ON A CLASS OF LINEAR OPTIMUM PROBLEMS

(OB ODNOM KLASSE LINEINYKH OPTIMAL'NYKH ZADACH)

PMM Vol.28, № 1, 1964, pp.25-38

E.P. GRIGOR'EV
(Leningrad)

(Received April 1, 1963)

The problem of minimizing functionals of particular form for constraints given by linear differential equations with variable coefficients and linear boundary conditions of general form is investigated. The functionals are given in the form of integrals of functions of the length of the control vector. The controls enter linearly into the right-hand sides of the differential equations and are bounded in modulus. The question of uniqueness and existence of optimum controls is solved for similar problems and the connection of these problems with problems of fast response is also established; methods of finding the gradients of the functions to be minimized by using the solution of certain systems of differential equations are indicated.

Let us consider the following variational problem : To find the m-dimensional vector function $\mathbf{U}(t)$ satisfying the restriction

$$|\mathbf{U}(t)| = \Big(\sum_{i=1}^{m} u_i^2(t)\Big)^{1/2} \leqslant U_*(t) \qquad (0.1)$$

which will achieve a minimum of the integral

$$I(\mathbf{U}) = \int_0^T f(|\mathbf{U}|)\,dt \qquad (0.2)$$

and will transform a point of the G_0 plane

$$L_0\mathbf{X}_0 = \mathbf{Y}_0, \qquad \mathbf{X}_0(x) = \mathbf{X}_0 \qquad (0.3)$$

of n-dimensional space into a point of the plane

$$L_1\mathbf{X}_1 = 0, \qquad \mathbf{X}(T) = \mathbf{X}_1 \qquad (0.4)$$

along the trajectory of a system of differential Equations

$$\mathbf{X}^{\cdot} = A\mathbf{X} + B\mathbf{U} \qquad (\mathbf{X}^{\cdot} = d\mathbf{X}/dt) \qquad (0.5)$$

within the time T. Here $\mathbf{X}_1$, $\mathbf{X}_0$ and $\mathbf{Y}_0$ are n-dimensional vectors, $\mathbf{X}(t)$ is an n-dimensional vector function, L_0 and L_1 are constant nth order square matrices of ranks $l_0 \leqslant n$, and $l_1 \leqslant n$, respectively, $A(t)$ is an nth order square matrix of the functions $a_{ik}(t)$, $B(t)$ is a matrix of the functions $b_{ij}(t)$ and has n rows and m columns. We shall consider the function $f(\sigma)$ to be twice continuously differentiable, given for $\sigma \in [0, +\infty)$ and such that

$$f(0) = 0, \quad \frac{df}{d\sigma}(0) = 0, \quad \frac{d^2f}{d\sigma^2} > 0, \quad \lim_{\sigma \to +\infty} \frac{df}{d\sigma} = +\infty \tag{0.6}$$

We shall consider the functions $a_{ik}(t)$, $b_{ij}(t)$, and $U_*(t)$ given for $t \in [0, +\infty)$, where we shall henceforth limit ourselves to the case when these functions are bounded, analytic and, moreover

$$U_*(t) \geqslant c > 0, \qquad t \in [0, +\infty) \tag{0.7}$$

where c is a fixed number.

In order for Equations (0.3) which defines the initial plane G_0 to be consistent, condition

$$\operatorname{rank}(L_0, \mathbf{Y}_0) = \operatorname{rank} L_0 = l_0 \tag{0.8}$$

must be satisfied.

The basic purpose of this paper is to obtain a method of finding the vector function $\mathbf{U}(t)$ which solves the formulated variational problem. However, to do this it is first necessary to investigate the question of the uniqueness of the desired function $\mathbf{U}(t)$ and to prove its existence. We shall designate the formulated variational problem as problem 1. At the same time let us consider the problem of minimization of the integral (0.2) when $f(|U|) \equiv 1$, under the restriction (0.1) and the constraints (0.3) to (0.5), i.e. the fast-response problem, which we shall henceforth designate as problem 2.

1. Uniqueness questions. In order to derive the uniqueness conditions for the desired vector function $\mathbf{U}(t)$ (the optimum control), let us use the scheme proposed in [1]. According to the maximum principle, the optimum control may be found from the condition for the maximum of the function

$$H(\Lambda, \mathbf{X}, \mathbf{U}) = \Lambda^* A\mathbf{X} + \Lambda^* B\mathbf{U} - f(|\mathbf{U}|)$$

in $\mathbf{U}$ in the domain defined by the restriction (0.1), where $\Lambda(t)$ is a non-zero solution of the system of n differential equations

$$\dot{\Lambda} = -A^*\Lambda \tag{1.1}$$

which satisfies transversality conditions of the form

$$\Lambda(0) = \Lambda_0, \quad K_0\Lambda_0 = 0, \qquad K_l\Lambda(T) = 0 \tag{1.2}$$

where K_0 and K_1 are real, constant nth order square matrices of ranks $n - l_0$ and $n - l$, respectively, which satisfy conditions

$$L_0K_0^* = 0, \qquad L_1K_1^* = 0 \tag{1.3}$$

Here and henceforth the asterisk will denote the transposition operation. If the inverse function for $df/d\sigma = h(\sigma)$ is denoted by $\sigma(h)$, then the optimum control $\mathbf{U}(t)$ may be expressed from the condition for the maximum of the function H in terms of the solution of the system (1.1) by means of Formula

$$\mathbf{U} = \begin{cases} \frac{B^*\Lambda}{|B^*\Lambda|}\sigma(|B^*\Lambda|) & \text{for } \sigma(|B^*\Lambda|) < U_*(t) \\ \frac{B^*\Lambda}{|B^*\Lambda|}U_*(t) & \text{for } \sigma(|B^*\Lambda|) \geqslant U_*(t) \end{cases} \tag{1.4}$$

The analogous Formula for problem 2 is

$$\mathbf{U}(t) = \frac{B^*\Lambda}{|B^*\Lambda|}U_*(t) \tag{1.5}$$

Let us note that Expressions (1.4) and (1.5) uniquely determine a function U, continuous on the left, only when the vector $B^*\Lambda$ vanishes at only a finite number of points of the interval $[0, T]$. Otherwise, the analytic vector function $B^*\Lambda \equiv 0$ for $t \in [0, T]$. However, according to the maximum principle, the vector $\Lambda(t)$, corresponding to the optimum control, should not be zero. Hence, in order to eliminate the possible ambiguity of the optimum control, it is necessary to require that the relation $\Lambda \equiv 0$ should result from condition $B^*\Lambda \equiv 0$. If the vector function $B^*\Lambda$ is expanded into the series

$$B^*\Lambda = \sum_{k=0}^{\infty} G_k t^k \Lambda_0$$

then the obtained condition will be equivalent to the fact that the system of Equations

$$G_k\Lambda_0 = 0 \qquad (k = 0, 1, \ldots) \tag{1.6}$$

has only the zero solution $\Lambda_0 = 0$. When the martices A and B are constant, the matrices G_k are determined by Formulas

$$G_k = \frac{(-1)^k}{k!} B^* (A^*)^k$$

Hence, it is seen that the system (1.6) will be equivalent to a finite system of Equations

$$B^* (A^*)^k \Lambda_0 = 0 \qquad (k = 0, 1, \ldots, r-1) \tag{1.7}$$

where r is the degree of the minimum polynomial of the matrix A^*, $r \leqslant n$. The system (1.7) will have only the zero solution $\Lambda_0 = 0$ if and only if n linearly independent vectors are found among the $r \times m$ vector columns of the matrices

$$B, AB, \ldots, A^{r-1}B \tag{1.8}$$

Henceforth, the condition of unique definability of $\mathbf{U}(t)$ in terms of $\Lambda(t)$ will be called condition A.

The vector function $\mathbf{U}(t)$ satisfying condition (0.1) and transforming a certain point $\mathbf{X}_0 \in G_0$ within the time T along the trajectory of the system (0.5) into any point $\mathbf{X}_1 \in G_1$ (G_1 is a plane given by (0.4)), will be called the admissible control. The admissible control satisfying the maximum principle will be called the extremal control.

T h e o r e m 1.1. Both the optimum and the extremal controls are unique when condition A is satisfied.

P r o o f . Let us assume that the extremal control $\mathbf{U}(t)$ transforms the point $\mathbf{X}_0 \in G_0$ along the trajectory of the system (0.5) into the point $\mathbf{X}_1 \in G_1$ within the time T and Λ_0 is the initial value of the vector $\Lambda(t)$ corresponding to $\mathbf{U}(t)$ according to maximum principle. Multiplying (0.5) by Λ^* and (1.1) by $\mathbf{X}^*$ on the left and combining the results, we obtain

$$(\Lambda^*\mathbf{X})^\cdot = \Lambda^{*\cdot}B\mathbf{U}$$

Hence

$$\Lambda^* (T)\ \mathbf{X}\ (T) - \Lambda_0^*\mathbf{X}_0 = \int_0^T \Lambda^*B\mathbf{U}\, dt \qquad (1.9)$$

Since the vectors $\mathbf{X}(T)$ and $\Lambda(T)$ satisfy conditions (0.4) and (1.2), respectively, then

$$\Lambda^* (T)\ \mathbf{X}\ (T) = 0$$

Analogously, by virtue of conditions (0.3) and (1.2) on the vectors $\mathbf{X}_0$ and Λ_0, the value $\Lambda_0{}^*\mathbf{X}_0$ remains constant for any $\mathbf{X}_0 \in G_0$. Hence, if $\mathbf{U}_1$ and $\mathbf{U}_2$ are two extremal controls, where $\Lambda(t)$ is a vector corresponding to the control $\mathbf{U}_1$, then the equality

$$\int_0^T \Lambda^*B\mathbf{U}_1 dt = \int_0^T \Lambda^*BU_2 dt \qquad (1.10)$$

is valid.

Now let us prove that

$$\int_0^T f\,(|\,\mathbf{U}_1\,|)\, dt = \int_0^T f\,(|\,\mathbf{U}_2\,|)\, dt \qquad (1.11)$$

In fact, it may otherwise be considered that

$$\int_0^T f\,(|\,U_1\,|)\, dt > \int_0^T f\,(|\,\mathbf{U}_2\,|)\, dt$$

from which taking into account (1.10) and the notation

$$H_1\,(\Lambda,\ \mathbf{U}) = \Lambda^*B\mathbf{U} - f\,(|\,\mathbf{U}\,|)$$

the inequality

$$\int_0^T H_1\,(\Lambda,\ \mathbf{U}_1) dt < \int_0^T H_1\,(\Lambda,\ \mathbf{U}_2)\, d \qquad (1.12)$$

follows.

On the other hand, the function H_1 achieved a maximum simultaneously with H, therefore

$$\int_0^T H_1\,(\Lambda,\ \mathbf{U}_1)\, dt \geqslant \int_0^T H_1\,(\Lambda,\ \mathbf{U}_2)\, dt$$

which contradicts inequality (1.12). From equalities (1.10) and (1.11) there results

$$\int_0^T H_1\,(\Lambda,\ \mathbf{U}_1)\, dt = \int_0^T H_1\,(\Lambda,\ \mathbf{U}_2)\, dt$$

and since $H_1(\Lambda,\ \mathbf{U})$ achieves a maximum for $\mathbf{U} = \mathbf{U}_1$ in $t \in [0,\ T]$, then $H_1\ (\Lambda,\ \mathbf{U}_1) \equiv \equiv H\ (\Lambda,\ \mathbf{U}_2)$ if the continuity of the functions $\mathbf{U}_1$ and $\mathbf{U}_2$ on the left is taken into account. The latter identity means that the same vector function $\Lambda(t)$

corresponds to the extremal controls U_1 and U_2 according to the maximum principle and hence, $U_1 \equiv U_2$, by virtue of condition A. Since the optimum control is the extremal control, the uniqueness of the optimum control is thereby proved also. The theorem is proved.

An analogous theorem for problem 2 is valid also upon compliance with condition A.

2. Existence theorems. Let us first establish a theorem which will permit the reduction of the problem of the existence of the optimum control to a problem of the presence of admissible controls [1].

T h e o r e m 2.1. If an admissible control exists in problem 1 and condition

$$\operatorname{rank}(L_0^*, \Phi^*(T) L_1^*) = n \tag{2.1}$$

is satisfied, where $\Phi(T)$ is fundamental, normalized matrix of the system $\dot{X} = AX$, then an optimum control exists also.

P r o o f . Let us consider that a class of measurable, bounded vector functions, given in $t \in [0, T]$ is selected as the class of admissible controls. The set of all admissible controls is non-empty, hence, the set of values of the functional (0.2) corresponding to the set of all admissible controls has the exact lower bound I_0. By definition of the exact lower bound, there exists a sequence of admissible controls $U_k(t)$ such that

$$\lim_{k \to +\infty} I(U_k) = I_0 \tag{2.2}$$

Each control $U_k(t)$ transfers a certain point $X_k(0) \in G_0$ into the point $X_k(T) \in G_1$. By the Cauchy formula we have

$$X_k(T) = \Phi(T)\left[X_k(0) + \int_0^T \Psi^* B U_k \, dt\right] \tag{2.3}$$

from which by virtue of the boundary conditions in (0.3) and (0.4) we have

$$L_0 X_k(0) = Y_0, \qquad L_1 \Phi(T) X_k(0) = -L_1 \Phi(T) \int_0^T \Psi^* B U_k dt \equiv Z_k$$

Here Ψ^* denotes the matrix Φ^{-1}. From the obtained system of $2n$ equations in $X_k(0)$ it is possible to select n equations such that the determinant of the latter system would be different from zero (this is possible since condition (2.1) is satisfied). It is easy to see that the components of the vector Z_k are bounded in a set, therefore the components of the vectors $X_k(0)$ and $X_k(T)$ are also bounded in a set. Hence, there exist sub-sequences of the sequences $X_k(0)$ and $X_k(T)$ such that

$$\lim_{q \to +\infty} X_q(0) = X_0, \qquad \lim_{q \to +\infty} X_q(T) = X(T) \tag{2.4}$$

Moreover, the relation (2.3) remains valid for the vectors $X_q(T)$ and $X_q(0)$. The components of the vectors $U_q(t)$ will be elements of the Hilbert space $L_2[0,T]$ bounded both in absolute value and in norm, hence, a sequence of vectors whose appropriate components converge weakly to certain limits forming the vector function $U_0(t)$, may be separated out. According to Banach-Saks theorem [2], a sequence U_{q_j} may be separated out of the obtained weakly convergent sub-sequence such that the sequence

$$U_l' = \frac{1}{l}\sum_{j=1}^{l} U_{q_j}$$

possesses the property that the appropriate components of the vectors $U_l'(t)$ converge in the norm of the space L_2 to components of the vector $U_0(t)$.

Finally, it is possible to separate out a sub-sequence $\mathbf{U}_k'(t)$ from the sequence $\mathbf{U}_l'(t)$, which will converge almost everywhere to $\mathbf{U}_0(t)$ [2]. Let us note that relation (2.3) remains true for the vectors

$$\mathbf{X}_l'(T) = \frac{1}{l}\sum_{j=1}^{l}\mathbf{X}_{q_j}(T),\quad \mathbf{X}_l'(0) = \frac{1}{l}\sum_{j=1}^{l}\mathbf{X}_{q_j}(0)$$

and $\mathbf{U}_l'(t)$. Moreover, it is evident that $\mathbf{U}_l'(t)$ satisfy the restriction (0.1), $\mathbf{X}_l'(0)\in G_0$, $\mathbf{X}_l'(T)\in G_1$. Consequently, the controls $\mathbf{U}_k'(t)$ will be admissible and, passing to the limit in the equalities (2.2) and (2.3), we obtain

$$\mathbf{X}(X) = \Phi(T)\left(\mathbf{X}_0 + \int_0^T \Psi^* B\mathbf{U}_0\,dt\right),\qquad \int_0^T f(|\mathbf{U}_0|)\,dt = I_0 \tag{2.5}$$

Here we used the assertion that the Cesaro sequences $\mathbf{X}_l'(0)$, $\mathbf{X}_l'(T)$ constructed by means of the sequences $\mathbf{X}_{q_j}(0)$, $\mathbf{X}_{q_j}(T)$, converge to the same limits, as do the original sequences [3], the relations (2.4) and the convexity f. Moreover, passing to the limit in the inequality

$$|\mathbf{U}_k'(t)| \leqslant U_*(t)$$

and, in case of need, changing the values of $\mathbf{U}_0(t)$ in a set of measure zero, we obtain

$$|\mathbf{U}_0(t)| \leqslant \mathbf{U}_*(t) \tag{2.6}$$

If it is taken into account that by virtue of the closedness of the manifolds G_0, G_1, $\mathbf{X}_0\in G_0$, $\mathbf{X}(T)\in G_1$, the optimality of the control $\mathbf{U}_0(t)$ results from relations (2.5) and (2.6). The theorem is proved.

N o t e . An analogous theorem is also valid for problem 2 with the sole difference that condition (2.1) must be satisfied for any t.

In order to investigate the question of the existence of admissible controls, let us consider the classical Lagrange problem of minimizing the functional

$$I = \int_0^T \mathbf{U}^2\,dt \tag{2.7}$$

under the differential constraints (0.5) and boundary conditions of the form

$$\mathbf{X}(0) = \mathbf{X}_0,\qquad \mathbf{X}(T) = 0 \tag{2.8}$$

The Euler equations for this problem, which we shall henceforth call problem 3, have the form

$$\mathbf{X}^{\cdot} = A\mathbf{X} + B\mathbf{U},\qquad \mathbf{\Lambda}^{\cdot} = -A^*\mathbf{\Lambda},\qquad \mathbf{U} = B^*\mathbf{\Lambda} \tag{2.9}$$

If $\Psi(t)$ denotes the fundamental normalized matrix of the system

$$\dot{\mathbf{\Lambda}} = -A^*\mathbf{\Lambda},$$

then by the Cauchy formula taking condition (2.8) into account, we obtain

$$\mathbf{X}(T) = \Phi(T)\left(\mathbf{X}_0 + \int_0^T \Psi^* BB^*\Psi\mathbf{\Lambda}_0\,dt\right) = 0$$

Hence, we have a system of linear algebraic equations

$$\mathbf{X}_0 + V(T)\,\mathbf{\Lambda}_0 = 0 \qquad \left(V(T) = \int_0^T \Psi^* BB^*\Psi\,dt\right) \tag{2.10}$$

for the determination of Λ_0.

If condition A is satisfied, the matrix $V(T)$ is nonsingular.

In fact, a vector $\Lambda_0 \neq 0$ exists otherwise, such that $V(T)\Lambda_0 = 0$, hence there results the relation

$$\Lambda_0^* V(T)\Lambda_0 = \int_0^T (B^*\Psi\Lambda_0)^2\,dt = 0 \qquad \text{or} \qquad B^*\Psi\Lambda_0 \equiv 0 \quad \text{for } t \in [0, T]$$

which contradicts condition A.

Consequently, for any T and arbitrary X_0 the unique solution of the system (2.10)

$$\Lambda_0(T, X_0) = -V^{-1}(T)X_0 \tag{2.11}$$

exists, which determines the solution of problem 3 by means of Formula

$$U(t, T, X_0) = B^*(t)\Psi(t)\Lambda_0(T, X_0) \tag{2.12}$$

The value of the functional I for the found extremum control is expressed in the form

$$I(T, X_0) = \int_0^T [B^*\Psi\Lambda_0(T, X_0)]^2\,dt =$$

$$= \Lambda_0^*(T, X_0)V(T)\Lambda_0(T, X_0) = -\Lambda_0^*(T, X_0)X_0 \tag{2.13}$$

Let us show that $I(T, X_0)$ will be non-increasing function of T. Actually, under the assumption made $I(T, X_0)$ will be an analytic function of T whose derivative is

$$\frac{dI}{dT} = U^2(T, T, X_0) + 2\Lambda_0^*(T, X_0)V(T)\frac{d\Lambda_0}{dT}$$

On the other hand, differentiating the left-hand side of the identity (2.10), obtained after the substitution of the function $\Lambda_0(T, X_0)$ in place of Λ_0, we obtain

$$\frac{dV}{dT}\Lambda_0(T, X_0) + V(T)\frac{d\Lambda_0}{dT} \equiv 0$$

Hence, the identity

$$U^2(T, T, X_0) + \Lambda_0^*(T, X_0)V(T)\frac{d\Lambda_0}{dT} \equiv 0$$

results.

Taking the obtained identity into account, the derivative dI/dT can be rewritten as

$$dI/dT = -U^2(T, T, X_0) \tag{2.14}$$

Hence, the function

$$I(T, X_0) = X_0^* V^{-1}(T)X_0$$

which decreases monotonically, tends to the finite limit $I(X_0)$ as $T \to +\infty$. By virtue of the arbitrariness of X_0 and the symmetry of the matrix $V^{-1}(T)$ it is easy to see that the finite limit

$$\lim_{T\to+\infty} V^{-1}(T) = V_0 \qquad (2.15)$$

and the finite limit

$$I(\mathbf{X}_0) = \lim_{T\to+\infty} \mathbf{X}_0{}^* V^{-1}(T)\,\mathbf{X}_0 = \mathbf{X}_0{}^* V_0 \mathbf{X}_0 \qquad (2.16)$$

exist

Moreover, the relations

$$\lim_{T\to+\infty} \mathbf{\Lambda}_0(T,\mathbf{X}_0) = -V_0\mathbf{X}_0, \qquad \lim_{T\to+\infty} \mathbf{U}^2(T,\,T,\,\mathbf{X}_0) = 0 \qquad (2.17)$$

are valid.

Since the matrix $V^{-1}(T)$ is positive-definite and symmetric, all its characteristic numbers will be positive. Hence, there results from the convergence of the matrix $V^{-1}(T)$ to the matrix V_0 as $T \to +\infty$ that the characteristic numbers of the matrix $V^{-1}(T)$ tend to the characteristic numbers of the symmetric matrix V_0. The characteristic numbers of the martix V_0 will therefore be non-negative real numbers. That case is of interest when all the characteristic numbers of the matrix V_0 would be zero. This last property will hold if, and only if, the greatest characteristic number of the matrix $V^{-1}(T)$ approaches zero as $T \to +\infty$. Since the characteristic numbers of the inverse matrix $V^{-1}(T)$ are inverse to the characteristic numbers of the matrix $V(T)$, the derived condition is equivalent to the fact that the least characteristic number of the matrix $V(T)$ tends to $+\infty$ as $T \to +\infty$.

It is known [4] that the least characteristic number of a symmetric matrix N is found by means of Formula

$$\nu_1 = \min_{|\mathbf{\Lambda}_0|=1} \mathbf{\Lambda}_0{}^* N \mathbf{\Lambda}_0$$

Therefore, all the characteristic numbers of the matrix V_0 will equal zero if, and only if, condition

$$\lim_{T\to+\infty} \int_0^T (B^*\Psi\mathbf{\Lambda}_0)^2\,dt = +\infty \qquad (2.18)$$

is satisfied for any $\mathbf{\Lambda}_0 \neq 0$

Now, it is not difficult to resolve the question of the existence of admissible controls.

T h e o r e m 2.2. 1) For any T there exists a $\rho > 0$ such that for an arbitrary point $\mathbf{X}_0$ from the neighborhood of the origin of radius ρ, $|\mathbf{X}_0|<\rho$, the solution of problem 3 determines the admissible control in problems 1 and 2 by means of Formula (2.12) if condition A is satisfied.

2) If conditions A and (2.18) are satisfied, then for any point $\mathbf{X}_0$ there exists a $T_0 > 0$ such that for $T > T_0$ the solution of problem 3 determines

the admissible control in problems 1 and 2 by means of Formula (2.12).

P r o o f . The first statement of the theorem results from the relation

$$\lim \mathbf{U}^2 (t, T, \mathbf{X}_0) = 0 \quad \text{for} \quad |X_0| \to 0$$

and Formula (0.7), therefore, let us prove the second statement of the theorem. According to what has been proved the matrix $V_0 = 0$, consequently

$$\lim \Lambda_0 (T, \mathbf{X}_0) = 0 \quad \text{for} \quad T \to +\infty \tag{2.19}$$

The function

$$I (\mathbf{X}_0, T) = \int_0^T \mathbf{U}^2 (t, T, \mathbf{X}_0)\, dt$$

being a meromorphic function of T, tends to zero as $T \to +\infty$, therefore, the function

$$\varphi (T, \Delta, \mathbf{X}_0) = \int_T^{T+\Delta} \mathbf{U}^2 (t, T+\Delta, \mathbf{X}_0)\, dt$$

which also will be meromorphic function of T, tends to zero as $T \to +\infty$ uniformly in $\Delta \in [0, +\infty]$; hence, we have

$$\varphi (T, \Delta, \mathbf{X}_0) = O (T^{-1}) \tag{2.20}$$

for fixed $\mathbf{X}_0$. Differentiating (2.20), we obtain the relation

$$d\varphi / dT = O(T^{-2}) \tag{2.21}$$

which is also satisfied uniformly in Δ. Evaluating $d\varphi/dT$ and taking account of the relation (2.21), we find that the relation

$$\lim [\mathbf{U}^2(T, T+\Delta, \mathbf{X}_0) + U^2(T+\Delta, T+\Delta, \mathbf{X}_0)] = 0 \quad \text{for} \quad T \to +\infty$$

is satisfied uniformly in $\Delta \in [0 + \infty)$

It is seen from (2.17) that $\mathbf{U}^2 (T+\Delta, T+\Delta, \mathbf{X}_0) \to 0$ as $T \to +\infty$ uniformly in $\Delta \in (0, +\infty)$, therefore $\mathbf{U}^2 (T, T+\Delta, \mathbf{X}_0) \to 0$ as $T \to +\infty$, uniformly in

$$\Delta \in [0 + \infty).$$

Consequently, for any $\varepsilon > 0$ a $T_1 > 0$ is found such that for $T > T_1$ the inequality

$$\mathbf{U}^2(t, T, \mathbf{X}_0) < \varepsilon \quad \text{for } t \in [T_1, T]$$

is valid.

On the other hand, taking account of the continuity of the function $\mathbf{U}^2(t, T, \mathbf{X}_0)$ in t, there results from (2.19) that for any $\varepsilon > 0$ a $T_2 > 0$ is found such that $T > T_2$ the inequality

$$\mathbf{U}^2(t, T, \mathbf{X}_0) < \varepsilon \quad \text{for } t \in [0, T_1]$$

is satisfied.

Selecting the constant σ^2 from inequality (0.7) as ε, we obtain a $T_0 = \max(T_1, T_2)$ such that for $T > T_0$ the inequality

$$|\mathbf{U}(t, T, \mathbf{X}_0)| \leqslant \mathbf{U}^*(t) \quad \text{for } t \in [0, T]$$

is satisfied.

The obtained inequality shows that for $T > T_0$ Formula (2.12) determines the admissible control, q.e.d.

N o t e . When condition (2.18) is not satisfied for all $\Lambda_0 \neq 0$, it can similarly be shown that

$$\lim \mathbf{U}^2(t, T, \mathbf{X}_0) = [B^*(t)\ \psi(t)\ V_0 \mathbf{X}_0]^2 = \mathbf{U}^2(t, \mathbf{X}_0) \quad \text{for} \quad T \to +\infty$$

The obtained relation may be used to verify the restriction (0.1). If $U_*^2(t) > \mathbf{U}^2(t, X_0)$, there exists a T' such that for $T > T'$ there exists an admissible control in the considered problem for given $\mathbf{X}_0$.

Theorem 2.2 affords a possibility of solving the question of the existence of optimum controls. Henceforth, we shall consider condition A and condition (2.1) to be satisfied.

If condition (2.18) is satisfied for any $\Lambda_0 \neq 0$, then by virtue of Theorem 2.2 for the problem of fast response with the constraints (0.5), (0.3) and (0.4) and the restriction (0.1), there exist admissible controls and, therefore, there exists a unique optimum control. The fast-response time $T_0 = T_0(\mathbf{Y}_0)$ will here be a function $\mathbf{Y}_0$. The optimum fast-response control supplemented by the null vector $t \in [T_0, T]$, will be admissible in the problem of minimizing the functional (0.2) under the same conditions. Therefore, for $T \geqslant T_0(\mathbf{Y}_0)$ a unique optimum control in problem 1 exists.

If condition (2.18) is satisfied not for any $\Lambda_0 \neq 0$, let us consider problem 2 with the boundary conditions (2.8). The set of all $\mathbf{X}_0$ for which this problem is solvable will be called the domain of controllability (it was shown in [1] that the mentioned set will be convex). If the plane (0.3) has a non-empty intersection with the domain of controllability, then problem 2 has a unique optimum control, otherwise, problem 2 is not solvable. When problem 2 is solvable, problem 1 has a unique optimum control only for

$$T \geqslant T_0(\mathbf{Y}_0).$$

Let us note that by virtue of Theorem 2.2 the domain of controllability contains a set of points $\mathbf{X}_0$ satisfying the inequality

$$\mathbf{U}^2(t, \mathbf{X}_0) < \mathbf{U}_*^2(t) \qquad t \in [0 + \infty) \tag{2.22}$$

In concluding this section, let us consider the case when the matrices A and B are constant. Let us assume that condition (2.18) is satisfied not for all $\Lambda_0 \neq 0$, then there exists a vector $\Lambda_0' \neq 0$, such that the integral

$$\int_0^T (B^* \Psi \Lambda_0')^2 \, dt \qquad \text{for } T \to + \infty$$

tends to finite limit. Since this integral is an analytic function of T, its derivatives tend to zero as $T \to +\infty$, from which we have

$$\lim_{T \to +\infty} B^* (A^*)^k \Lambda_0' = 0 \qquad (k = 0, 1, \ldots) \tag{2.23}$$

Taking into account compliance with condition A, there follows from the relation (2.23)

$$\lim_{T \to +\infty} \Psi(T) \Lambda_0' = 0 \tag{2.24}$$

The equality (2.24) shows that the zero solution of the homogeneous system (1.1) is provisionally asymptotically stable. This latter is possible if, and only if, the matrix of this system has eigen numbers with negative real parts. This is equivalent to the matrix A having characteristic numbers with positive real parts. Hence, the following theorem is valid.

T h e o r e m 2.3. If the matrices A and B in the system (0.5) are constants, where the system of vector columns of the matrix (1.8) has n linearly independent vectors, all the characteristic numbers of the matrix A have non-positive real parts and

$$\operatorname{rank}(L_0^*, e^{A^* t} L_1^*) = n$$

for any t, then problem 2 has a unique solution. Problem 1 is uniquely solvable only for $T \geqslant T_0(\mathbf{Y}_0)$, where $T_0(\mathbf{Y}_0)$ is the fast-response time in the corresponding problem 2.

3. Method of solution. Let us first consider the method of solving problem 2. Since the optimum control has the form (1.5), to look for it is sufficient to find the corresponding initial value Λ_0 of the vector $\Lambda(t)$. It was shown in [5] that in the case $L_0 = L_1 = E$, where E is the unit matrix of order n, the solution of the prcblem of seeking the initial Λ_0 may be reduced to the solution of the problem of seeking the conditional extremum of a certain function, where the latter problem may be solved by the gradient method.

The following assertion may be proved.

Theorem 3.1. The value of the vector Λ_0 defining the optimum control in problem 2 by means of (1.1) achieves a relative minimum of the function

$$F_1(X_0, T) = \min F(\Lambda_0, X_0, T) = 0 \qquad \left(F_1 = \Lambda_0^* X_0 + \int_0^T u_*(t) \, |B^* \Psi \Lambda_0| \, dt\right) \tag{3.1}$$

under the conditions

$$|\Lambda_0| = 1, \quad K_0 \Lambda_0 = 0, \quad K_1 \Psi(T) \Lambda_0 = 0$$

This minimum will be a unique relative extremum of the function F and the values Λ_0' and T' achieving the minimum of F, which is zero, are independent of $X_0 \in G_0$.

Proof It follows from the condition of Theorem 2.1 of section 2 that rank $(K_0, K_1, \Psi(T)) \leqslant n$. Evidently, the theorem has meaning only when rank $(K_0, K_1 \Psi(T)) < n$, as we shall henceforth assume. Let us assume that Λ_0' is the value of the vector Λ_0 which defines the optimum fast-response control by means of (1.5) and which transforms the point $X_0 \in G_0$ into the point $X_1 \in G_1$ in the time T' along the trajectory of the system (0.5). Then by virtue of the boundary conditions

$$X^*(T') \Lambda(T') = 0, \qquad \Lambda_0'^* X_0 = \Lambda_0'^* X_0' \qquad \Lambda(T') = \Psi(T') \Lambda_0', \ X_0' \in G_0 \tag{3.2}$$

Hence

$$X_0'^* \Lambda_0' + \int_0^T u_*(t) \, | B^* \Psi \Lambda_0' | \, dt = 0 \tag{3.3}$$

where the value $F(\Lambda_0', X_0, T')$ is independent of $X_0 \in G_0$. If the assertion in the theorem is incorrect, then $F_1(X_0, T') < 0$. It is not difficult to see that $F_1(\mu X_0, T')$ will be continuous function of μ, where $\lim F_1(\mu X_0, T') > 0$ as $\mu \to +0$, since condition A is satisfied. Hence, there is found a μ_0, $0 < \mu_0 < 1$, such that

$$F_1(\mu_0 X_0, T') = 0 \tag{3.4}$$

Using the method of Lagrange multipliers, the necessary conditions for the minimum (3.4) is obtained in the form

$$\mu_0 X_0 + \int_0^{T'} \frac{\Psi^* B B^* \Psi \Lambda_0''}{|B^* \Psi \Lambda_0''|} u_* \, dt + K_0^* Y + \Psi^*(T') K_1^* Z + \lambda \Lambda_0'' = 0$$

$$K_0 \Lambda_0'' = 0, \qquad K_1 \Psi(T') \Lambda_0'' = 0, \qquad \Lambda_0''^2 = 1$$

$$Y^* = (y_1, \ldots, y_n), \qquad Z^* = (z_1, \ldots, z_n)$$

Here Y^*, Z^* and λ are Lagrange multipliers. Let us multiply the first of these relations by $\Lambda_0''^*$; then taking the remaining relations as well as (3.4) into account, we obtain $\lambda = 0$ and then

$$\mu_0 X_0 + \int_0^{T'} \frac{\Psi^* B B^* \Psi \Lambda_0''}{|B^* \Psi \Lambda_0''|} u_* \, dt + K_0^* Y + (K_1 \Psi(T'))^* Z = 0 \tag{3.5}$$

Let us transform (3.5) to the form

$$\Phi(T') \left[\mu_0 X_0 + K_0^* Y + \int_0^{T'} \Psi^* B U(t, \Lambda_0'') \, dt\right] + K_1^* Z = 0$$

where $\Phi(t)$ is fundamental, normalized matrix of the system $\dot{X} = AX$ and $U(t, \Lambda_0'')$ is defined by means of (1.5). Let us rewrite the last expression as

$$X(\mu_0, T') + K_1^* Z = 0$$

then it is not difficult to see that the control $U(t, \Lambda_0'')$ transforms the point $\mu_0 X_0 + K_0^* Y$, from the $L_0(\mu_0 X_0 + K_0^* Y) = \mu_0 Y_0$, plane into the point $X_0(\mu_0, T') \in G_1$ along the trajectory of the system (0.5) within the time T' and, hence, it will be an optimum fast-response control. On the other hand, the control $\mu_0 U(t, \Lambda_0')$ transforms the point $\mu_0 X_0$ from the $L_0 \mu_0 X_0 = \mu_0 Y_0$ plane into the point $\mu_0 X(T') \in G_1$ along the trajectory of the system (0.3) within the time T'; this contradicts the uniqueness of the optimum control $U(t, \Lambda_0')$. The theorem is proved.

Theorem 3.1 says nothing about the set of values Λ_0 which achieve the relative minimum (3.1). It is easy to prove that this set will be convex and closed. In fact, in view of the convexity of the function F the set of Λ_0 such that $F(\Lambda_0, X_0, T) \leqslant 0$, will be convex and closed. If zero is here the least value of the function F then $F(\Lambda_0, X_0, T) \geqslant 0$ for any Λ_0, from which it follows that the set of those Λ_0 for which $F(\Lambda_0, X_0, T) = 0$ agrees with the set Λ_0 for which $F(\Lambda_0, X_0, T) \leqslant 0$. q.e.d.

It was mentioned above that the domain of controllability defined by Equation (0.5), restriction (0.1) and boundary conditions (0.4) plays a large part in the clarification of the solvability conditions. Theorem 3.1 shows that the domain of controllability coincides with the set of all X_0 ($L_0 = E$) for which the function F (for $K = 0$) achieves the least non-positive value on the intersection of the sphere $|\Lambda_0| = 1$ with the set $k_1 \Psi(T) \Lambda_0 = 0$.

Let us now consider problem 1 by assuming that $T > T_0(Y_0)$, where $T_0(Y_0)$ is the fast-response time in the corresponding problem 2. The case $T = T_0(Y_0)$ is of no interest because the control which is optimum with respect to fast response, will then be the single admissible control. The optimum control in problem 1 is defined bv Formula

$$U(t) = \begin{cases} \dfrac{B^* \Psi \Lambda_0}{|B^* \Psi \Lambda_0|} \sigma(|B^* \Psi \Lambda_0|) & \text{for } t \in S \\ \dfrac{B^* \Psi \Lambda_0}{|B^* \Psi \Lambda_0|} u_*(t) & \text{for } t \in S' \end{cases} \tag{3.6}$$

where S and S' denote the sets of all $t \in [0, T]$, for which the inequalities

$$\sigma(|B^* \Psi \Lambda_0|) < u_*(t) \quad \text{or} \quad \sigma(|B^* \Psi \Lambda_0|) \geqslant u_*(t)$$

are satisfied, respectively.

If the control (3.6) transfers the point $X_0 \in G_0$ into the point $X(T) \in G_1$ along the trajectory of the system (0.5), then as has been shown in the proof of Theorem 2.1 in Section 2, the following equality

$$W_1 = \Lambda_0^* X_0 + \int_S |B^*\Psi\Lambda_0| \sigma(|B^*\Psi\Lambda_0|)\, dt + \int_{S'} |B^*\Psi\Lambda_0| u_* \, dt = 0 \quad (3.7)$$

is valid.

It is now easy to establish a theorem analogous to Theorem 1.

Theorem 3.2. The value of the initial vector Λ_0 defining the optimum control in problem 1 by means of (3.6) achieves the maximum of the function

$$W_2 = \int_S f(\sigma(|B^*\Psi\Lambda_0|))\, dt$$

under conditions $W_1 = 0$, $K_0\Lambda_0 = 0$, $K_1\Psi(T)\Lambda_0 = 0$. The mentioned value of Λ_0 is determined in a unique manner and is independent of $X_0 \in G_0$.

Proof. Let us show that for a fixed $X_0 \in G_0$ the set of Λ_0 satisfying the relation (3.7) and the transversality conditions will be a bounded set of the $X_0^*\Lambda_0 < 0$ half-space. In fact for any Λ_0 such that $X_0^*\Lambda_0 < 0$ the measure of the set $S(\Lambda_0)$ will be not less than a certain positive quantity since otherwise the relation (3.1) is violated. Hence, by virtue of condition (0.6) and condition A as $\mu \to \infty$ the inequality $\lim \mu^{-1} W_1(X_0, \mu\Lambda_0) > 0$ is correct uniformly in $|\Lambda_0| = 1$, satisfying the transversality conditions. It follows from above that a $\mu_0 > 0$ is found such that $W_1(X_0, \mu\Lambda_0) > 0$ for $\mu > \mu_0$ for all Λ_0 subjected to the mentioned restrictions, which proves the required boundedness of the set of Λ_0 defined above. Therefore, the function W_2 achieves its greatest value under the restrictions mentioned in the conditions of the theorem. Let us assume that this maximum is achieved for Λ_0'. Using the method of Lagrange multipliers, the necessary conditions for the relative extremum of function W_2 may be written in the form

$$\mu\left[X_0 + \int_S \sigma(|B^*\Psi\Lambda_0'|) \frac{\Psi^*BB^*\Psi\Lambda_0'}{|B^*\Psi\Lambda_0'|}\, dt + \int_{S'} \frac{\Psi^*BB^*\Psi\Lambda_0'}{|B^*\Psi\Lambda_0'|} u_* \, dt + \right.$$

$$\left. + (\mu + 1)\int_S \sigma'(|B^*\Psi\Lambda_0'|)\, \Psi^*BB^*\Psi\Lambda_0' \, dt + K_0^* Y + \Psi^*(T) K_1^* Z = 0 \right. \quad (3.8)$$

$$K_0\Lambda_0' = 0, \qquad K_1\Psi(T)\Lambda_0' = 0, \qquad W_1 = 0$$

Multiplying the first of equalities (3.8) on the left-hand side by $\Lambda_0'^*$, we obtain

$$(\mu + 1)\int_S \sigma'(|B^*\Psi\Lambda_0'|)(|B^*\Psi\Lambda_0'|)^2\, dt = 0$$

Hence, there results $\mu + 1 = 0$. Consequently, the equality

$$\Phi(T)\left[X_0 - K_0^* Y_0 + \int_0^T \Psi^* B U(t, \Lambda_0')\, dt\right] - K_1^* Z = 0$$

where $U(t, \Lambda_0')$ is defined by Formula (3.6), is applicable. The obtained equality shows that the control $U(t, \Lambda_0')$ will be an extremum control in problem 1. Hence, if W_2 achieves a relative extremum for $\Lambda_0 \neq 0$ under conditions (1.2) and (3.7), then Λ_0 determines the extremal control in problem 1 by means of Formula (3.6). Since the extremal control is unique, Formula (3.6) defines the identical control $U(t, \Lambda_0)$ for any Λ_0 achieving the provisional extremum of the function W_2. From the agreement of the controls

$\mathbf{U}(t, \Lambda_0)$ for different Λ_0 in a set S of non-zero measure follows the agreement of the vectors $B^*\Psi\Lambda_0$ in S for different Λ_0, which is impossible by virtue of condition A. Hence, the uniqueness of the Λ_0 which achieves the conditional extremum of the function W_2 is establisned and this proves the required assertion.

Theorems 3.1 and 3.2 show that to find the optimum controls in problems 1 and 2 it is sufficient to find the relative extremum of certain functions. In practice, however, it is considerably more convenient to solve the problem of seeking the absolute extremum of certain functions. In the cases under consideration, finding the relative extremums can be successfully reduced to problems of seeking the absolute extremums. In fact, in order to find the optimum control in problem 1,it is sufficient to find the absolute maximum of the function

$$W_2 - |W_1| - {}^1/_2 (K_0\Lambda_0)^2 - {}^1/_2 (K_1\Psi(T)\Lambda_0)^2 \tag{3.9}$$

Analogously, in order to find the optimum control in problem 2, it is sufficient to find the absolute minimum of the function

$$|F(\Lambda_0, \mathbf{X}_0, T)| + {}^1/_2 (K_0\Lambda_0)^2 + {}^1/_2 (K_1\Psi(T)\Lambda_0)^2 \tag{3.10}$$

which equals zero. The mentioned passage from the relative to the absolute extremums is possible because of the uniqueness of the relative extremums. In finding the extremums of (3.9) and (3.10) the gradient method may be used as its convergence raises no doubts since the desired extremum is unique. In the first case the $\Lambda_0 = \lambda \mathbf{X}_0 / |\mathbf{X}_0|$, where $\lambda < 0$ such that $W_1 \geqslant 0$, might be chosen as the initial approximation and the $\Lambda_0 = -\mathbf{X}_0 / |\mathbf{X}_0|$ in the second case. In both cases the point $\mathbf{X}_0$ may be chosen on the G_0 plane so that the quantity $|\mathbf{X}_0|$ would be a minimum. There remains to describe the method of finding the gradient of the functions (3.9) and (3.10). In case $W_1 \geqslant 0$, the gradient of the function (1.10) has the form

$$\Delta\Lambda_0 \equiv \mathbf{X}_0 + \int_0^T \Psi^* B\mathbf{U}(t, \Lambda_0)\, dt + K_0^* K_0\Lambda_0 + \Psi^*(T) K_1^* K_1 \Psi(T) \Lambda_0 \tag{3.11}$$

The gradient of the function (1.11) has the form

$$\Delta\Lambda_0 \equiv \operatorname{sign} F\Big(\mathbf{X}_0 + \int_0^T \Psi^* B\mathbf{U}(t, \Lambda_0)\, dt\Big) + \\ + K_0^* K_0\Lambda_0 + \Psi^*(T) K_1^* K_1 \Psi(T) \Lambda_0 \tag{3.12}$$

The function $\mathbf{U}(t, \Lambda_0)$ in (3.11) is defined by Expression (3.6) and the function $\mathbf{U}(t, \Lambda_0)$ in (3.12) has the form of (1.5). Let us dwell in more detail on the calculation of the gradient of the function (3.9) by means of Formula (3.11). The solution of the system of equations (0.5), where $\mathbf{U}$ is defined by (3.6), with the initial data $\mathbf{X}(0) = \mathbf{X}_0$ may be written

$$\mathbf{X}(T) = \Phi(T)\Big(\mathbf{X}_0 + \int^T \Psi^* B\mathbf{U}(t, \Lambda_0)\, dt\Big) \tag{3.13}$$

Let us consider the solution of the system of equations

$$\mathbf{Z}^{\cdot} = -A\mathbf{Z} + B(T-t)\,\mathbf{U}(T-t) \tag{3.14}$$

with the initial conditions $\mathbf{Z}(0) = \mathbf{Z}_0$. Then for $t = T$ we obtain

$$\mathbf{Z}(T) = \Phi(-T)\left[Z_0 + \int_0^T \Phi(t)\,B(T-t)\,\mathbf{U}(T-t)\,dt\right] \tag{3.15}$$

If the change of variable $v = T - t$ is made in the intergral of the obtained expression, then (3.15) becomes

$$\dot{\mathbf{Z}}(T) = \Psi^*(T)\,\mathbf{Z}_0 + \int_0^T \Psi^*(t)\,B(t)\,\mathbf{U}(t)\,dt \tag{3.16}$$

Assuming $\mathbf{Z}_0 = K_1{}^*K_1\Psi(T)\,\Lambda_0$ in Formula (3.16), we then obtain Expression

$$\Delta\Lambda_0 = \mathbf{Z}(T) + \mathbf{X}_0 + K_0{}^*K_0\Lambda_0 \tag{3.17}$$

for the gradient (3.10).

In deducing (3.17), it was taken into account that $U(T-t)$ is obtained by means of Formula (3.6) by replacing t by $T-t$. It is easy to see that the vector $\Psi(T-t)\,\Lambda_0 = \Psi(-t)\,\Psi(T)\,\Lambda_0$ is obtained in integrating the system of Equations

$$\Omega^{\cdot} = A^*\Omega \tag{3.18}$$

with the initial data $\Omega(0) = \Psi(T)\,\Lambda_0$. Hence, the desired gradient $\Delta\Lambda_0$ may be obtained by means of Formula (3.17), where $\mathbf{Z}(T)$ is found by integrating the system of Equations

$$\dot{\mathbf{Z}} = -A\mathbf{Z} + B(T-t)\,\mathbf{U}(T-t,\Omega), \qquad \dot{\Omega} = A^*\Omega \tag{3.19}$$

between 0 and T with the initial conditions $\mathbf{Z}(0) = K_1{}^*K_1\Psi(T)\,\Lambda_0$, and $\dot{\Omega}(0) = \Psi(T)\,\Lambda_0$. Here $\mathbf{U}(T-t,\Omega)$ has the form

$$\mathbf{U}(T-t,\Omega) = \begin{cases} \dfrac{B^*(T-t)\,\Omega}{|B^*(T-t)\,\Omega|}\,\sigma(|B^*(T-t)\,\Omega|) & \text{for } T-t \in S \\ \dfrac{B^*(T-t)\,\Omega}{|B^*(T-t)\,\Omega|}\,u^*(T-t) & \text{for } T-t \in S' \end{cases}$$

Hence, the vector $\Psi(T)\Lambda_0$ is found by integrating the system $\Lambda^{\cdot} = -A^*\Lambda$ between 0 and T with the initial conditions $\Lambda(0) = \Lambda_0$. The gradient (3.12) is calculated analogously with the sole difference that $\mathbf{U}(t,\Lambda_0)$ is found by means of Formula (1.5).

BIBLIOGRAPHY

1. Pontriagin, L.S., Boltianskii, V.G., Gamkrelidze, R.V. and Mishchenko, E.F., Matematicheskaia teoriia optimal'nykh protsessov (Mathematical Theory of Optimum Processes). Fizmatgiz, 1961.

2. Riesz, F. and Sekefalvy-Nagy, B., Lektsii po funktsional'nomu analizu (Lectures on the functional analysis). IL, p. 91, 1954.

3. Polya, G. and Szego, G., Zadachi i teoremy iz analiza (Problems and theorems from the analysis). Gostekhteoretizdat, Vol.1, p.33, 1956.

4. Gantmakher, F.R., Teoriia matrits (The Theory of Matrix). Gostekhteoretizdat, 1954.

5. Krasovskii, N.N., Vybor parametrov optimal'nykh ustoichivykh sistem (Selection of parameters of the optimum stable systems). Proc. of the First Internat. Congr. on automatic control. Izd-vo Akad.Nauk SSSR, 1961.

Translated by M.D.F.