UDC 62-50

ON THE APPROXIMATE SYNTHESIS OF THE OPTIMAL CONTROL OF STOCHASTIC QUASILINEAR SYSTEMS WITH AFTEREFFECT

PMM Vol. 42, No. 6, 1978, pp. 978-988

V. B. KOLMANOVSKII and L. E. SHAIKHET
(Moscow and Donetsk)

(Received September 27, 1977)

Control problems for quasilinear deterministic systems without time lag were analyzed in [1, 2]. In the present paper the control of quasilinear stochastic systems, whose theory has been presented in [3 —6], is studied. The approximate synthesis of the control of stochastic systems with aftereffect is of importance since the construction of their exact optimal control is successful only in exceptional cases [7, 8]. In the paper an approximate optimal control synthesis algorithm is proposed and a method for obtaining error bounds, different from ones previously obtained [9, 10], is developed.

1. Let $\{Q, \sigma, P\}$ be a fixed probability space; $\{Q_t, t \geqslant 0\}$ be a monotonically decreasing family of σ-algebras, $Q_t \subset \sigma$, $W(t) = (W_1(t), \ldots, W_N(t))$ be an N-dimensional standard Wiener process; $\nu^\circ(t, A)$ be a centered Poisson measure with parameter $t\Pi(A)$; the process $W(t)$ and the measure $\nu^\circ(t, A)$ be mutually independent and Q_t-measurable when $t \geqslant 0$. The measure $\Pi(A)$ is defined on Borel sets in Euclidean space R^n; H_0 is the set of deterministic functions $\varphi(s)$ $(-h \leqslant s \leqslant 0)$ with values in R^n, having limits from the left and also being right-continuous when $s < 0$. The norm in H_0 is defined by the equality

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

The functionals encountered later on in the paper, specified on $[0, T] \times H_0$, are reckoned to be measurable relative to the σ-algebra of Borel sets of space $[0, T] \times H_0$. By θ_t, $0 \leqslant t \leqslant T$, we denote the family of operators associating the function $\theta_t\xi = \xi(t + s)$ with an arbitrary function $\xi(s_1)$, $-h \leqslant s_1 \leqslant T$. Here s ranges the values $-h \leqslant s \leqslant 0$ for each fixed t.

Our purpose in the paper is to construct an approximate optimal control and to estimate the error for a stoohastic system of the form

$$d\xi(t) = (\varepsilon f(t, \theta_t\xi) + B(t)u)\,dt + d\eta(t) \tag{1.1}$$

$$d\eta(t) = \sum_{r=1}^{N} b_r(t)\,dW_r(t) + \int_{R^n} C(z, t)\,\nu^\circ(dt, dz), \quad \theta_0\xi = \varphi_0 \in H_0$$

($\xi(t) \in R^n$ is the phase vector and $u \in R^l$ is the control). The initial condition φ_0 and the number T are prescribed and $\varepsilon \geqslant 0$ is a small parameter. The functional $f(t, \varphi)$ is measurable and defined on $[0, T] \times H_0$. It is assumed that a function $r(t, \tau)$, nonnegative and nondecreasing in τ, exists, for which the inequalities

$$|f(t,\varphi)|^2 \leqslant a_0 + \int_0^h |\varphi(-\tau)|^2 d_\tau r(t,\tau) \tag{1.2}$$

$$|f(t,\varphi) - f(t,\psi)|^2 \leqslant \int_0^h |\varphi(-\tau) - \psi(-\tau)|^2 d_\tau r(t,\tau)$$

$$\sup_{0\leqslant t\leqslant T} \int_0^h d_\tau r(t,\tau) < \infty$$

are valid. The $n \times l$ -matrix $B(t)$ and the n -dimensional vectors $b_r(t)$ ($r = 1,\ldots, N$) and $C(z, t)$ ($z \in R^n$) are measurable and bounded in $[0, T]$. We note that the system

$$d\xi_1(t) = (A(t)\xi_1(t) + \varepsilon f(t, \theta_t\xi_1) + B(t)u)\,dt + d\eta(t)$$

is easily reduced to a system of form (1.1) by the change of variables $\xi_1(t) = Z(t)\xi(t)$. Here $Z(t)$ is a solution of the matrix differential equation $Z^\cdot = AZ$ with initial condition $Z(0) = I$, where I is the unit matrix.

Let D be the class of functionals $V(t, \varphi)$ in $[0, T] \times H_0$, such that for any function $\varphi(\tau)$, fixed for $-h \leqslant \tau < 0$, and for an arbitrary vector $x = \varphi(0)$ the function $V_\varphi(t, x) = V(t, \varphi)$ is twice continuously differentiable in x and has a continuous derivative in t for almost all t from $[0, T]$. With system (1.1) we connect an integro-differential operator L_u defined on D and having the form

$$L_u V(t,\varphi) = L_0 V_\varphi(t,x) + (\varepsilon f(t,\varphi) + B(t)u)' \nabla V_\varphi(t,x)$$

$$L_0 V_\varphi(t,x) = \frac{\partial V_\varphi(t,x)}{\partial t} + \frac{1}{2}\sum_{r=1}^N b_r'(t)\nabla^2 V_\varphi(t,x)\, b_r(t) +$$

$$\int_{R^n} [V_\varphi(t, x + C(z,t)) - V_\varphi(t,x) - C'(z,t)\nabla V_\varphi(t,x)]\, \Pi(dz)$$

Here the prime is the sign of transposition and $\partial V_\varphi/\partial t$ is the partial derivative in t, while ∇V_φ and $\nabla^2 V_\varphi$ are, respectively, the vector of first derivatives and the matrix of second derivatives with respect to $x = \varphi(0)$ of the function $V_\varphi(t, x) = V(t, \varphi)$ with function $\varphi(\tau)$ fixed on $-h \leqslant \tau < 0$.

An arbitrary control u is said to be admissible if under this control system (1.1) has a solution (not necessarily unique) and the functional $G(0, u)$, where

$$G(t,u) = M_\varphi \left\{ H(\xi^u(T)) + \int_t^T F(s, \theta_s\xi, u(s))\, ds \right\}$$

is finite. Here M_φ is the mean, computed under the condition that the trajectory of process $\xi^u(s)$ on $[t - h, t]$ is fixed and coincides with a specified function $\varphi \in H_0$. The functions $H(x)$ and $F(t, \varphi, u)$ are prescribed and are nonnegative. Let U be the class of admissible controls. The optimal control problem consists in choosing from U the control u under which functional $G(0, u)$ is minimal. In general, the optimal control depends upon time t and upon the trajectory $\theta_t\xi^u$ of the controlled process up to instant t, i.e., has the form of a functional $u(t, \varphi)$, measurable on $[0, T] \times H_0$, such that $u(t) = u(t, \theta_t\xi) \in U$. The following theorem is valid [7].

Theorem. Let there exist a functional $V(t, \varphi) \in D$ and a control $v = v(t, \varphi) \in U$ satisfying the conditions

$$L_u V(t, \varphi) + F(t, \varphi, u) \geqslant 0, \quad L_v V(t, \varphi) + F(t, \varphi, u(t, \varphi)) = 0, \quad V(T, \varphi) = H(\varphi(0)) \tag{1.3}$$

for almost all $t \in [0, T]$, for all $\varphi \in H_0$ and for all $u \in U$. Then control v is optimal in the sense of performance index $G(t, u)$, and the relation

$$G(t, v) = \inf_{u \in V} G(t, u) = V(t, \varphi)$$

is valid for all $t \geqslant 0$ and $\varphi \in H_0$.

In what follows it is assumed that

$$F(t, \varphi, u) = \varphi'(0) F(t) \varphi(0) + u'N(t) u, \quad H(x) = x'Hx$$

where $N(t)$ and $F(t)$ are measurable and bounded, $N(t)$ is uniformly positive definite and $F(t)$ and H are nonnegative-definite matrices. Let $V(t, \varphi) = V_\varphi(t, x)$ be the minimal value of the performance index under the initial condition $\theta_t\xi = \varphi$. Conditions (1.3) can then be combined into one relation that is an analog of Bellman's equation for the problem being analyzed

$$\inf_{u \in V} [L_0 V_\varphi(t, x) + (\varepsilon f(t, \varphi) + B(t) u)' \nabla V_\varphi(t, x) + x'F(t) x + u'N(t) u] = 0, \quad x = \varphi(0)$$

Whence it follows that $V_\varphi(t, x)$ is determined by the equation

$$L_0 V_\varphi(t, x) + \varepsilon f'(t, \varphi) \nabla V_\varphi(t, x) + x'F(t) x = {}^1/_4 \nabla V_\varphi'(t, x) B_1(t) \nabla V_\varphi(t, x) \tag{1.4}$$
$$V_\varphi(T, x) = x'Hx, \quad x = \varphi(0), \quad B_1 = BN^{-1}B'$$

The optimal control $v(t, \varphi)$ equals

$$v(t, \varphi) = -{}^1/_2 N^{-1}(t) B'(t) \nabla V(t, \varphi)$$

2. It is well known [11] that when $\varepsilon = 0$ Eq. (1.4) has an exact solution of form $V_0(t, \varphi) = \varphi'(0) P(t) \varphi(0) + P_1(t)$. Here the matrices P and P_1 are bounded, nonnegative-definite and depend only on the parameters of system (1.1) and of the performance index. When $\varepsilon = 0$ the optimal control is

$$u_0(t, \varphi) = u_0(t, \varphi(0)) = -N^{-1}(t) B'(t) P(t) \varphi(0)$$

Let us show that this control is the zeroth approximation to the optimal one, i.e., yields an error of the order of ε in the performance index.

We introduce the following notation ξ_ε^u is the solution of system (1.1) with $\varepsilon > 0$ and control u; ξ_0^u is the solution of system (1.1) with $\varepsilon = 0$ and control u; u and v are, respecitvely, the control and the optimal control in system (1.1) with $\varepsilon > 0$. We assume

$$I_\varepsilon(u) = M_{\varphi_0}\left[\xi'(T)H\xi(T) + \int_0^T (\xi'(s)F(s)\xi(s) + u'(s)N(s)u(s))\,ds\right]$$

$$V_\varepsilon(\varphi_0) = I_\varepsilon(v), \quad V_0(\varphi_0) = I_0(u_0)$$

Choosing the measurable random processes $a(t) = u_0(t, \xi_0^{u_0}(t))$ and $b(t) = v(t, \theta_t \xi_\varepsilon^v)$ as the controls in Q_t, we conclude that

$$V_\varepsilon(\varphi_0) = \inf_{u \in U} I_\varepsilon(u) \leqslant I_\varepsilon(a) = V_0(\varphi_0) + [I_\varepsilon(a) - I_0(a)]$$
$$V_0(\varphi_0) = \inf_{u \in U} I_0(u) \leqslant I_0(b) = V_\varepsilon(\varphi_0) + [I_0(b) - I_\varepsilon(b)]$$

Consequently,

$$|V_0(\varphi_0) - V_\varepsilon(\varphi_0)| \leqslant \max[|I_0(a) - I_\varepsilon(a)|, \; |I_0(b) - I_\varepsilon(b)|] \tag{2.1}$$

Further, using the arguments used in [6], we can show that for any Q_t-measurable random process $\gamma(t)$ for which

$$\int_0^T M_{\varphi_0}|\gamma(t)|^2\,dt \leqslant C(1 + \|\varphi_0\|^2) \equiv C^\circ \tag{2.2}$$

there holds the inequality

$$M_{\varphi_0}\{\sup_{0 \leqslant t \leqslant T} |\xi_\varepsilon^\gamma(t)|^2\} \leqslant C^\circ \tag{2.3}$$

Here and everywhere subsequently C stands for certain distinct positive constants depending on the control problem's parameters but not on the initial condition of system (1.1). The inequality

$$M_{\varphi_0}\{\sup_{0 \leqslant t \leqslant T} |\xi_\varepsilon^\gamma(t) - \xi_0^\gamma(t)|^2\} \leqslant \varepsilon^2 C^\circ \tag{2.4}$$

can be proved similarly. From (2.3) and (2.4) it follows that

$$|I_\varepsilon(\gamma) - I_0(\gamma)| = \Big| M_{\varphi_0}\Big[(\xi_\varepsilon^\gamma(T) - \xi_0^\gamma(T))' H (\xi_\varepsilon^\gamma(T) + \xi_0^\gamma(T)) + \int_0^T (\xi_\varepsilon^\gamma(s) - \xi_0^\gamma(s))' F(s)(\xi_\varepsilon^\gamma(s) + \xi_0^\gamma(s))\,ds \Big| \leqslant C[M_{\varphi_0}\{\sup_{0 \leqslant t \leqslant T} |\xi_\varepsilon^\gamma(t) + \xi_0^\gamma(t)|\} \times M_{\varphi_0}\{\sup_{0 \leqslant t \leqslant T} |\xi_\varepsilon^\gamma(t) - \xi_0^\gamma(t)|\}]^{1/2} \leqslant \varepsilon C^\circ \tag{2.5}$$

Let us show that controls $a(t)$ and $b(t)$ satisfy an inequality of form (2.2). Let

$$q = \inf_{t \in [0, T], |u|=1} u' N(t) u$$

Since $N(t)$ is positive definite uniformly in t, then $q > 0$. Hence

$$\int_0^T M_{\varphi_0}|a(t)|^2\,dt = \int_0^T M_{\varphi_0}|u_0(t, \xi_0^{u_0}(t))|^2\,dt \leqslant \frac{1}{q}\int_0^T M_{\varphi^\circ} \times$$

$$u_0'(t, \xi_0^{u_0}(t)) N(t) u_0(t, \xi_0^{u_0}(t))\, dt \leqslant \frac{1}{q} I_0(u_0) \leqslant$$
$$\frac{1}{q} I_0(0) \leqslant CM_{\varphi_0}\{\sup_{0\leqslant t\leqslant T} |\xi_0^\circ(t)|^2\} \leqslant C^\circ$$

Similarly for $b(t)$. Consequently, the estimates of form (2.3) –(2.5) are valid for controls $a(t)$ and $b(t)$. Hence from (2.1) follows

$$|V_0(\varphi_0) - V_\varepsilon(\varphi_0)| \leqslant \varepsilon C^\circ$$

Estimates of form (2.3) –(2.5) can be obtained similarly for the control u_0, making use of its linearity with respect to $\xi(t)$. Consequently,

$$|J_\varepsilon(u_0) - J_0(u_0)| \leqslant \varepsilon C^\circ$$

Thus

$$0 \leqslant J_\varepsilon(u_0) - J_\varepsilon(v) \leqslant |J_\varepsilon(u_0) - J_0(u_0)| + |V_0(\varphi_0) - V_\varepsilon(\varphi_0)| \leqslant \varepsilon C^\circ \qquad (2.6)$$

Q.E.D.

We note that in this proof we made essential use of a certain auxiliary controlled system for which control u_0 is optimal and functional $V_0(t, \varphi) = V_\varphi^\circ(t, \varphi(0))$ is Bellman's function. We use this proof method later on. To be precise, at each step an auxiliary controlled system is constructed, for which the next approximation u_k to the optimal control is itself optimal and some functional $Q_k(t, \varphi) = Q_\varphi^k(t, \varphi(0))$ is Bellman's function. For $k \geqslant 1$ we choose Eq. (1.1) as the auxiliary controlled system, while the functional Q_k differs from the Bellman's function for the original problem by an amount of order ε^{k+1}. The need for bounds of form (2.4) is now eliminated. In addition, no assumptions are made on the Bellman equation for the original problem when proving the error estimates.

To illustrate what we have said we present another proof of estimate (2.6) different from the preceding one. It is precisely this proof that will be generalized later on to higher-order successive approximations. First of all we note that the equations

$$L_0 W_\varphi(t, x) + \varepsilon f'(t, \varphi)(\nabla W_\varphi(t, x) - 2P(t)x) + x'F(t)x = {}^1/_4 \nabla W_\varphi'(t, x) B_1(t) \nabla W_\varphi(t, x) \qquad (2.7)$$
$$W_\varphi(T, x) = x'Hx, \quad x = \varphi(0)$$

defines Bellman's function for the optimal control problem with equation of motion (1.1) and performance index I_ε of the form

$$I_\varepsilon(u) = J_\varepsilon(u) - 2\varepsilon M_{\varphi_0}\int_0^T f'(s, \theta_s, \xi_\varepsilon^u) P(s) \xi_\varepsilon^u(s)\, ds \qquad (2.8)$$

From relations (1.4) with $\varepsilon = 0$ it follows that the functional $V_0(t, \varphi) = \varphi'(0)P(t)\varphi(0) + P_1(t)$ is a solution of Eq. (2.7) for any $\varepsilon > 0$. Similarly to [9] it can be shown that the solution of Eq. (2.7) is unique for sufficiently small ε. Thus, V_0 is

Bellman's function and u_0 is the optimal control for problem (1.1), (2.8).

Now let $W_\varepsilon(\varphi_0) = I_\varepsilon(u_0)$ and $C(t) = u_0(t, \xi_\varepsilon^{u_0}(t))$. As before, we have

$$0 \leqslant J_\varepsilon(u_0) - J_\varepsilon(v) \leqslant | J_\varepsilon(u_0) - I_\varepsilon(u_0) | + \\ | W_\varepsilon(\varphi_0) - V_\varepsilon(\varphi_0) | \leqslant | J_\varepsilon(u_0) - I_\varepsilon(u_0) | + \\ \max [| J_\varepsilon(C) - I_\varepsilon(C) |, | J_\varepsilon(b) - I_\varepsilon(b) |]$$

We assume

$$\alpha(u) = M_{\varphi_0}\{\sup_{0\leqslant t\leqslant T} | \xi_\varepsilon^u(t) |^2\}$$

From (2.8) and (1.2), for any $u \in U$ follows

$$| I_\varepsilon(u) - J_\varepsilon(u) | \leqslant \varepsilon C [\alpha(u)(1 + \| \varphi_0 \|^2 + \\ \alpha(u))]^{1/2} \leqslant \varepsilon C (1 + \| \varphi_0 \|^2 + \alpha(u))$$

Hence (2.6) follows from the fact that estimates (2.3) hold for controls u_0, C and b.

3. We now pass on to the higher ($k \geqslant 1$) approximations to the optimal control. The algorithm for constructing these approximations is as follows. We represent functional V as the series

$$V(t, \varphi) = V_0(t, \varphi) + \varepsilon V_1(t, \varphi) + \varepsilon^2 V_2(t, \varphi) + \ldots$$

where $V_i \in D$. We substitute this expansion into Eq. (1.4) and we equate the coefficients of like powers of ε to zero. Allowing for Eq. (1.4) for $V_0(t, \varphi)$ with $\varepsilon = 0$, we obtain that functionals $V_i(t, \varphi) = V_\varphi^i(t, x)$ $(i = 1, 2, \ldots)$ are determined by the recurrence equations

$$L_0 V_\varphi^i(t, x) + f'(t, \varphi)\nabla V_\varphi^{i-1}(t, x) = \frac{1}{4}\sum_{j=0}^{i} \nabla V_\varphi^{j\prime}(t, x) B_1(t) \nabla V_\varphi^{i-j}(t, x) \quad (3.1)$$

$$V_\varphi^i(T, x) = 0, \quad x = \varphi(0)$$

Having thus determined $V_i(t, \varphi)$ $(i = 1, \ldots, k)$, we specify the k-th approximation to the optimal control as

$$u_k(t, \varphi) = -{}^1/_2 N^{-1}(t) B'(t) [\nabla V_0(t, \varphi) + \ldots + \varepsilon^k \nabla V_k(t, \varphi)]$$

The effectiveness of the algorithm presented depends upon the ability to compute the functionals $V_i(t, \varphi)$. For $V_0(t, \varphi)$ there is an explicit formula. For $i \geqslant 1$ by virtue of (3.1) we have

$$LV_i(t, \varphi) + S_i(t, \varphi) = 0, \quad V_i(T, \varphi) = 0 \quad (3.2)$$

$$S_i(t, \varphi) = f'(t, \varphi)\nabla V_{i-1}(t, \varphi) - \frac{1}{4}\sum_{j=1}^{i-1} \nabla V_j'(t, \varphi) B_1(t) \nabla V_{i-j}(t, \varphi)$$

In (3.2) L denotes the generating operator of the stochastic differential equation without time lag

$$d\xi(s) = -B_1(s) P(s) \xi(s)\, ds + d\eta(s), \quad s \in [t, T], \quad \theta_t \xi = \varphi \quad (3.3)$$

In deriving (3.2) we used the following indentity:

$$\sum_{j=0}^{i} \nabla V_\varphi^{j'}(t,x) B_1(t) \nabla V_\varphi^{i-j}(t,x) =$$
$$4B_1(t)P(t)x + \sum_{j=1}^{i-1} \nabla V_\varphi^{j'}(t,x) B_1(t) \nabla V_\varphi^{i-j}(t,x)$$

Lemma 1 [7]. Let $V(t,\varphi) \in D$ and L_1 be the generating operator of system

$$d\xi(s) = a(s, \theta_s\xi)\,ds + d\eta(s), \quad \theta_t\xi = \varphi \ (t \leqslant s \leqslant T) \qquad (3.4)$$

Then for any t_1 and t_2

$$M_\varphi V(t_2, \theta_{t_2}\xi) - M_\varphi V(t_1, \theta_{t_1}\xi) = \int_{t_1}^{t_2} M_\varphi L_1 V(s, \theta_s\xi)\,ds \quad (t \leqslant t_1 \leqslant t_2 \leqslant T)$$

From Lemma 1 follows

Lemma 2. Let $V(t,\varphi) \in D$ and for any $t \in [0, T]$

$$L_1 V(t,\varphi) + r(t,\varphi) = 0, \quad V(T,\varphi) = 0$$

where L_1 is the generating operator of system (3.4). Then functional $V(t,\varphi)$ is representable as

$$V(t,\varphi) = M_\varphi \int_t^T r(s, \theta_s\xi)\,ds$$

where $\xi(s)$ is a solution of system (3.4).

For $i = 1$ we write (3.2) as

$$LV_\varphi^1(t,x) + 2f'(t,\varphi)P(t)x = 0 \qquad (3.5)$$
$$V_\varphi^1(T,x) = 0, \quad x = \varphi(0)$$

If V_φ^1 and W_φ^1 are two solutions of Eq. (3.5), then for $R_\varphi^1 = V_\varphi^1 - W_\varphi^1$ it follows [4] from $LR_\varphi^1(t,x) = 0$ and $R_\varphi^1(T,x) = 0$ that $R_\varphi^1(t,x) \equiv 0$, i.e., the solution of Eq. (3.5) is unique. The uniqueness of the solution of Eq. (3.2) for all $i \geqslant 0$ can be proved similarly by mathematical induction. On the basis of Lemma 2, from this and from relations (3.2) and (3.3) follows the representation

$$V_i(t,\varphi) = M_\varphi \int_t^T S_i(\tau, \theta_\tau\xi)\,d\tau \qquad (3.6)$$

Here $\xi(\tau)$ is a solution of Eq. (3.3), and for $\tau \leqslant t$ the process $\xi(\tau)$ is determined by the equality $\xi(\tau) = \varphi(\tau)$, where $\varphi(\tau)$ is a prescribed deterministic function.

In some cases the computation of the right hand side of (3.6) reduces to a quadrature. For instance, let $f(t, \theta_t\xi) = f(t, \xi(t-h))$, where $h \geqslant 0$ is a specified constant, and let $p(t, x, s, y)$ be the transfer probability density of the process specified by Eq. (3.3). Then when $i = 1$ representation (3.6) can be written as $(0 \leqslant t + h \leqslant T)$

$$V_1(t, \varphi) = 2 \int_t^{t+h} \int_{R^n} f'(s, \varphi(s-h)) P(s) y p(t, \varphi(t), s, y)\, dy\, ds +$$

$$2 \int_{t+h}^{T} \int_{R^n} \int_{R^n} f'(s, z) P(s) y p(t, \varphi(t), s-h, z) p(s-h, z, s, y)\, dz\, dy\, ds$$

We note that the density $p\ (t, x, s, y)$ occurring in the last formula also can be computed in explicit analytic form for certain systems of form (3.3). Estimates of the error in the functional, admissible under control u_k for $k \geqslant 1$, are established analogously. Therefore, we give a detailed proof of the error estimate only for the first-approximation control u_1 so important from the practical point of view.

As already noted, the main idea of the proof is the construction of an auxiliary control problem for which u_1 is the optimal control and functional Q_1, equalling $Q_1 = V_0 + \varepsilon V_1$, is Bellman's function. Let us construct the auxiliary control problem. We add Eq. (1.4), in which $\varepsilon = 0$, and Eq. (3.1) multiplied by ε, in which $i = 1$. Then for functional $Q_1\ (t, \varphi) = Q_\varphi^1\ (t, x)$ we obtain the relations

$$L_0 Q_\varphi^1(t, x) + x'F(t)x + \varepsilon f'(t, \varphi) Q_\varphi^1(t, x) + \frac{\varepsilon^2}{4} \nabla V_\varphi^{1\prime}(t, x) B_1(t) \nabla V_\varphi^1(t, x) -$$

$$\varepsilon^2 f'(t, \varphi) \nabla V_\varphi^1(t, x) =$$

$$\frac{1}{4} \nabla Q_\varphi^{1\prime}(t, x) B_1(t) \nabla Q_\varphi^1(t, x), \qquad Q_\varphi^1(T, x) = x'Hx, \qquad x = \varphi(0)$$

Consequently, u_1 is the optimal control and $Q_1\ (t, \varphi)$ is Bellman's function for the control problem with equation of motion (1.1) and performance index

$$I_\varepsilon^1(u) = J_\varepsilon(u) + \varepsilon^2 M_{\varphi_0} \int_0^T \delta_1(s, \theta_s \xi_\varepsilon^u)\, ds$$

$$\delta_1 = {}^1/_4 \nabla V_1' B_1 \nabla V_1 - f' \nabla V_1$$

Let us assume that functional $V_1\ (t, \varphi)$ satisfies the inequality

$$|\nabla V_1\ (t, \varphi)|^2 \leqslant C\ (1 + \|\varphi\|^2) \tag{3.7}$$

We set

$$\alpha(u) = M_{\varphi_0}\{\sup_{0 \leqslant t \leqslant T} |\xi_\varepsilon^u(t)|^2\}, \cdot \quad u \in U$$

Then, as above, it is easy to establish the inequality

$$|I^1_\varepsilon(u) - J_\varepsilon(u)| \leqslant \varepsilon^2 C\ (1 + \|\varphi_0\|^2 + \alpha(u))$$

In addition, for $V_\varepsilon^1\ (\varphi_0) = I_\varepsilon^1\ (u_1)$ and $c_1\ (t) = u_1\ (t, \theta_t \xi_\varepsilon^{u_1})$

$$|V_\varepsilon(\varphi_0) - V_\varepsilon^1(\varphi_0)| \leqslant \max\left[|I_\varepsilon^1(c_1) - J_\varepsilon(c_1)|, |J_\varepsilon^1(b) - J_\varepsilon(b)|\right]$$

Bounds of form (2.3) are fulfilled for controls $b\ (t)$ and $c_1\ (t)$. Consequently,

$$|V_\varepsilon(\varphi_0) - V_\varepsilon^1(\varphi_0)| \leqslant \varepsilon^2 C^\circ$$

Using (3.7) we can show that a bound of form (2.3) is valid for control u_1. Thus

$$0 \leqslant J_\varepsilon(u_1) - J(v) \leqslant |J_\varepsilon(u_1) - I_\varepsilon^1(u_1)| + |V_\varepsilon^1(\varphi_0) - V_\varepsilon(\varphi_0)| \leqslant \varepsilon^2 C^\circ$$

where the constant C° can be estimated in terms of the parameters of the original problem. Thus we have shown that for the original control problem the control u_1 yields an error in the functional of order ε^2. To complete the proof it remains to show that functional V_1 does indeed satisfy condition (3.7). From (3.6) with $i = 1$ we have

$$V_1(t, \varphi) = 2M_\varphi \int_t^T f'(s, \theta_s\xi) P(s) \xi(s)\, ds$$

where $\xi(s)$ is a solution of Eq. (3.3).

Now let $\xi(s)$ be the solution of Eq. (3.3) for $\theta_t\xi = \varphi$ and $\xi_1(s)$ for $\theta_t\xi_1 = \varphi_1$. Then

$$|V_1(t,\varphi) - V_1(t,\varphi_1)|^2 = |M(V_1(t,\varphi) - V_1(t,\varphi_1))|^2 =$$

$$4\left|M\left(M_\varphi \int_t^T f'(s,\theta_s\xi) P(s)\xi(s)\,ds - M_{\varphi_1}\int_t^T f'(s,\theta_s\xi_1) P(s)\xi_1(s)\,ds\right)\right|^2 =$$

$$4\left|M\int_t^T (f'(s,\theta_s\xi) P(s)\xi(s) - f'(s, \theta_s\xi_1) P(s)\xi_1(s))\,ds\right|^2 \leqslant$$

$$8T\left[\int_t^T M((f'(s,\theta_s\xi) - f'(s,\theta_s\xi_1)) P(s)\xi(s))^2 ds + \right.$$

$$\left.\int_t^T M(f'(s, \theta_s\xi_1) P(s)(\xi(s) - \xi_1(s)))^2\, ds\right] \leqslant$$

$$C\left[\int_t^T M|\xi(s)|^2 M|f(s,\theta_1\xi) - f(s,\theta_s\xi_1)|^2 ds + \right.$$

$$\left.\int_t^T M|f(s,\theta_s\xi_1)|^2 M|\xi(s) - \xi_1(s)|^2 ds\right] \leqslant$$

$$C[(1 + \|\varphi\|^2)\|\varphi - \varphi_1\|^2 + (1 + \|\varphi_1\|^2)\|\varphi - \varphi_1\|^2] \leqslant$$

$$C(1 + \|\varphi\|^2 + \|\varphi_1\|^2)\|\varphi - \varphi_1\|^2$$

Let $\varphi_1(s) = \varphi(s)$ when $-h \leqslant s < 0$, $\varphi(0) = x$ and $\varphi_1(0) = x + \Delta x$. Then $\|\varphi - \varphi_1\|^2 = |\Delta x|^2$ and

$$|\nabla V_1(t,\varphi)|^2 = |\nabla V_\varphi^1(t,x)|^2 = \lim_{\Delta x\to 0} \frac{|V_\varphi^1(t,x) - V_\varphi^1(t, x+\Delta x)|^2}{|\Delta x|^2} =$$

$$\lim_{\Delta x\to 0}\frac{|V_1(t,\varphi) - V_1(t,\varphi_1)|^2}{|\Delta x|^2} \leqslant \lim_{\Delta x\to 0} C(1 + \|\varphi\|^2 + \|\varphi_1\|^2) \leqslant$$

$$C(1 + \|\varphi\|^2)$$

whence follows (3.7)

4. For the zeroth and first approximations we have constructed auxiliary control problems with equation of motion (1.1) and performance indices differing from that of the original problems by amounts of the order of ε and ε^2, respectively. Let us construct the auxiliary problem for which u_k is the optimal control and the functional

$$Q_k = V_0 + \varepsilon V_1 + \ldots + \varepsilon^k V_k \tag{4.1}$$

is Bellman's function for all $k \geqslant 1$. We add Eq. (1.4), in which $\varepsilon = 0$, and Eq. (3.1) multiplied by ε^i $(i = 1, \ldots, k)$. In the resulting equality we add and subtract the expression $1/4\ \nabla Q_k' B_1 \nabla Q_k$. As a result of this, with $x = \varphi(0)$ we obtain

$$\begin{gathered} L_0 Q_\varphi^{\ k}(t,x) + x'F(t)x + \varepsilon f'(t,\varphi)\nabla Q_\varphi^{\ k}(t,x) + \\ \frac{1}{4}\Big[\nabla Q_\varphi^{\ k\prime}(t,x)B_1(t)\nabla Q_\varphi^{\ k}(t,x) - \\ \sum_{i=1}^{k}\varepsilon^i \sum_{j=0}^{i}\nabla V_\varphi^{\ j\prime}(t,x)B_1(t)\nabla V_\varphi^{i-j}(t,x)\Big] - \\ \varepsilon^{k+1}f'(t,\varphi)\nabla V_\varphi^{\ k}(t,x) = \frac{1}{4}\nabla Q_\varphi^{\ k\prime}(t,x)B_1(t)\nabla Q_\varphi^{\ k}(t,x) \end{gathered}$$

Using (4.1) we transform the epxression within brackets in the following manner:

$$\begin{gathered} \sum_{m=0}^{k}\sum_{n=0}^{k}\varepsilon^{n+m}\nabla V_m' B_1 \nabla V_n - \sum_{i=0}^{k}\varepsilon^i\sum_{j=0}^{i}\nabla V_j' B_1 \nabla V_{i-j} = \\ \sum_{j=0}^{k}\sum_{i=j}^{k+j}\varepsilon^i \nabla V_j' B_1 \nabla V_{i-j} - \sum_{j=0}^{k}\sum_{i=j}^{k}\varepsilon^i\nabla V_j' B_1 \nabla V_{i-j} = \\ \sum_{j=1}^{k}\sum_{i=k+1}^{k+j}\varepsilon^i\nabla V_j' B_1 \nabla V_{i-j} = \varepsilon^{k+1}\sum_{j=1}^{k}\sum_{i=0}^{j-1}\varepsilon^i \nabla V_j' B_1 \nabla V_{i-j} \end{gathered}$$

Thus, for functional $Q_k(t, \varphi)$ we have obtained the equation

$$\begin{gathered} L_0 Q_\varphi^{\ k}(t,x) + \varepsilon f'(t,\varphi)\nabla Q_\varphi^{\ k}(t,x) + x'F(t)x + \\ \varepsilon^{k+1}\delta_k(t,\varphi) = {}^1/_4 \nabla Q_\varphi^{\ k\prime}(t,x) B_1 \nabla Q_\varphi^{\ k}(t,x) \\ \delta_k = \frac{1}{4}\sum_{j=1}^{k}\sum_{i=0}^{j-1}\varepsilon^i \nabla V_j' B_1 \nabla V_{k+1+i-j} - f'\nabla V_k \end{gathered}$$

Consequently, $Q_k(t, \varphi)$ is Bellman's function for the control problem with equation of motion (1.1) and with the functional

$$I_\varepsilon^{\ k}(u) = J_\varepsilon(u) + \varepsilon^{k+1}M_{\varphi_0}\int_0^T \delta_k(s, \theta_s\xi_\varepsilon^{\ u})\,ds$$

to be minimized.

Representation (3.6) enables us to establish certain sufficient conditions and constraints on $f(t, \varphi)$, under whose fulfilment the functionals $V_t(t, \varphi)$ satisfy a bound of form (3.7). After this, analogously to the preceding, we can prove that

$$0 \leqslant J_\varepsilon(u_k) - J_\varepsilon(v) \leqslant \varepsilon^{k+1} C^\circ$$

In conclusion we note that it is not difficult to generalize the results obtained to systems of form

$$d\xi(t) = (\varepsilon f(t, \theta_t \xi) + B(t) u) dt + d\eta(t) + d\eta_1(t) \xi(t)$$

where $\eta_1(t)$ is a matrix-valued process with independent increments, while the remaining parameters have the same sense as in Eq. (1.1), as well as to systems with noice in the control.

REFERENCES

1. Krasovskii, N. N., Theory of Control of Motion. Moscow, "Nauka", 1967. (See also N. N. Krasovskii, Stability of Motion, Stanford University Press, Stanford, California, 1963).
2. Al'brekht, E. G. On the control of motion of nonlinear systems. Differents. Uravn., Vol. 2, No. 3, 1966.
3. Gikhman, I. I. and Skorokhod, A. V., Introduction to the Theory of Random Processes. Moscow, "Nauka", 1965.
4. Gikhman, I. I. and Skorokhod, A. V., Stochastic Differential Equations, Kiev, "Naukova Dumka", 1968.
5. Gikhman, I. I. and Skorokhod, A. V., Theory of Random Processes, Vol. 3. Moscow, "Nauka", 1975.
6. Gikhman, I. I. and Skorokhod, A. V., Controlled Random Processes. Kiev, "Naukova Dumka", 1977.
7. Kolmanovskii, V. B. and Maizenberg, T. L., Optimal control of stochastic systems with aftereffect. Avtomat. Telemekh., No. 1, 1973.
8. Shaikhet, L. E., Solving a stabilization problem for controlled stochastic time-lag systems by the Liapunov functional method. In: Theory of Random Processes, No. 4, Kiev, "Naukova Dumka", 1976.
9. Kolmanovskii, V. B., On the approximate synthesis of certain stochastic quasilinear systems. Avtomat. Telemekh., No. 1, 1975.
10. Kolmanovskii, V. B., Optimal control of certain quasilinear stochastic systems. PMM Vol. 39, No. 4, 1975.
11. Roitenberg, Ia. N., Certain Problems of the Control of Motion. Moscow, Fizmatgiz, 1963.

Translated by N. H. C.