6. Lur'e, A.I., Analytical Mechanics. M., Fizmatgiz, 1961.
7. Landau, L.D. and Lifshits, E.M., Electrodynamics of Continuous Media. M., Fizmatgiz, 1959.
8. Rozenvasser, E.N., Application of integral equations in the theory of nonlinear oscillations. Proceedings of the 2nd All-Union Congress on the Theoretical and Applied Mechanics., 1964, Vol.2, M. Nauka, 1965.

Translated by L.K.

ON OPTIMIZATION OF THE TRACKING PROCESS

PMM Vol. 33, No. 1, 1969, pp. 101-111
F.L. CHERNOUS'KO
(Moscow)
(Received September 3, 1968)

The present paper concerns the optimization of the tracking process (with allowance for measurement errors) in a system whose motion is described by linear differential equations. It is shown that under certain assumptions the problem reduces to one of ordinary optimal control. Further analysis using the maximum principle enables us to reduce the initial problem to a system of transcendental equations. Examples illustrating optimal tracking strategy in specific cases are discussed.

Problems of optimal control in the absence of complete information, i.e. with incomplete and inexact measurements or observations, are of great interest in control engineering. Various approaches to optimal control and tracking problems with incomplete information are considered, for example, in [1-8], whose authors employ both probabilistic and minimax formulations.

1. The initial relations. Let the state of a system at any instant be defined by an n-dimensional phase coordinate vector x. The law of variation of $x(t)$ takes the form of a determinate linear system of ordinary differential equations,

$$dx / dt = A(t)x + b(t) \qquad (1.1)$$

where A is an $n \times n$ matrix and b is an dimensional vector. System (1.1) can be regarded in many cases as a system in variations near the theoretical (nominal) trajectory of the initial nonlinear system.

The motion of the system is considered over the time interval $[t_0, T]$; the phase coordinates of the system are observed (measured) at the fixed instants $t_0, t_1, \ldots, t_N = T$. Here $t_k < t_{k+1}$ for $k = 0, 1, \ldots, N-1$. By "observation" at each instant of time t_k we mean the approximate measurement of certain linear combinations of the components of the vector $x(t_k)$, i.e. measurement of the vector $Q_k x(t_k)$. Here Q_k is a given rectangular matrix with l_k rows and n columns. The integer $l_k \geqslant 0$ is the number of scalar parameters measured at the instant t_k, $k = 0, 1, \ldots, N$. We assume that the error of each observation is a random l_k-dimensional vector quantity distributed according to a normal law with zero mathematical expectation and a known correlation matrix B_k. The term "correlation matrix" is used throughout the present paper to refer to an unnormalized correlation matrix (a second-moment matrix). The measurement error at a given instant is assumed to be independent of the errors at the other instants.

Thus, the result of observation at the instant t_k is a random l_k-dimensional vector quantity y_k with a normal distribution law. Its mathematical expectation is equal to the true value of $Q_k x(t_k)$, and its $l_k \times l_k$ correlation matrix is known and equal to B_k.

The values of the phase vectors at the instants of observation are related by the linear expressions

$$x(t_{k+1}) = A_k x(t_k) + b_k \qquad (k = 0, 1, \ldots, N-1) \tag{1.2}$$

Equations (1.2) follow from the linearity of system (1.1). The matrix A_k and the vector b_k depend on the matrix A and the vector b in the interval $[t_k, t_{k+1}]$. Explicit expressions for A_k and b_k can be readily written by the method of variation of arbitrary constants provided we know the matrix of fundamental solutions for the homogeneous equation corresponding to (1.1). Relations (1.2) are also valid if the system motion is described by finite-difference equations or if equations of motion (1.1) are impulsive in character (e.g. if $b(t)$ is a delta function).

Let the probabilistic distribution of the initial value of the phase vector $x(t_0 - 0)$ just before the start of the process be known. We consider this distribution to be normal with the mathematical expectation x_0 and the correlation matrix D_0. The purpose of tracking is to enable us to indicate the mathematical expectation and correlation matrix for an unknown instantaneous value of the phase coordinate vector at any instant.

These quantities (the mathematical expectation and correlation matrix) vary, first, by virtue of the equations of motion, and second, as a result of the measurements. All of the probabilistic distributions are assumed to be normal, and the results are treated by the method of maximum plausibility [4]. The analytical scheme which we describe and conversion formulas (1.3) are given in [3].

Let us denote by x_k and x_k^* the mathematical expectations for the unknown vector $x(t_k)$ at the instants $t_k - 0$ and $t_k + 0$, respectively, i.e. just before and just after the k-th measurement $(k = 0,1,\ldots,N)$. By D_k and D_k^* we denote the correlation matrices for the vector x at the instants $t_k - 0$ and $t_k + 0$.

It can be shown that the following conversion formulas are valid (the primes denote transposed matrices);

$$x_k^* = x_k + D_k^* Q_k' B_k^{-1}(y_k - Q_k x_k), \quad D_k^* = (D_k^{-1} + Q_k' B_k^{-1} Q_k)^{-1} \tag{1.3}$$

To derive (1.3) we note that the quantity x_k can be regarded as the result of measuring the phase vector $x(t_k)$ with the correlation matrix D_k, and y_k as the result of measuring the vector $Q_k x(t_k)$ with the correlation matrix B_k. Let us construct the plausibility function for these two measurements by the method of maximum plausibility [4],

$$\begin{aligned} L = C \exp\left[-\tfrac{1}{2}\left(D_k^{-1}(x_k^* - x_k), (x_k^* - x_k)\right)\right] \times \\ \times \exp\left[-\tfrac{1}{2}\left(B_k^{-1}(Q_k x_k^* - y_k), (Q_k x_k^* - y_k)\right)\right] \end{aligned} \tag{1.4}$$

Here the constant C does not depend on x_k^*. The comma denotes scalar multiplication of vectors. The required mathematical expectation can be determined from the maximum condition for function (1.4) with respect to x_k^*. Setting the gradient of the function L with respect to x_k^* equal to zero, we obtain

$$D_k^{-1}(x_k^* - x_k) + Q_k' B_k^{-1}(Q_k x_k^* - y_k) = 0$$

This yields the relation

$$x_k^* = F_k D_k^{-1} x_k + F_k Q_k' B_k^{-1} y_k, \qquad F_k = (D_k^{-1} + Q_k' B_k^{-1} Q_k)^{-1} \tag{1.5}$$

The quantities x_k^*, x_k, y_k can be regarded as normally distributed random quantities with the correlation matrices D_k^*, D_k, B_k, respectively; the quantities x_k, y_k are independent. Linear relation (1.5) then implies the following relationship [4] for the second-moment matrices:

$$\begin{aligned} D_k^* &= (F_k D_k^{-1}) D_k (F_k D_k^{-1})' + (F_k Q_k' B_k^{-1}) B_k (F_k Q_k' B_k^{-1})' = \\ &= F_k D_k^{-1} F_k + F_k Q_k' B_k^{-1} Q_k F_k = F_k (D_k^{-1} + Q_k' B_k^{-1} Q_k) F_k = \\ &= F_k F_k^{-1} F_k = F_k = (D_k^{-1} + Q_k B_k Q_k)^{-1} \end{aligned} \qquad (1.6)$$

In writing out transformations (1.6) we used Formula (1.5) for F_k, as well as the symmetry property of the matrices D_k, B_k, F_k and their inverses. By some elementary transformations we find from (1.5) and (1.6) that

$$\begin{aligned} x_k^* &= x_k + F_k Q_k' B_k^{-1} y_k - x_k + F_k D_k^{-1} x_k = x_k + F_k Q_k' B_k^{-1} y_k - \\ &- F_k (F_k^{-1} - D_k^{-1}) x_k = x_k + F_k Q_k' B_k^{-1} y_k - F_k Q_k' B_k^{-1} Q_k x_k = \\ &= x_k + F_k Q_k' B_k^{-1} (y_k - Q_k x_k) = x_k + D_k^* Q_k' B_k^{-1} (y_k - Q_k x_k) \end{aligned}$$

Since no measurements are made in the invervals between the instants t_k and t_{k+1}, we can write the following expressions on the basis of linear relations (1.2)

$$x_{k+1} = A_k x_k^* + b_k, \qquad D_{k+1} = A_k D_k^* A_k' \qquad (1.7)$$

Relations (1.3) and (1.7) are our point of departure; they describe the variation of the mathematical expectation and correlation matrix as a result of tracking process (1.3) and motion process (1.7). In order to carry out computations using recursion formulas (1.3) and (1.7) we need to choose initial values x_0, D_0 and measurement data y_k. The variation of the correlation matrix does not depend on the measurement data y_k and can be computed in advance.

2. Taking the limit. Let us set

$$\tau = (T - t_0) / N, \quad t_k = t_0 + k\tau \quad (k = 0, 1, \ldots, N)$$

and introduce matrices $B(t)$ and $Q(t)$ such that

$$B_k = B(t_k)\tau^{-1}, \quad Q_k = Q(t_k) \quad (k = 0, 1, \ldots, N) \qquad (2.1)$$

We now take the limit, making $\tau \to 0$, $N \to \infty$ with $N\tau = T - t_0 = \text{const}$. This passage to the limit corresponds to the case of very frequent observations (continuous observation in the limiting case). The error of each observation is then large (the elements of the matrix B_k are proportional to τ^{-1}), but the accuracy over a finite interval is finite.

From Eq. (1.1) we find that to within higher-order small terms

$$x(t_{k+1}) = x(t_k) + \tau[A(t_k) x(t_k) + b(t_k)]$$

Comparing this relation with Eq. (1.2), we find that

$$A_k = E + \tau A(t_k), \qquad b_k = \tau b(t_k) \qquad (2.2)$$

where E is an $n \times n$ unit matrix. Substituting Eqs. (2.1), (2.2) into Eqs. (1.3), we find that to within higher-order small terms

$$x_k^* = x_k + \tau D_k \; Q'(t_k)B^{-1}(t_k) \; [y_k - Q(t_k)x_k]$$

$$D_k^* = \{D_k^{-1}[E + \tau D_k Q'(t_k)B^{-1}(t_k)Q(t_k)]\}^{-1} = [E - \tau D_k Q'(t_k) \times \times B^{-1}(t_k)Q(t_k)]D_k = D_k - \tau D_k Q'(t_k)\, B^{-1}(t_k)Q(t_k)D_k$$

Substituting these relations and Eqs. (2.2) into Eqs. (1.7) and again omitting small terms of order τ^2 and higher, we obtain

$$x_{k+1} = x_k + \tau\{A(t_k)x_k + b(t_k) + D_k Q'(t_k)B^{-1}(t_k)[y_k - Q(t_k)x_k]\} \quad (2.3)$$
$$D_{k+1} = D_k + \tau[A(t_k)D_k + D_k A'(t_k) - D_k Q'(t_k)B^{-1}(t_k)Q(t_k)D_k]$$

We denote by $\xi(t)$ and $D(t)$ the mathematical expectation and correlation matrix for the phase vector $x(t)$ computed at the instant t. Then $x_k = \xi(t_k)$, $D_k = D(t_k)$ for $k = 0, 1, \ldots, N$. Moreover, we denote by $y(t)$ the result of measurements at the instant t, so that $y_k = y(t_k)$ for $k = 0, 1, \ldots N$. Taking the limit as $\tau \to 0$, we obtain from (2.3) the differential equations

$$d\xi / dt = A\xi + b + DQ'B^{-1}(y - Q\xi) \quad (2.4)$$

$$dD / dt = AD + DA' - DQ'B^{-1}QD \quad (2.5)$$

The initial conditions for Eqs. (2.4), (2.5) are the equations $\xi(t_0) = x_0$, $D(t_0) = D_0$, which specify the mathematical expectation and the correlation matrix for the phase vector $x(t_0)$ before the start of the process. Equation (2.4) is a stochastic differential equation, and Eq. (2.5), which does not depend on the random measurement result, is an ordinary differential equation.

If the measurements are sufficiently frequent, then it is convenient to consider differential equations (2.4), (2.5) instead of finite-difference equations (1.3), (1.7). The function $B(t)$ characterizes the measurement error per unit time. However, Eqs. (2.4), (2.5) can also be used when the measurements are made at discrete instants. In this case the function $B^{-1}(t)$ is impulsive (of the delta function type).

We note that $B(t)$ is an $l \times l$ square matrix, and that Q is an $l \times n$ rectangular matrix. The number $l(t)$ is equal to the number of scalar parameters measured at the instant t and can vary in the course of motion.

Equations (2.4), (2.5) are also valid when the matrix A and the vector b in system (1.1) depend on the control or on the external perturbations. It is important to note that Eq. (2.5) does not depend on the function b, which can be arbitrary.

3. Tracking control. Let us write Eq. (2.5) in the form

$$dD/dt = AD + DA' - DVD \qquad (V = Q'B^{-1}Q) \quad (3.1)$$

It is easy to see that like the matrices B and D, the $n \times n$ matrix V is symmetric and positive-definite. It characterizes the accuracy or "intensity" of the tracking process ($V = 0$ if no observations are made). The matrix V depends on how many and which parameters are measured (this is determined by the matrix Q), and on the error brackets of these measurements (these are characterized by the matrix B).

If the tracker is able to vary the selection of observed parameters or the accuracy of their measurement, then the matrix V in Eq. (3.1) can be regarded as a controlling function. It can be subjected to the restrictions

$$V(t) \in U(t) \quad (t_0 \leqslant t \leqslant T) \tag{3.2}$$

where $U(t)$ is a closed set of matrices characterizing the tracker's freedom of choice. Let us also introduce the integral functional

$$J_0 = \int_{t_0}^{T} f(V, t)\, dt \tag{3.3}$$

Here the scalar function f is defined for all $t \in [t_0, T]$, $V \in U$ and characterizes the cost of the observations.

For example, let the set U for any t consist of two fixed matrices O and V_0, and let the function f be given by the relations

$$f(0, t) = 0, \qquad f(V_0, t) = 1$$

In this case the tracker can either conduct his observations in a fixed way (using the matrix V_0) at every instant, or he can refrain from observation. The functional J of (3.3) then simply represents the total duration of the tracking process.

The purpose of tracking is usually to determine the values of certain functions of the phase coordinates at specific instants with a prescribed or a minimal error. Let $T_1, \ldots, T_m$ be instants specified in the interval $[t_0, T]$, and let $z_1, \ldots, z_m$ be the scalar parameters of interest to the tracker at these instants. Some of the quantities T_i may coincide, which means that several parameters are of interest at some of the instants. Limiting ourselves, as in system (1.1), to a linear approximation (in the neighborhood of some nominal trajectory), we represent the parameters z_i as linear functions of the phase coordinates,

$$z_i = (q_i, x(T_i)) + a_i \qquad (i = 1, \ldots, m) \tag{3.4}$$

Here q_i are specified n-dimensional vectors and a_i are constants. Recalling that the correlation matrix for the vector $x(T_i)$ is equal to $D(T_i)$, we apply certain familiar rules [4] to (3.4) to obtain the dispersion J_i of the quantity z_i,

$$J_i = \sum_{j,\,k=1}^{n} D^{jk}(T_i)\, q_i^j q_i^k \qquad (i = 1, \ldots, m) \tag{3.5}$$

Here and below the superscripts represent the numbers of elements in the vectors and matrices. We note that the quantities given by (3.5) are linear with respect to the elements of the matrix D, which play the role of the phase coordinates. Functions (3.5) characterize the errors involved in determining the parameters of interest.

Now let us formulate some variants of the optimal tracking problem. We can pose the problem of finding the control $V(t)$ which satisfies restrictions (3.2) for all $t \in [t_0, T]$ and minimizes functional (3.3) under the condition that functionals (3.5) assume specified values. The phase coordinates $D(t)$ are determined by Eqs. (3.1) and the initial condition $D(t_0) = D_0$. This problem is a conventional optimal problem with an integral functional, a restriction on the control, and linear multipoint boundary conditions. The number of phase coordinates and controlling functions, i.e. the number of elements in the matrices D and V is n^2. However, since these matrices are symmetric, we need consider only $n(n+1)/2$ of their components.

Instead of minimizing the functional J_0 of (3.3) (the tracking cost) we can require minimization of one of the functionals J_{i_*} of (3.5), i.e. we can minimize the error of determining one of the parameters (3.4). The remaining functions (3.5), as well as functional (3.3) may be given in this case.

If the measurements are made at discrete instants, then

$$V(t) = \sum_{k=1}^{r} V_k(t)\delta(t-t_k) \tag{3.6}$$

Here r is the number of measurements, δ is the delta function, $V_k(t)$ are the prescribed matrix functions (which may be constant), and t_k are the instants of measurement. We can pose the problem of optimal choice of the numbers t_k from the interval $[t_0, T]$ or from a part of this interval in such a way as to minimize one of the functionals of the form (3.5) (possibly for given values of the other functionals). This problem is one in nonlinear programming. Still other formulations of tracking process optimization are possible. For example, we can consider discrete optimal tracking problems (with the aid of Eqs. (1.3) and (1.7)) as multistep discrete controlled processes.

4. Analysis of the equations. The problems posed in Sect. 3 can be simplified considerably. Nonlinear system (3.1) can be reduced to a linear one by the substitution of variables $D = Y^{-1}$. Differentiating the identity $DY = E$ (E is a unit matrix), we obtain

$$dY/dt = -Y(dD/dt)Y \qquad (Y = D^{-1}) \tag{4.1}$$

Substituting Eq. (3.1) into (4.1), we obtain

$$dY/dt = -A'Y - YA + V, \qquad Y(t_0) = Y_0 = D^{-1}(t_0) \tag{4.2}$$

Making use of Pontriagin's maximum principle [5], we construct the Hamiltonian H for the optimal tracking problem with functional (3.3) and Eqs. (4.2),

$$H = P*(-A'Y - YA + V) + p_0 f(V, t) \quad \left(A_1 * A_2 = \sum_{j,k=1}^{n} A_1^{jk} A_2^{jk}\right) \tag{4.3}$$

Here p_0 is a constant, P is the matrix of associated variables (a symmetric $n \times n$ matrix), and the asterisk denotes the scalar (element-by-element) multiplication of matrices. The corresponding associated system is of the form

$$dP/dt = AP + PA' \tag{4.4}$$

Let our problem be that of minimizing functional (3.3) under restriction (3.2) with the constant closed set U and the following conditions imposed on functionals (3.5) at the end of the process:

$$J_i = \sum_{j,k=1}^{n} D^{jk}(T) q_i^j q_i^k = \sum_{j,k=1}^{n} [Y^{-1}(T)]^{jk} q_i^j q_i^k = c_i \quad (i = 1, \ldots, m;\ m < n) \tag{4.5}$$

Here c_i are given constants, and the superscripts indicate the number of elements. By the maximum principle [5], the optimal tracking problem is reducible to a boundary value problem for system (4.2), (4.4) under boundary conditions (4.2), (4.5) and the transversality conditions (λ_i are constants)

$$P^{jk}(T) = \sum_{i=1}^{m} \lambda_i \frac{\partial J_i}{\partial Y^{jk}} \qquad (j, k = 1, \ldots, n) \tag{4.6}$$

The control V can be eliminated with the aid of the maximum condition for the function H of (4.3), i.e.

$$P * V + p_0 f(V, t) \to \sup \text{ with respect to } V \in U \qquad (t_0 < t < T) \tag{4.7}$$

Here we can usually set $p_0 \equiv -1$. We compute the derivatives in (4.6) by means of an identity similar to (4.1),

$$\partial D / \partial Y^{jk} = -D(\partial Y / \partial Y^{jk}) D \qquad (j, k = 1, \ldots, n)$$

Recalling the symmetry of the matrices D, Y, we immediately infer from this that

$$\partial D^{rs} / \partial Y^{jk} = -D^{jk} D^{ks} \qquad (r, s, j, k = 1, \ldots, n)$$

Substituting this equation and relation (4.5) into condition (4.6), we obtain

$$P^{jk}(T) = -\sum_{i=1}^{m} \lambda_i \sum_{r,s=1}^{n} D^{jr}(T) D^{ks}(T) q_i^r q_i^s =$$
$$= -\sum_{i=1}^{m} \lambda_i [D(T) q_i]^j [D(T) q_i]^k = -\sum_{i=1}^{m} \lambda_i [Y^{-1}(T) q_i]^j [Y^{-1}(T) q_i]^k \tag{4.8}$$
$$(j, k = 1, \ldots, n)$$

The resulting nonlinear matrix boundary value problem (4.2), (4.4), (4.5), (4.7), (4.8) can be reduced to a system of transcendental equations. Let the fundamental matrix $X(t)$ of solutions for initial vector system (1.1) be known. Making use of identity (4.1), we can write out the equations for the fundamental matrix and the matrices associated with it,

$$dX / dt = AX, \quad X(t_0) = E, \quad dX' / dt = X'A'$$
$$dX^{-1} / dt = -X^{-1}A, \qquad d(X')^{-1} / dt = -A'(X')^{-1} \tag{4.9}$$

Now let us consider the matrices

$$X_1(t) = (X')^{-1} C X^{-1}, \qquad X_2(t) = XCX' \tag{4.10}$$

where C is a constant $n \times n$ matrix. Computing the derivatives of matrices (4.10) and making use of relations (4.9), we can readily verify that these matrices satisfy the equations

$$dX_1 / dt = -A'X_1 - X_1A, \qquad dX_2 / dt = AX_2 + X_2A'$$
$$X_1(t_0) = X_2(t_0) = C \tag{4.11}$$

Comparing (4.11) with (4.2) and (4.4), we see that the matrix $X_1(t)$ is the general solution of Eq. (4.4) for the associated variables. Hence,

$$P(t) = X(t)\ CX'(t) \tag{4.12}$$

where the constant matrix C is symmetric by virtue of the symmetry of the matrix $P(t)$.

The solution of inhomogeneous equation (4.2) can be found by the method of variation of the arbitrary constants entering into the matrix X_1 of (4.10). Setting $Y = (X')^{-1}Y_1X^{-1}$, we find with the aid of Eqs. (4.2), (4.9) that

$$(X')^{-1}(dY_1 / dt)X^{-1} = V$$

From this we obtain the matrix Y_1, and then the solution $Y(t)$ of Cauchy problem (4.2) and the matrix $D(t)$ which is the inverse of $Y(t)$,

$$\begin{aligned} Y(t) &= [X'(t)]^{-1}\left[Y_0 + \int_{t_1}^{t} X'(\tau)V(\tau)X(\tau)d\tau\right]X^{-1}(t) \\ D(t) &= X(t)\left[D_0^{-1} + \int_{t_1}^{t} X'(\tau)V(\tau)X(\tau)d\tau\right]^{-1}X'(t) \end{aligned} \tag{4.13}$$

The second equation of (4.3) determines the general solution of nonlinear matrix equation (3.1). Let the function V in Eqs. (4.13) be eliminated with the aid of condition (4.7) in which $p_0 = -1$. Substituting solution (4.13) for $D(t)$ and (4.12) for $P(t)$ into conditions (4.5) and (4.8), we obtain a total of $n^2 + m$ algebraic equations with $n^2 + m$ unknowns (n^2 elements of the constant matrix C and m constants λ_i). The constants λ_i occur linearly in this system and are easy to eliminate. Moreover, recalling that the matrix C is symmetric, we see that the problem can be reduced to a system of $n(n+1)/2$ algebraic equations with the same number of unknowns. In similar fashion it is possible by means of solutions (4.12), (4.13) to reduce other optimal tracking problems to systems of transcendental equations.

For example, let us require minimization of a functional of the form (3.5),

$$J = \sum_{j,k=1}^{n} D^{jk}(T)\, q^j q^k \tag{4.14}$$

under conditions (4.5), restriction (3.2), and a restriction on integral functional (3.3) (the tracking cost or time),

$$J_0 = \int_{t_0}^{T} f(V, t)\, dt = c_0 \tag{4.15}$$

In Eqs. (4.14), (4.15) q is a given n-dimensional vector and c_0 is a constant.

This problem is reciprocal to that considered above. The condition of the maximum principle is of the same form (4.7), but p is now an unknown constant. The transversality conditions can be written in a form similar to (4.8),

$$P^{jk}(T) = [D(T)q]^j [D(T)q]^k - \sum_{i=1}^{m} \lambda_i [D(T)q_i]^j [D(T)q_i]^k \qquad (4.16)$$
$$(j, k = 1, \ldots, n)$$

We have $n^2 + m + 1$ conditions (4.5), (4.15), (4.16) for determining $n^2 + m + 1$ constants C, λ_i, p_0 into which we must substitute solutions (4.12), (4.13) for $P(t)$ and $D(t)$ and eliminate V by means of condition (4.7). If the problem posed does not include restriction (4.15), then we must set $p_0 = 0$ in expression (4.7).

Various problems of tracking optimization by the choice of discrete instants of measurement, i.e. in case (3.6), can also be simplified by means of solution (4.13). Substituting (3.6) into (4.13), we obtain

$$D(T) = X(T)\left[D_0^{-1} + \sum_{k=1}^{r} X'(t_k) V_k(t_k) X(t_k)\right]^{-1} X'(T) \qquad (4.17)$$

The problem of the minimum of functional (4.14) for tracking method (3.6) can now be reduced by substituting expression (4.17) into Eq. (4.14) to the problem of minimizing a function of the variables t_k.

5. Examples. 1. Let equation of motion (1.1) be of the simplest form

$$dx / dt = ax + b(t) \qquad (5.1)$$

where x is the only phase coordinate $(n = 1)$, a is a constant, and $b(t)$ is a function of time. The tracking process consists in measuring the instantaneous value of the phase coordinate; the measurement error dispersion per unit time is either equal to a constant b_0 at every instant, or no measurements are made. The sum duration of the observations is given and equal to $T_0 < T$, where T is the duration of the tracking process. Without limiting generality we can also set $t_0 = 0$.

We are required to indicate a control method which will minimize the dispersion $D(T)$ of the phase coordinate $x(T)$ at the end of the process. In the notation of Sects. 2 and 3 we have $l = 1$, $Q = 1, V = B^{-1}$. The matrices $D, Q, B, V,$ and P here become scalars, and the set U' of (3.2) consists of two points: 0 and b_0^{-1}; the function f of (3.3) is of the form

$$f = 0 \quad \text{for } V = 0, \; f = 1 \quad \text{for } V = b_0^{-1} \qquad (5.2)$$

The vector q in functional (4.14) reduces to the scalar $q = 1$ and condition (4.5) does not apply, so that $m = 0$. The fundamental matrix $X(t)$ defined by relations (4.9) reduces to the scalar function $X(t) = e^{at}$, and solutions (4.12), (4.13) become

$$P(t) = ce^{2at} \qquad D(T) = e^{2aT}\left[D_0^{-1} + \int_0^T e^{2a\tau} V(\tau)\, d\tau\right]^{-1} \qquad (5.3)$$

Here c and D_0 are constants. Condition (4.16) yields

$$P(T) = ce^{2aT} = D^2(T) \qquad (5.4)$$

Recalling Eqs. (5.2) and (5.3) for P, we apply maximum principle (4.7) to obtain

$$V = b_0^{-1} \quad \text{for} \quad \varphi(t) > 0, \qquad V = 0 \quad \text{for} \quad \varphi(t) < 0,$$
$$\varphi = ce^{2aT}b_0^{-1} + p_0 \tag{5.5}$$

Since $c > 0$ by virtue of (5.4), it follows by (5.4) that the function $\varphi(t)$ of (5.5) increases monotonically for $a > 0$ and decreases monotonically for $a < 0$. Hence, by (5.5), the optimal tracking law is

$$V = 0 \quad \text{for} \quad t < \theta, \qquad V = b_0^{-1} \quad \text{for} \quad t > \theta \quad (a > 0) \tag{5.6}$$
$$V = b_0^{-1} \quad \text{for} \quad t < \theta, \quad V = 0 \quad \text{for} \quad t > \theta \quad (a < 0)$$

Here θ is the unique switching instant corresponding to the root of the monotonic function $\varphi(t)$. Since the sum observation time is T_0, it follows that

$$\theta = T - T_0 \quad \text{for} \quad a > 0, \qquad \theta = T_0 \quad \text{for} \quad a < 0 \tag{5.7}$$

Substituting relations (5.6), (5.7) into (5.3), we obtain

$$D(T) = e^{2aT}\left[D_0^{-1} + {}^1/_2\, b_0^{-1}a^{-1}e^{2aT}\left(1 - e^{-2aT_0}\right)\right]^{-1} \quad (a > 0)$$
$$D(T) = e^{2aT}\left[D_0^{-1} + {}^1/_2\, b_0^{-1}a^{-1}\left(e^{2aT_0} - 1\right)\right]^{-1} \quad (a < 0)$$
$$D(T) = \left(D_0^{-1} + b_0^{-1}T_0\right)^{-1} \quad (a = 0) \tag{5.8}$$

It is clear from (5.3) that for $a = 0$ the functional $D(T)$ to be minimized does not depend on the tracking law and is given by formula (5.8) for any law. Equations (5.6) - (5.8) completely determine the solution of the problem, i.e. the optimal tracking law and the funcional, for all cases. The constants c, p_0 can be found with the aid of Eq. (5.4) and the equation $\varphi(\theta) = 0$, but this is not necessary. The meaning of solution (5.6), (5.7) is obvious: for $a > 0$, when the trajectories of Eq. (5.1) diverge, the measurements are best made at the end of the process; for $a < 0$, when they converge, measurements at the beginning of the process are more advantageous.

2. Let us take the equations of motion of the system in the form

$$dx_1 / dt = x_2, \qquad dx_2 / dt = b(t) \qquad (n = 2) \tag{5.9}$$

Here x_1 is the coordinate, x_2 the velocity, and $b(t)$ the acceleration for a mechanical system with one degree of freedom. Let the measurements be made at discrete instants, the coordinate x_1 with the error dispersion b_1 being measured at the instants t_i, $i = 1, \ldots, r$, and the velocity x_2 with the error dispersion b_2 at the instants t_j', $j = 1, \ldots, r'$. The dispersions d_1, d_2 for the coordinate $x_1(0)$ and the velocity $x_2(0)$, respectively, at the initial instant $t_0 = 0$ are given; the error brackets for these quantities are independent. In the notation of Sects. 1 - 4 we have

$$D_0 = \begin{Vmatrix} d_1 & 0 \\ 0 & d_2 \end{Vmatrix}, \qquad X(t) = \begin{Vmatrix} 1 & t \\ 0 & 1 \end{Vmatrix}$$
$$V = \sum_{i=1}^{r} \begin{Vmatrix} b_1^{-1} & 0 \\ 0 & 0 \end{Vmatrix} \delta(t - t_i) + \sum_{j=1}^{r'} \begin{Vmatrix} 0 & 0 \\ 0 & b_2^{-1} \end{Vmatrix} \delta(t - t_j') \tag{5.10}$$

The fundamental matrix $X(t)$ for system (5.9) is defined in accordance with general relation (4.9). Substituting expressions (5.10) into general solution (4.13), we obtain

$$D(T)=\begin{Vmatrix}1 & T\\ 0 & 1\end{Vmatrix}\cdot\begin{Vmatrix}d_1^{-1}+rb_1^{-1} & b_1^{-1}\sum_1\\ b_1^{-1}\sum_1 & d_2^{-1}+r'b_2^{-1}+b_1^{-1}\sum_2\end{Vmatrix}\cdot\begin{Vmatrix}1 & 0\\ T & 1\end{Vmatrix}$$
$$\sum_1=t_1+\ldots+t_r,\qquad \sum_2=t_1^2+\ldots+t_r^2 \tag{5.11}$$

From (5.11) we see that the matrix $D(T)$ does not depend on the instants t_j' at which the velocities are measured; their choice is therefore arbitrary. Let us substitute variables, setting

$$t_i=T\theta_i,\qquad d_1^{-1}=b_1^{-1}a,\qquad d_2^{-1}+r'b_2^{-1}=b_1^{-1}T^2b \tag{5.12}$$

Here θ_i are the dimensionless instants of measurement of the coordinate, and $a>0, b>0$ are dimensionless constants. After substitutions (5.12) and simplification expression (5.11) can be written as

$$D(T)=\frac{b_1}{T^2[(a+r)(b+y)-x^2]}\begin{Vmatrix}T^2(a+r+b+y-2x) & T(a+r-x)\\ T(a+r-x) & a+r\end{Vmatrix}$$
$$x=\theta_1+\ldots+\theta_r,\qquad y=\theta_1^2+\ldots+\theta_r^2 \tag{5.13}$$

Let us pose the following optimal tracking problem. We are to choose the instants t_i of coordinate measurement in such a way as to minimize the dispersion of the velocity $x_2(T)$ at the end of the process. Let us construct the appropriate functional of the form (4.14), making use of Eq. (5.13),

$$J=\frac{b_1(a+r)}{T^2[(a+r)b+\psi]},\qquad \psi=(a+r)y-x^2 \tag{5.14}$$

The problem of a minimum velocity dispersion, i.e. that of minimizing expression (5.14), reduces to the problem of maximizing the function

$$\psi(\theta_1,\ldots,\theta_m)=(a+r)\sum_{i=1}^{r}\theta_i^2-\Big(\sum_{i=1}^{r}\theta_i\Big)^2 \tag{5.15}$$

under the conditions $0\leqslant\theta_i\leqslant 1$, $i=1,\ldots,r$. Applying the Cauchy-Buniakowski inequality, we obtain

$$\psi=a\sum_{i=1}^{r}\theta_i^2+\Big(\sum_{i=1}^{r}1^2\Big)\Big(\sum_{i=1}^{r}\theta_i^2\Big)-\Big(\sum_{i=1}^{r}\theta_i\Big)^2\geqslant a\sum_{i=1}^{r}\theta_i^2\geqslant 0$$

This implies that ψ is a positive-definite quadratic form in θ_i. It is clear that its maximum on the r-dimensional cube $0\leqslant\theta_i\leqslant 1$, $i=1,\ldots,r$, occurs at one of the vertices of the cube. Hence, some of the quantities θ_i in the optimal solution are equal to zero, and the rest are equal to unity; at least one of the θ_i is equal to unity, since the point $\theta_i=0$, $i=1,\ldots,r$ minimizes function (5.15).

We therefore have

$$t_i = T\theta_i = 0 \text{ for } 1 \leqslant i \leqslant r - k$$
$$t_i = T\theta_i = T \quad \text{for} \quad r - k < i \leqslant r \qquad (1 \leqslant k \leqslant r) \tag{5.16}$$

Here k is the number of measurements at the end of the process. To determine k we substitute (5.16) into (5.15),

$$\psi = (a + r)\, k - k^2 \tag{5.17}$$

and then find the maximum of function (5.17) with respect to integers k from the interval $1 \leqslant k \leqslant r$. The maximum of (5.17) occurs for the integer k closest to $(a + r)\,/\,2$ which does not exceed r. Denoting the whole part of a number by square brackets, we write

$$k = \min\{r,\ [(a + r + 1)\,/\,2]\},\ a = b_1 d_1^{-1} \tag{5.18}$$

Relations (5.16) - (5.18) completely determine the optimal tracking law which reduces to the fact that k measurements are made at the end, and remaining measurements at the beginning, of the process. Substituting expression (5.17) into (5.14) we obtain the value of the functional to be minimized (i.e. the dispersion of the velocity error at the end of the process).

$$J = \frac{b_1 (a + r)}{T^2 [(a + r)(b + k) - k^2]}, \quad a = b_1 d_1^{-1}, \quad b = \frac{d_2^{-1} + r' b_2^{-1}}{b_1^{-1} T^2}$$

BIBLIOGRAPHY

1. Fel'dbaum, A.A., Principles of Optimal Automatic Systems Theory. Moscow, Fizmatgiz, 1963.
2. Krasovskii, N.N., The Theory of Motion Control. Moscow, "Nauka", 1968.
3. Lidov, M.L., Optimal pulse correction in linear systems with incomplete information. (p.192) Third All-Union Congress on Theoretical and Applied Mechanics: Abstracts of Papers. Moscow, 1968.
4. Kramer, G., The Mathematical Methods of Statistics. Moscow, Izd. inostr. lit., 1948.
5. Pontriagin, L.S., Boltianskii, V.G., Gamkrelidze, R.V. and Mishchenko, E.F., The Mathematical Theory of Optimal Processes. Moscow, Fizmatgiz, 1961.

Translated by A. Y.