}, \qquad i, k=1, 2,$$

such that

$$\sigma^2(a)=B(a).$$

Marginal distributions of the field $\beta_\varepsilon(a, t)$ converge weakly to the marginal distributions of the field

$$\beta_0(a, t)=\{\beta_0^1(a, t), \beta_0^2(a, t)\}$$

which can be defined by means of the relations:

$$\beta_0^1(a, t)=\sigma_{11}(a)w_1(t)+\sigma_{12}(a)w_2(t),$$
$$\beta_0^2(a, t)=\sigma_{12}(a)w_1(t)+\sigma_{22}(a)w_2(t).$$

Here $w_1(t)$ and $w_2(t)$ are two independent Wiener processes.

Thus if the remaining conditions of Theorem 14 dealing with the regularity properties of functions $f_1(x, \dot{x})$ and $f_2(x, \dot{x})$ are fulfilled, one can assert the following.

A solution of equation (39) *can be represented as*

$$x\left(\frac{t}{\varepsilon}\right)=a_\varepsilon(t)\cos\left(\frac{\omega t}{\varepsilon}+\theta(t)\right),$$

$$\dot{x}\left(\frac{t}{\varepsilon}\right)=-\omega a_\varepsilon(t)\sin\left(\frac{\omega t}{\varepsilon}+\theta_\varepsilon(t)\right),$$

and, moreover, the distributions of the random processes $(a_\varepsilon(t), \theta_\varepsilon(t))$ *are weakly convergent as* $\varepsilon \to 0$ *to the measure associated with the process* $(\bar{a}(t), \bar{\theta}(t))$ *which is a solution of the stochastic differential equation*

$$d\bar{a}=\bar{f}_1(\bar{a})\,dt+\sigma_{11}(\bar{a})\,dw_1+\sigma_{12}(\bar{a})\,dw_2,$$
$$d\theta=\bar{f}_2(\bar{a})\,dt+\sigma_{12}(\bar{a})\,dw_1+\sigma_{22}(\bar{a})\,dw_2.$$