UDC 62-50

SYNTHESIS OF OPTIMAL STOCHASTIC CONTROL SYSTEMS BY THE METHOD OF SUCCESSIVE APPROXIMATIONS

PMM Vol. 43, No. 1, 1979, pp. 7-16
G. E. KOLOSOV
(Moscow)
(Received March 21, 1978)

An optimal control problem is analyzed for dynamic systems with random perturbations. Control performance is evaluated by the magnitude of the mean value of a functional on the system's motion trajectories. The optimal controls are synthesized by solving a parabolic semilinear partial differential equation (the Bellman equation). A method is suggested for solving this equation (and the synthesis problem), based on the calculation of successive approximations. It is shown that the suboptimal systems constructed in such a way coincide asymptotically with the optimal system. The method suggested can be used for solving the synthesis problem in systems with bounded controls. The method's effectiveness is illustrated by an example.

1. Statement of the problem. We shall analyze dynamic systems whose behaviors can be described by a vector-matrix differential equation of form

$$x^{\cdot} = \bar{b}(x, t) + \bar{c}(t)\, u(t) + \bar{a}(x, t)\, \xi(t), \quad x = (x_1, \ldots, x_n) \tag{1.1}$$

Here x is the system's phase coordinate vector, u is the m-dimensional control vector, $\xi(t)$ in the n-dimensional vector of random perturbations of white noise type with independent components, zero mean and unit intensity, $\bar{b}(x, t)$ is a vector-valued function of the phase coordinates and of time t, and $\bar{c}(t)$ and $\bar{a}(x, t)$ are $(n \times m)$- and $(n \times n)$-matrices whose elements depend on t and on (x, t), respectively. The requirements on the functions $\bar{b}(x, t)$, $\bar{c}(t)$ and $\bar{a}(x, t)$ are given in detail below. Here we merely note that these functions are always assumed to be such that a unique solution $x(t)$ of the stochastic Eq. (1.1) exists for $t \geqslant t_0$, satisfying the condition $x(t_0) = x_0$ and understood in at least the weak sense (see [1]).

The problem is to find a vector-valued control function u with values in some closed bounded domain U, that would ensure the minimum of a certain performance index of the system (the optimality criterion), which we take in the form

$$I[u(t)] = M\left\{\int_0^T \omega_1(x(t), u(t))\, dt + \psi(x(T))\right\} \tag{1.2}$$

Here M denotes the mean, $[0, T]$ is the time interval on which the system's operation is to be analyzed, and ω_1 and ψ are scalar penalty functions whose actual forms are determined by the nature of the problem being solved (see below for the requirements on ω_1 and ψ). The unkown vector-valued function u^* minimizing (1.2) must at each instant t be expressed in terms of the current values of the phase coordinate vector and of time t, i.e., $u^* = u^*(t, x(t))$ (the synethesis problem).

According to the dynamic programming method [2] the solving of the problem posed is equivalent to solving the Bellman equation which for system (1. 1) and criterion (1. 2) has the form [3]

$$-\frac{\partial F}{\partial t} = \bar{b}^T(x, t)\frac{\partial F}{\partial x} + \frac{1}{2} Tr\left[\bar{a}(x, t)\bar{a}^T(x, t)\frac{\partial^2 F}{\partial x\,\partial x^T}\right] + \min_{u \in U}\left[\omega_1(x, u) + u^T \bar{c}^T(t)\frac{\partial F}{\partial x}\right] \quad (1.3)$$

In writing (1. 3) it is assumed that all the stochastic integrals are understood in Itô's sense, that the operator $\partial / \partial x$ is a column vector with components $(\partial / \partial x_1, \ldots, \partial / \partial x_n)$, and that the superscript ^{T} denotes transposition. The function F (henceforth called the loss function) characterizes the operating performance of the optimal system and by definition equals

$$F(x, t) = \min_{\substack{u(s) \in U \\ s \geqslant t}} M\left\{\left[\int_t^T \omega_1(x(s), u(s))\,ds + \psi(x(T))\right] \mid x(t) = x\right\} \quad (1.4)$$

Here $M\{(\cdot) \mid x(t) = x\}$ denotes the averaging of $(\cdot)$ over all possible realizations of the controlled random process $x(s)$ $(s \geqslant t)$, starting from point x when $s = t$. From (1. 4) it follows that

$$F(x, T) = \psi(x) \quad (1.5)$$

Reversing time by means of the substitution $\tau = T - t$, Eq. (1. 3) and condition (1. 5) for function $F(x, \tau)$ are transformed to

$$LF = -\min_{u \in U}\left[\omega_1(x, u) + u^T c^T \frac{\partial F}{\partial x}\right] \quad (1.6)$$

$$L = -\frac{\partial}{\partial \tau} + b_i(x, \tau)\frac{\partial}{\partial x_i} + a_{ij}(x, \tau)\frac{\partial^2}{\partial x_i\,\partial x_j}$$

$$b_i(x, \tau) = \bar{b}_i(x, T - \tau), \quad c(\tau) = \bar{c}(T - \tau)$$

$$F(x, 0) = \psi(x) \quad (1.7)$$

Here $a_{ij}(x, \tau)$ is a general element of matrix $^1/_2\,\bar{a}(x, T - \tau)\cdot\bar{a}^T(x, T - \tau)$; summation from 1 to n is assumed over repeated indices. Taking the quantity $\partial F / \partial x$ as known and carrying out the minimization in (1. 6), we have

$$LF = \Phi_1(x, \tau, \partial F / \partial x) \quad (1.8)$$

The function

$$u^* = \varphi(x, \tau, \partial F / \partial x) \quad (1.9)$$

obtained here, satisfying the condition

$$-\min_{u \in U}\left[\omega_1(x, u) + u^T c^T \frac{\partial F}{\partial x}\right] = -\omega_1(x, \varphi) - \varphi^T c^T \frac{\partial F}{\partial x} = \Phi_1\left(x, \tau, \frac{\partial F}{\partial x}\right)$$

solves the synthesis problem (after Eq. (1.8) has been solved). The forms of functions Φ and φ depend on ω_1 and on domain U.

Example 1. Let $\omega_1(x, u) = \omega(x) + u^T B u$, where B is a positive-definite $(m \times m)$-matrix for any $x \in E_n$ and $\tau \in [0, T]$. Controls u are not bounded $(U = E_m)$. Then $(B^{-1}$ is the matrix inverse to $B)$

$$\varphi = -\frac{1}{2} B^{-1} c^T \frac{\partial F}{\partial x}, \quad \Phi_1 = -\omega(x) + \frac{1}{4} \frac{\partial F}{\partial x^T} c B^{-1} c^T \frac{\partial F}{\partial x}$$

Example 2. Let $\omega_1(x, u) = \omega(x)$ and U be an m-dimensional parallelepiped: $|u_i| \leqslant u_{0i}, i = 1, \ldots, m$. In this case [4]

$$\varphi = -\{u_{01}, \ldots, u_{0m}\} \operatorname{sign}\left(c^T \frac{\partial F}{\partial x}\right), \quad \Phi_1 = -\omega(x) + u_0^T \left| c^T \frac{\partial F}{\partial x} \right|$$

where sign A and $|A|$ are matrices formed from A by replacing a_{ij} by sign a_{ij} and by $|a_{ij}|$ and $\{u_{01}, \ldots, u_{0m}\}$ is a diagonal matrix.

Example 3. Let $\omega_1(x, u) = \omega(x)$ and U be an m-dimensional sphere of radius R. Then

$$\varphi = -Rc^T \frac{\partial F}{\partial x} \left(\frac{\partial F}{\partial x^T} cc^T \frac{\partial F}{\partial x}\right)^{-1/2}, \quad \Phi = -\omega(x) + R\left(\frac{\partial F}{\partial x^T} cc^T \frac{\partial F}{\partial x}\right)^{1/2}$$

2. Successive approximations. If the matrix $\tilde{a}(x, t)$ in (1.1) is nonsingular for any $(x, t) \in E_n \times [0, T]$, then the Bellman Eq. (1.8) is a parabolic semilinear inhomogeneous equation. We rewrite it as

$$LF = -\omega(x) + \Phi(x, \tau, \partial F / \partial x), \quad F(x, 0) = \psi(x) \tag{2.1}$$

We shall seek the solution by the method of successive approximations (Picard's method; see [5]), found recurrently by the following scheme:

$$LF_0 = \omega(x), \quad F_0(x, 0) = \psi(x) \tag{2.2}$$

$$LF_{N+1} = -\omega(x) + \Phi(x, \tau, \partial F_N / \partial x), \quad F_{N+1}(x, 0) = \psi(x) \tag{2.3}$$

$$N = 0, 1, \ldots$$

Here, simultaneously with the functions $F_0, F_1, \ldots$ we find the functions

$$u_N(x, \tau) = \varphi(x, \tau, \partial F_N / \partial x), \quad N = 0, 1, 2, \ldots \tag{2.4}$$

which permit us to synthesize systems close to optimal. Such a method of approximate synthesis has already been used repeatedly, beginning with [6, 7]. Control problems for forces small in magnitude were examined in [7]. Such systems were later called weakly controllable [8]. In this case the nonlinear summand in the Bellman equation contains a small parameter $\Phi = \varepsilon\bar{\Phi}$, which allows us to stay with low order approximations for the synthesis. Error estimates of the method are shown in [9].

The present paper's purpose is to establish convergence conditions for procedure (2.3) when the nonlinear term Φ is not small. In this case the suboptimal system synthesized on the basis of (2.4) can turn out to be near-optimal only for large N; therefore, the need arises to investigate the asymptotic behavior of $F_N(x, \tau)$ and $u_N(x, \tau)$ as $N \to \infty$. Such an investigation was made in [6] for a bounded phase space, using the maximum principle for the solutions of parabolic equations with bounded norm. Below we use estimates of the fundamental solution in unbounded domains, enabling us to establish the convergence of procedure (2.3) as $N \to \infty$ for solutions $F(x, \tau)$ growing unboundedly as $|x| = (x_1^2 + \ldots + x_n^2)^{1/2} \to \infty$.

3. **Investigation of convergence.** We consider an Eq. (2.1) in which operator L is defined in (1.6). Its solution $F(x, \tau)$ and the coefficients $a_{ij}(x, \tau)$ and $b_i(x, \tau)$ of operator L are defined in domain $\Omega = E_n \times [0, T] \equiv \{(x, \tau);\ x \in E_n,\ 0 \leqslant \tau \leqslant T\}$. We take it that everywhere in Ω the matrix $\| a_{ij}(x, \tau) \|$ satisfies the condition of uniform parabolicity of L, i.e., everywhere in Ω, for any real vector χ

$$\lambda |\chi|^2 \leqslant \chi^T a(x, \tau) \chi \leqslant \bar{\lambda} |\chi|^2 \tag{3.1}$$

where λ and $\bar{\lambda}$ are some positive constants. We assume, in addition, that $a_{ij}(x, \tau)$ and $b_i(x, \tau)$ are functions bounded in Ω, satisfying the conditions

$$|a_{ij}(x, \tau) - a_{ij}(x^\circ, \tau)| \leqslant A |x - x^\circ|^\alpha \tag{3.2}$$

$$|b_i(x, \tau) - b_i(x^\circ, \tau)| \leqslant A |x - x^\circ|^\alpha, \quad 0 < \alpha < 1, \qquad A = \text{const}$$

everywhere in Ω. We take it that the functions ω, ψ and Φ are continuous in Ω and that ω and ψ are bounded in growth as $|x| \to \infty$

$$|\omega(x)| \leqslant K_1 e^{h|x|}, \qquad |\psi(x)| \leqslant K_1 e^{h|x|} \tag{3.3}$$

(h is some positive constant), while the function $\Phi(x, \tau, v)$ satisfies a Lipschitz condition in variable $v = (v_1, \ldots, v_n)$ uniformly with respect to $(x, \tau) \in \Omega$ and $v \in E_n$

$$|\Phi(x, \tau, v) - \Phi(x, \tau, v^0)| \leqslant K_2 |v - v^0|, \quad \Phi(x, \tau, 0) = 0 \tag{3.4}$$

(in particular, the functions Φ in Examples 2 and 3 satisfy condition (3.4) if the coefficients of matrix c are bounded in Ω).

The following three results, stemming from the assumptions made, are well known [10].

1) The unique fundamental solution $G(x, \tau; \eta, \sigma)$ of the linear Eqs. (2.2) and (2.3) exists. It is defined for all $(x, \tau) \in \Omega$ and $(\eta, \sigma) \in \Omega$ $(\tau > \sigma)$, satisfies the homogeneous equation $LG = 0$ (with respect to variables (x, τ)) and possesses the property

$$\lim_{\sigma \to \tau - 0} \int_{E_n} G(x, \tau; \eta, \sigma) f(\eta)\, d\eta = f(x) \tag{3.5}$$

for any continuous function $f(x)$ having a majorant (here λ is from (3.1))

$$|f(x)| \leqslant \text{const} \cdot \exp(\bar{h} |x|^2), \quad \bar{h} < \frac{\lambda}{4T} \tag{3.6}$$

2) The solutions of inhomogeneous Eqs. (2.2) and (2.3) are expressed in terms of $G(x, \tau; \eta, \sigma)$

$$F_0(x, \tau) = \int_{E_n} G(x, \tau; \eta, 0) \psi(\eta)\, d\eta + \int_0^\tau d\sigma \int_{E_n} G(x, \tau; \eta, \sigma) \omega(\eta)\, d\eta \tag{3.7}$$

$$F_N(x, \tau) = \int_{E_n} G(x, \tau; \eta, 0) \psi(\eta)\, d\eta + \int_0^\tau d\sigma \int_{E_n} G(x, \tau; \eta, \sigma) \times \tag{3.8}$$

$$\left[\omega(\eta) - \Phi\left(\eta, \sigma, \frac{\partial F_{N-1}}{\partial \eta}\right)\right] d\eta$$

(formula (3.7) is unconditionally valid by virtue of (3.3); (3.8) is true only under the condition that an inequality of type (3.3) (or, at least, (3.6)) is fulfilled for the derivatives $\partial F_N / \partial x_i$, which, as shown below, does obtain). The solutions F_N ($N = 0,1, \ldots$) are twice continuously differentiable with respect to variable x, and the derivatives $\partial F_N / \partial x_i$ and $\partial^2 F_N / \partial x_i \partial x_j$ can be computed by differentiating the right-hand sides of (3.7) and (3.8) under the integral sign.

3) The following inequalities are valid (for any $\lambda < \lambda$ from (3.1)):

$$|G(x, \tau; \eta, \sigma)| \leqslant K_3 (\tau - \sigma)^{-n/2} \exp\left[-\frac{\lambda |x - \eta|^2}{4(\tau - \sigma)}\right] \tag{3.9}$$

$$\left|\frac{\partial G(x, \tau; \eta, \sigma)}{\partial x_i}\right| \leqslant K_3 (\tau - \sigma)^{-(n+1)/2} \exp\left[-\frac{\lambda |x - \eta|^2}{4(\tau - \sigma)}\right] \tag{3.10}$$

Results 1)-3) are valid for the linear equations of successive approximations (2.2) and (2.3). Returning now to the original nonlinear Bellman Eq. (2.1) and to the synthesis problem, we look at two stages of the solving of the general problem. First, using the majorant bounds (3.9) and (3.10) for the fundamental solution we prove the convergence as $N \to \infty$ of the successive approximations $F_N(x, \tau)$ determined from (2.2) and (2.3), to the solution of the Bellman Eq. (2.1) (the existence and the uniqueness of the solution of (2.1) can be proved simultaneously). After this we show that the suboptimal systems constructed in conformity with control laws (2.4) are equivalent, asymptotically as $N \to \infty$, to the optimal system.

1°. First of all we prove the uniform convergence of the sequence of functions $F_0(x, \tau)$, $F_1(x, \tau), \ldots$, determined by the recurrence formulas (3.7) and (3.8), as well as of their partial derivatives $\partial F_N(x, \tau) / \partial x_i$ ($N = 0,1,2, \ldots$). To do this we form the differences

$$Q_N = F_{N+1}(x, \tau) - F_N(x, \tau) = \int_0^\tau d\sigma \int_{E_n} G(x, \tau; \eta, \sigma) \times \left[\Phi\left(\eta, \sigma, \frac{\partial F_{N-1}}{\partial \eta}\right) - \Phi\left(\eta, \sigma, \frac{\partial F_N}{\partial \eta}\right)\right] d\eta \tag{3.11}$$

$$\frac{\partial Q_N}{\partial x_i} = \frac{\partial F_{N+1}}{\partial x_i} - \frac{\partial F_N}{\partial x_i} = \int_0^\tau d\sigma \int_{E_n} \frac{\partial G(x, \tau; \eta, \sigma)}{\partial x_i} \times \left[\Phi\left(\eta, \sigma, \frac{\partial F_{N-1}}{\partial \eta}\right) - \Phi\left(\eta, \sigma, \frac{\partial F_N}{\partial \eta}\right)\right] d\eta \tag{3.12}$$

(in the latter formulas we can take $N = 0, 1, 2, \ldots$, if we agree that $\partial F_{-1} / \partial x_i = 0$). Using (3.4), for (3.11) and (3.12) we have the inequalities

$$|Q_N(x, \tau)| \leqslant K_2 \int_0^\tau d\sigma \int_{E_n} |G(x, \tau; \eta, \sigma)| \left|\frac{\partial Q_{N-1}(\eta, \sigma)}{\partial \eta}\right| d\eta \tag{3.13}$$

$$\left|\frac{\partial Q_N(x,\tau)}{\partial x_i}\right| \leqslant K_2 \int_0^\tau d\sigma \int_{E_n} \left|\frac{\partial G(x,\tau;\eta,\sigma)}{\partial x_i}\right| \left|\frac{\partial Q_{N-1}(\eta,\sigma)}{\partial \eta}\right| d\eta \qquad (3.14)$$

Formulas (3.13), (3.14) and (3.9), (3.10) enable us to find bounds for the differences (3.11) and (3.12) recurrently. For this it is necessary only to estimate $|\partial F_0 / \partial x_i|$. Here a bound of type (3.3) is valid, i.e.

$$\left|\frac{\partial Q_0}{\partial x_i}\right| \leqslant K e^{h|x|} \qquad (3.15)$$

Indeed, since when $\lambda > 0$

$$t^{-n/2} \int_{E_n} \exp\left[-\frac{\lambda}{t}\eta^2 + h|\eta|\right] d\eta \leqslant K_4 \qquad (3.16)$$

for the derivative $\partial F_0 / \partial x_i$ we obtain, with due regard to (3.3), (3.7), (3.9) and (3.10),

$$\left|\frac{\partial F_0}{\partial x_i}\right| \leqslant K_1 K_2 \left\{ \int_{E_n} \tau^{-(n+1)/2} \exp\left[-\frac{\lambda}{\tau}|x-\eta|^2 + h|\eta|\right] d\eta + \right. \qquad (3.17)$$

$$\left. \int_0^\tau d\sigma \int_{E_n} (\tau-\sigma)^{-(n+1)/2} \exp\left[-\frac{\lambda|x-\eta|^2}{\tau-\sigma} + h|\eta|\right] d\eta \right\} \leqslant$$

$$K_1 K_2 K_4 (\tau^{-1/2} + 2\tau^{1/2}) e^{h|x|}$$

Acting analogously and keeping in mind the inequality

$$\left|\frac{\partial F_0}{\partial \eta}\right| \leqslant \sum_{i=1}^n \left|\frac{\partial F_0}{\partial \eta_i}\right|$$

and (3.4), from (3.12) and (3.17) we obtain

$$\left|\frac{\partial Q_0}{\partial x_i}\right| \leqslant n K_1 K_2 K_3^2 K_4 \int_0^\tau d\sigma \int_{E_n} (\tau-\sigma)^{-(n+1)/2} (\sigma^{-1/2} + 2\sigma^{1/2}) \times$$

$$\exp\left[-\frac{\lambda|x-\eta|^2}{\tau-\sigma} + h|\eta|\right] d\eta \leqslant n K_1 K_2 K_3^2 K_4^2 e^{h|x|} (1+\tau)\pi$$

whence, by vitue of the boundedness of τ, we arrive at (3.15).

Using (3.15) and applying formulas (3.13) and (3.14) repeatedly, we can obtain the following bounds for differences (3.11) and (3.12) for an arbitrary number $N \geqslant 1$ ($\Gamma(\cdot)$ is the gamma-function):

$$|Q_N(x,\tau)| \leqslant \frac{K}{\sqrt{\pi}} \frac{\overline{K}^N \tau^{(N+1)/2}}{\Gamma((N+1)/2+1)} e^{h|x|}, \quad \overline{K} = \sqrt{\pi}\, n K_2 K_3 K_4 \qquad (3.18)$$

$$\left|\frac{\partial Q_N(x,\tau)}{\partial x_i}\right| \leqslant K \frac{\overline{K}^N \tau^{N/2}}{\Gamma(N/2+1)} e^{h|x|} \qquad (3.19)$$

Formulas (3.18) and (3.19) are proved by induction on N. The estimates obtained show that the sequences of functions

$$F_N(x, \tau) = F_0(x, \tau) + Q_0(x, \tau) + \ldots + Q_{N-1}(x, \tau) \tag{3.20}$$

$$\partial F_N(x, \tau) / \partial x_i = \partial F_0(x, \tau) / \partial x_i + \partial Q_0(x, \tau) / \partial x_i + \ldots \quad \partial Q_{N-1}(x, \tau) / \partial x_i \tag{3.21}$$

converge to some limit functions

$$F(x, \tau) = \lim_{N\to\infty} F_N(x, \tau), \quad W_i(x, \tau) = \lim_{N\to\infty} \partial F_N(x, \tau) / \partial x_i$$

Uniform convergence in any bounded region in Ω obtains for the partial sums of the series in the right hand side of (3.20), while in (3.21) only the sums of the summands beginning with the second converge uniformly. The first summand, according to (3.17) is majorized by a function having a singularity at $\tau = 0$. However, it is easy to see that this singularity is integrable and, therefore, we can pass to the limit in (3.8) and in the formula obtained by differentiating (3.8) with respect to x_i. As a result we obtain

$$F(x, \tau) = \int_{E_n} G(x, \tau; \eta, 0)\psi(\eta)\, d\eta + \int_0^\tau d\sigma \int_{E_n} G(x, \tau; \eta, \sigma) \times [\omega(\eta) - \Phi(\eta, \sigma, W(\eta, \sigma))]\, d\eta$$

$$W_i(x, \tau) = \int_{E_n} \frac{\partial G}{\partial x_i}\psi(\eta)\, d\eta + \int_0^\tau d\sigma \int_{E_n} \frac{\partial G}{\partial x_i}[\omega(\eta) - \Phi(\eta, \sigma, W(\eta, \sigma))]\, d\eta$$

whence it follows that $W_i(x, \tau) = \partial F(x, \tau) / \partial x_i$ $(i = 1, 2, \ldots, n)$, and, consequently, the limit function $F(x, \tau)$ satisfies the equation

$$F(x, \tau) = \int_{E_n} G(x, \tau; \eta, 0)\psi(\eta)\, d\eta + \int_0^\tau d\sigma \int_{E_n} G(x, \tau; \eta, \sigma)\left[\omega(\eta) - \Phi\left(\eta, \sigma, \frac{\partial F(\eta, \sigma)}{\partial \eta}\right)\right] d\eta \tag{3.22}$$

which is equivalent to the original Bellman Eq. (2.1), as is easily verified by a differentiation with the use of (3.5).

Thus, we have proved the existence of a solution of Eq. (2.1). From the proof it follows that the solution $F(x, \tau)$ and its derivatives $\partial F / \partial x_i$ have the majorants

$$|F(x, \tau)| \leqslant K_5 e^{h|x|}, \quad \left|\frac{\partial F(x, \tau)}{\partial x_i}\right| \leqslant K_5 \tau^{-1/2} e^{h|x|} \tag{3.23}$$

everywhere in Ω. Using (3.23) we can prove the uniqueness of the solution of (2.1). Indeed, admitting the existence of two solutions F_1 and F_2 of Eq. (2.1) (or (3.23)), for the difference $V = F_1 - F_2$ we obtain the equation

$$V(x, \tau) = \int_0^\tau d\sigma \int_{E_n} G(x, \tau; \eta, \sigma)\left[\Phi\left(\eta, \sigma, \frac{\partial F_1}{\partial \eta}\right) - \Phi\left(\eta, \sigma, \frac{\partial F_2}{\partial \eta}\right)\right] d\eta$$

which with due regard to (3.4) allows us to write

$$|V(x,\tau)| \leqslant K_2 \int_0^\tau d\sigma \int_{E_n} G'(x,\tau;\eta,\sigma) \left| \frac{\partial V(\eta,\sigma)}{\partial \eta} \right| d\eta$$

whence by arguments analogous to those above for functions F_N it is easy to obtain the following bound for the difference $V = F_1 - F_2$:

$$|V(x,\tau)| \leqslant K_5 \bar{K}^N \pi^{-1/2} e^{h|x|} \frac{\tau^{N/2}}{\Gamma(N/2+1)}$$

valid for any N. Hence it follows that $V(x,\tau) \equiv 0$, i.e., $F_1(x,\tau) = F_2(x,\tau)$. Thus we have proved that the successive approximations $F_0(x,\tau)$, $F_1(x,\tau)$, . . ., determined by the recurrence formulas (2.2) and (2.3), converge, asymptotically as $N \to \infty$, to the solution of the Bellman Eq. (2.1), which exists and is unique.

2°. We turn to the synthesis problem. Above, for the synthesis of the control system we suggested the use of functions $u_N(x,\tau)$ defined by (2.4). By $H_N(x,\tau)$ we denote functional (1.2) computed on the trajectories of system (1.1), passing through point x at instant $t = T - \tau$ under control $u = u_N$ (it is implicit that the lower bound of the integral in the right-hand side of (1.2) equals t). Function $H_N(x,\tau)$ determines the performance of control $u_N(x,\tau)$. It satisfies the linear equation

$$LH_N = -\omega(x) - u_N^T c^T \partial H_N / \partial x, \quad H_N(x,0) = \psi(x) \tag{3.24}$$

With due regard to the inequality $-u_N^T c^T \, \partial F_N / \partial x = \Phi(x,\tau,\partial F_N/\partial x)$, from (2.3) and (3.24) it follows that the difference $\Delta_N(x,\tau) = F_N(x,\tau) - H_N(x,\tau)$ satisfies the equation

$$L_N\Delta_N \equiv L\Delta_N + u_N{}^T c^T \partial\Delta_N / \partial x = \Phi(x,\tau,\partial F_{N-1}/\partial x) - \Phi(x,\tau,\partial F_N/\partial x) \quad \Delta_N(x,0) = 0 \tag{3.25}$$

Since for large N the right hand side of (3.25) is small (see (3.4) and (3.19))

$$\left|\Phi\left(x,\tau,\frac{\partial F_N}{\partial x}\right) - \Phi\left(x,\tau,\frac{\partial F_{N-1}}{\partial x}\right)\right| \leqslant \varepsilon_N' e^{h|x|} \tag{3.26}$$

$$\varepsilon_N' = \frac{nKK_2K^{N-1}T^{(N-1)/2}}{\Gamma((N+1)/2)}$$

and since the initial conditions are zero, we can expect that the difference $\Delta_N(x,\tau)$, as a solution of Eq. (3.25), is of the same order, i.e.,

$$|\Delta_N(x,\tau)| \leqslant \varepsilon_N' K_6' e^{h|x|} \tag{3.27}$$

When (as in Example 3) the functions $u_N(x,\tau)$ are bounded and smooth, ensuring the Hölder – continuity of the coefficients of operator L_N, the operator L_N, just as both L and the inequality (3.27), is an elementary consequence of formulas (3.7), (3.9) and (3.26). If, however, the $u_N(x,\tau)$ are discontinuous functions (but without singularities, e.g., such as those in Example 2), then inequality (3.27) follows from Theorem 2 in [11].

From the convergence of series (3.20) it follows that $|F(x,\tau) - F_N(x,\tau)| \leqslant$

$\varepsilon_N'' K_6'' e^{h|x|}$ ($\varepsilon_N'' \to 0$ as $N \to \infty$), whence with due regard to the inequality $|F - H_N| \leqslant |F - F_N| + |F_N - H_N|$ we finally obtain

$$|F(x, \tau) - H_N(x, \tau)| \leqslant \varepsilon_N K_6 e^{h|x|} \tag{3.28}$$

$$(\varepsilon_N = \max(\varepsilon_N', \varepsilon_N''), K_6 = \max(K_6', K_6''))$$

Formula (3.28) proves the asymptotic (as $N \to \infty$) optimality of the sub-optimal systems constructed with the use of control laws $u_N(x, \tau)$ computed by the recurrence formulas (2.2) – (2.4).

N o t e 1. If the coefficients of operator L are not bounded in Ω, then, in general, bounds (3.9) and (3.10) are false. However, it can happen that in this case the problem admits of reduction to the case considered above by a certain change of variables. If, for instance, the coefficients $\bar{b}_i(x, t)$ in (1.1) depend linearly on x (i.e., vector $\bar{b}(x, t) = B(t)x$, where $B(t)$ is an $(n \times n)$-matrix depending only on t), then the change of variables $y = Z^{-1}(t, t_0)\, x$ ($Z(t, t_0)$ is the fundamental matrix of system $x^{\cdot} = B(t)x$) eliminates the unbounded coefficients in operator L (in the new variables y), which permits us to investigate such systems by the methods in Sect. 3.

N o t e 2. The condition of uniform Lipschitz – continuity (3.4), used in the calculations in Sect. 3, is not fulfilled for the function $\Phi(x, \tau, \partial F / \partial x)$ of Example 1 (where instead of (3.4) we have

$$|\Phi(x, \tau, v) - \Phi(x, \tau, v^{\circ})| \leqslant K\,|v - v^{\circ}|^2)$$

It can be shown that in this case the arguments made in Sect. 3 are not violated if the restriction (3.9) on the growth of functions ω and ψ is replaced by the requirement $|\omega\cdot(x)|, |\psi(x)| \leqslant K_0 + K_1 |x|^l$.

E x a m p l e 4. We illustrate the effectiveness of the proposed method by example of a one-dimensional control problem for which the exact solution is known. Let the control system be described by the scalar equation

$$x^{\cdot} = u + \xi(t), \ M\xi(t)\xi(t - \tau) = a\delta(\tau), \ u \in [\alpha, \beta]$$

($\delta(\tau)$ is the delta-functions and the point $u = 0$ is located inside the segment $[\alpha, \beta]$). In this case the Bellman Eq. (1.3) has the form ($\psi(x) = 0$)

$$\frac{\partial F}{\partial \tau} = \omega(x) + \min_{u \in [\alpha, \beta]} \left[u \frac{\partial F}{\partial x}\right] + \frac{a}{2} \frac{\partial^2 F}{\partial x^2}, \quad F(x, 0) = 0$$

It was solved in [12]. Here we restrict our consideration to the symmetric problem when $\omega(x)$ is an even function having a minimum at $x = 0$ and $u = 0$ is the midpoint of segment $[\alpha, \beta]$, i.e., $\alpha = -u_0$ and $\beta = u_0$. In this case the optimal control is $u^* = -u_0 \operatorname{sign} x$, and the solution $F(x, \tau)$ of the Bellman equation is an even function of x and for $x \geqslant 0$ is determined by the formula [12]

$$F(x, \tau) = \int_0^{\tau} \frac{d\sigma}{\sqrt{2\pi a \sigma}} \int_0^{\infty} \omega(\mu) \exp\left[\frac{u_0}{a}(x - \mu) - \frac{u_0^2}{2a}\sigma\right] \times$$

$$\left\{\exp\left[-\frac{1}{2a\sigma}(x - \mu)^2\right] + \exp\left[-\frac{1}{2a\sigma}(x + \mu)^2\right] + \right.$$

$$\frac{2u_0}{a}\int_0^\infty \exp\left[-\frac{1}{2a\sigma}(x+\mu+\nu)^2+\frac{u_0}{a}\nu\right]d\nu\Big\}\,d\mu$$

Turning to the calculation of the successive approximations, we note that all the functions $F_0, F_1, \ldots$ are even in the variable x. Therefore, any approximate control(2.4) coincides with the optimal u^* and the method's effectiveness can be evaluated from the deviation of the successive approximations $F_0, F_1, \ldots$ from the exact solution. Selecting a quadratic penalty function $\omega(x) = x^2$, for the first two approximations we obtain the expressions

$$F_0(x,\tau)=\int_0^\tau \frac{d\sigma}{\sqrt{2\pi a(\tau-\sigma)}}\int_{-\infty}^{+\infty}\mu^2\exp\left[-\frac{(x-\mu)^2}{2a(\tau-\sigma)}\right]d\mu = x^2\tau+\frac{a}{2}\tau^2$$

$$F_1(x,\tau)=F_0(x,\tau)-2u_0\int_0^\tau \frac{\sigma\,d\sigma}{\sqrt{2\pi a(\tau-\sigma)}}\int_{-\infty}^{+\infty}|\mu|\exp\left[-\frac{(x-\mu)^2}{2a(\tau-\sigma)}\right]d\mu$$

Figure 1 shows the functions F_0, F_1 and F computed for $u_0 = a = \tau = 1$. We see that

$$\max_x \frac{|F(x)-F_0(x)|}{F(x)} \approx 1,$$

$$\max_x \frac{|F(x)-F_1(x)|}{F(x)} \approx 0.2$$

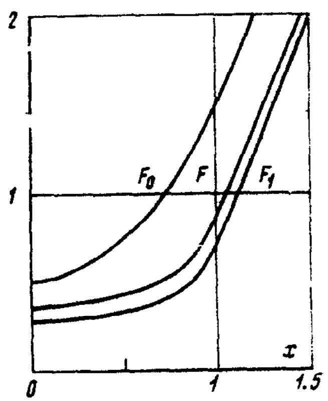

Fig. 1

i.e., even the second approximation yields a satisfactory approximation to the exact solution. The example considered shows that the actual convergence of the successive approximations to the solution of the Bellman equation can take place more rapidly than the theoretical one which is estimated by formulas (3.18) and (3.20). This face is a consequence of the grossness of the bounds (3.9) and (3.10) of the fundamental solution, on which we have based the proof of the convergence of the method of successive approximations in Sect. 3.

The author thanks V.B. Kolmanovskii for discussing the paper's results.

REFERENCES

1. Liptser, R. Sh. and Shiriaev, A. N., Statistics of Random Processes. Moscow, "Nauka", 1974.

2. Bellman, R., Dynamic Programming. Princeton, NJ, Princeton Univ. Press, 1957.

3. Krasovskii, N. N. and Lidskii, E. A., Analytic design of controllers in stochastic systems with velocity-limited controlling .action. PMM Vol. 25, No. 3, 1961.

4. Z u b o v, V. I., Lectures on Control Theory. Moscow, "Nauka", 1975.

5. B e l l m a n, R. and A n g e l, E., Dynamic Programming and Partial Differential Equations. New York, Academic Press, 1972.

6. F l e m i n g, W. H., Some Markovian optimization problems. J. Math. and Mech., Vol. 12, No. 1, 1963.

7. K o l o s o v, G.E. and S t r a t o n o v i c h, R. L., On an asymptotic method of solving the synthesis problem for optimal regulators. Avtomat. Telemekh., Vol., 25, No. 12, 1964.

8. C h e r n o u s ' k o, F. L. and K o l m a n o v s k i i, V. B., Computational and approximate methods of optimal control. In: Mathematical Analysis, Vol. 14. Moscow, Itogi Nauki i Tekhniki VINITI Akad. Nauk SSSR, 1977.

9. C h e r n o u s ' k o, F. L., Some problems of optimal control with a small parameter. PMM Vol. 32, No. 1, 1968.

10. I l ' i n, A. M., K a l a s h n i k o v, A. S. and O l e i n i k, O. L., Second-order parabolic linear equations. Uspekhi Mat. Nauk, Vol. 17, No. 3, 1962.

11. K a m y n i n, L. I., Method of heat potentials for a parabolic equation with discontinuous coefficients. Sib. Mat. Zh., Vol. 4, No. 5, 1963.

12. K o l o s o v, G. E., On a nonstationary optimal control problem. Izv. Akad. Nauk SSSR, Tekhn. Kibernet., No. 4, 1975.

Translated by N. H. C.