ADAPTIVE FUNCTIONING MECHANISMS OF ACTIVE SYSTEMS.

III. LEARNING AND CONTROL

V. V. Tsyganov

UDC 62-506

We consider the design of adaptive functioning mechanisms for two-level active systems in which the current operating information provided by the active elements is utilized to alter the functioning mechanism parameters and to achieve the headquarters objective. We pose the design problem and establish constructive sufficient conditions for the existence of learning mechanisms in which headquarters learning overlaps the divisional planning and incentive stages. We also consider an approach to the design of dual mechanisms, which additionally optimize some control performance criterion. Sufficient conditions of optimality are obtained.

1. Introduction

Adaptive identification and control of complex hierarchical systems must take into account the human factor or the active behavior of the system elements attributable to the existence of their individual objectives [1-5]. A far-sighted active element (AE), or division, in a system may predict the headquarters controls and choose its states so as to influence the outcome of the identification and control process while maximizing its own objective function (see formula (1) in [4]), which is dependent on the incentive procedures in each period.

For instance, with standard learning and control procedures [6-9], the headquarters uses the current operating information supplied by the divisions to alter the functioning mechanism parameters and to achieve the headquarters objective. Specifically, the control is established as a function of current learning – a recursive estimator of the active system parameters (a_t): $a_t = I_t(a_{t-1}, y_t)$, $t = 0, 1, \ldots, a_{-1} = a^0$. Here I_t is the learning procedure, I is the AE state as observed by the headquarters in period t. The learning procedure is known and the active element chooses its state from a set of solutions of a game [4, 5] so as to maximize in some sense its own objective function. This state, in general, is different from the state that would obtain without active behavior. Therefore, direct application by the headquarters of adaptation and learning procedures developed for automatic systems is ineffective. The problem is to construct functioning mechanisms $\Sigma = (I, \pi, Q, f)$ with learning (I), planning (π), control (Q), and incentive (f) procedures which ensure effective adaptation in active systems.

In [3-5] we have examined the construction of adaptive mechanisms which ensure structural identification of a vector AE dependent on a scalar parameter – the potential p. For the case of an AE with deterministic structure, we obtained the strong progressivity condition [3, 4], maximizing the convergence rate of the parameter estimator to the potential ($a_t = p$ for any $t \geq 0$). A weaker sufficient condition for a scalar AE was derived in [5].

The construction of the functioning mechanism of an independent scalar AE with nondeterministic structure for the case of reference model adaptive control was considered in Chapter 4 of [3]. Sufficient conditions for effectiveness of the mechanism were obtained, including constraints on identification and control procedures.

In this article, we examine the construction of adaptive functioning mechanisms (AFM) for interdependent vector AEs with nondeterministic structure determined by a random vector parameter, assuming various headquarters objectives. The notation is the same as in [4, 5].

2. The Solution of the Divisional Game

Consider an active system comprising N interdependent AEs. As previously, we denote the state of the i-th AE in period t by $y_t^i = (y_{1t}^i, \ldots, y_{mt}^i)$, $i = \overline{1, N}$ and the state of

Moscow. Translated from Avtomatika i Telemekhanika, No. 3, pp. 117-126, March, 1987. Original article submitted January 16, 1986.

 0005-1179/87/4803-0368$12.50

the system by $y_t=(\mathbf{y}_t^1, \ldots, \mathbf{y}_t^N)$. Let m = N and the headquarters observes only one component of the state vector of each AE (the divisional "output"). We denote the observed component for the i-th AE by $\tilde{y}_{it} \equiv y_{it}^i$, $i = \overline{1, N}$. The remaining divisional state components y_{jt}^i may be interpreted as the decentralized resources allocated by the i-th AE to the j-th AE prior to the beginning of the period t, $i, j = \overline{1, N}$, $i \neq j$, $y_{jt}^i \in [0, p_{jt}^i]$, where p_{jt}^i is a random parameter. The state $\tilde{y}_{it}$ observed by the headquarters is constrained by the i-th division's own potential — the random parameter p_{it}^i, and also by the resources allocated to the i-th AE by the other divisions,

$$\tilde{y}_{it} \in [0, W_t^i(\mathbf{u}_t^i)],\ \mathbf{u}_t^i = (y_{it}^1, \ldots, y_{it}^{i-1}, p_{it}^i, y_{it}^{i+1}, \ldots, y_{it}^N), \tag{1}$$

where W_t^i is a monotone increasing function of its arguments. Then the set of admissible states of the i-th AE is $Y_t^i(\mathbf{u}_t^i)=[0, W_t^i(\mathbf{u}_t^i)] \times \prod_{i=1}^{N} [0, p_j^i]$. It is easy to see that $W_t^i(\mathbf{u}_t^i) \leqslant z_t^i \equiv W_t^i(\mathbf{p}_t^i)$ where z_t^i is the observed state of the i-th AE in period t without active behavior (in short, the limiting state), $\mathbf{p}_t^i = (p_{1t}^i, \ldots, p_{Nt}^i)$ is its potential, $i = \overline{1, N}$. The set of feasible values of the composite parameter $p_t=(\mathbf{p}_t^1, \ldots, \mathbf{p}_t^N)$ — the system potential in period t — will be denoted by P_t.

We assume that the system functions in the following sequence in period t. First the headquarters, applying some preselected AFM to the past observations $\tilde{\mathbf{y}}_\tau=(\tilde{y}_{1\tau}, \ldots, \tilde{y}_{N\tau})'$,* $\tau < t$, estimates the active system parameters a_{t-1} and assigns the target vector $\mathbf{x}_t$ (the norms for awarding incentives) and the control vector $\mathbf{r}_t$ (the allocation of centralized resources). Then, given the realization of the random parameter $\mathbf{p}_t$, the i-th AE chooses the states y_{jt}^i, $i \neq j$, $1 \leq j \leq N$ and communicates this information to the other AEs. Then it selects its observed state $\tilde{y}_{it}$, with full knowledge of the actual realization of $W_t^i(\mathbf{u}_t^i)$, and collects the incentive $\varphi_t^i=f_t^i(\mathbf{x}_t, \tilde{\mathbf{y}}_t)$.

Let us now determine the solution of the divisional game. According to [5], the divisions optimize their own objective functions with potential, state, and mechanism prediction. The i-th AE uses its own predictors of the set of feasible system potentials $P_{i\tau}$ ($P_{i\tau} \subset P_\tau$) and also predictors of the feasible choice sets of the k-th AE, $Y_{i\tau}^k(\mathbf{u}_\tau^k)$, $k=\overline{1, N}$ ($p_\tau \in P_{i\tau}$, $Y_{i\tau}^k(\mathbf{u}_\tau^k) \subset Y_\tau^k(\mathbf{u}_\tau^k)$), $\tau=\overline{t+1, t+T}$, $i=\overline{1, N}$. From these predictors, the i-th AE calculates its guaranteed payoff in system state y_t:

$$\hat{w}_t^i(x_t, y_t) = \min_{p_\tau \in P_{i\tau},\ \tau=\overline{t+1, t+T}}\ \min_{\mathbf{y}_\tau^k \in Y_{i\tau}^k,\ k=\overline{1, N}} w_t^i(\varphi_t^i, \ldots, \varphi_{t+T}^i). \tag{2}$$

The solution of the divisional game is a Nash equilibrium. Specifically, the set of solutions of the game, regarded as the set of system states maximizing the guaranteed outcome for each division, has the form

$$R_t(\Sigma, p_t)=\{y_t^* \mid \hat{w}_t^i(x_t, y_t^*) \geqslant \hat{w}_t^i(x_t, y_t),\ \mathbf{y}_t^k \in Y_t^k(\mathbf{u}_t^k),\ i, k=\overline{1, N}\} \tag{3}$$

As in [3, 5], we assume that each AE extrapolates the relevant AFM to its far-sightedness horizon (see (3) in [5]): $\mathbf{a}_\tau=I_t(\mathbf{a}_{\tau-1}, y_\tau)$, $x_{\tau+1}=\pi_{t+1}(\mathbf{a}_\tau)$, $r_\tau=Q_t(\mathbf{a}_\tau)$, $\varphi_\tau^i=f_t^i(x_\tau, \tilde{y}_\tau)$, $\tau=\overline{t, t+T}$. The AE is moreover assumed to be friendly toward the headquarters: if $R_t(\Sigma, p_t) \ni v_t=(\mathbf{v}_t^1, \ldots, \mathbf{v}_t^N)$, $\mathbf{v}_t^i=(p_{1t}^i, \ldots, p_{i-1t}^i, z_t^i, p_{i+1t}^i, \ldots, p_{Nt}^i)$, then $y_t^* = v_t$. This means that the AEs will not lower their performance indexes if this is not beneficial for them. The vector v_t^i is the state of the i-th division without active behavior. Let us now investigate the design of adaptive mechanisms under various assumptions regarding the headquarters objectives.

3. Design of a Learning Mechanism: the Problem

In a learning functioning mechanism, the headquarters in an active system overlaps learning [6-8] with planning and awarding of incentives to the active elements. Suppose that the vectors $\mathbf{p}_t^i$ are functions of the input stimulus r_t observed by the headquarters, corrupted by some random unobservable noise ξ_t. Then the AE limiting state z_t^i is also a random function of r_t. Assuming that all the divisions function in a stationary mode (or normal operating mode) [8], let us consider an adaptive model of the i-th division constraints in the form

$$\hat{z}_t^i = a_{t-1}^i q^i(\mathbf{r}_t),\ \hat{\mathbf{z}}_t = (\hat{z}_t^1, \ldots, \hat{z}_t^N)',\ \mathbf{a}_t = (a_t^1, \ldots, a_t^N)', \tag{4}$$

*The prime denotes the transpose.

where $\hat{z}_t^i$ is the limiting state estimator of the i-th AE in period t, q^i is some function of the observed input stimuli (external resources) r_t chosen by the headquarters from more general considerations; a_t^i is a tunable parameter, $i = \overline{1, N}$. Suppose that while tuning the parameter a_t the headquarters observes the limiting state $z_t = (z_t^1, \ldots, z_t^N)'$ and its objective is to minimize the mean losses $J_t(a) = M_\xi\{\Phi_t|_{a_{t-1}=a}\}$, where the loss function $\Phi_t = \Phi(\varepsilon_t)$ is dependent on the error $\varepsilon_t = z_t - \hat{z}_t$, M is the expectation operator, Φ is a convex twice differentiable function. The corresponding parameter tuning procedure (or learning procedure) has the form [6-8]

$$\mathbf{a}_t = \mathbf{a}_{t-1} - A\boldsymbol{\gamma}_t \nabla_{\mathbf{a}} \Phi_t \equiv I_t(\mathbf{a}_{t-1}, \mathbf{z}_t), \tag{5}$$

where A is an operator which transforms the gains γ_t into a diagonal matrix. Let $J_t(\mathbf{a}, z^t) = \frac{1}{t}\sum_{\tau=0}^{t} \Phi_\tau|_{\mathbf{a}_{\tau-1}=\mathbf{a}}$ be the empirical mean losses which characterize the learning performance, $z^t = (\mathbf{z}_0, \ldots, \mathbf{z}_t)$. It is assumed that the optimal sample estimator of the vector parameter $\mathbf{a}_t$ converges in probability to the optimal estimator $\mathbf{a}^*$,

$$\mathbf{a}_t = \operatorname*{argmin}_{\mathbf{a}} J_t(\mathbf{a}, z^t) \to \mathbf{a}^* = \operatorname*{argmin}_{\mathbf{a}} J_t(\mathbf{a}). \tag{6}$$

If the divisions are active, the choose their states from the set of solutions of the game (3). Therefore, instead of $\mathbf{z}$, the headquarters actually observes the vector $\tilde{\mathbf{y}}_t^* = (y_{1t}^*, \ldots, y_{Nt}^*)$, so that the parameter estimators

$$\mathbf{a}_t = I_t(\mathbf{a}_{t-1}, \tilde{\mathbf{y}}_t^*), \tag{7}$$

obtained from the procedure (5) with the observations $\tilde{\mathbf{y}}_t^*$ in general do not converge to the optimal estimator (6).

The design of learning functioning mechanism (LFM) is thus posed as a problem of choosing a mechanism $\Sigma = (I, \pi, f)$ with learning procedure $I = \{I_t\}$, planning procedure $\pi = \{\pi_t\}$, and incentive procedure $f = \{f_t\}$ such that $\mathbf{a}_t = I_t(\mathbf{a}_{t-1}, \tilde{\mathbf{y}}_t^*)_t \to \mathbf{a}^*$, $y_t^* \in R_t(\Sigma, p_t)$.

4. Design of a Learning Mechanism: the Solution

Consider the mechanism $\Sigma = (I, \pi, f)$ with the learning procedure (7) and the following *planning and incentive procedures:*

$$\mathbf{x}_{t+1} = \mathbf{a}_t, \quad \varphi_t^i = f_t^i(\mathbf{x}_t, \tilde{\mathbf{y}}_t). \tag{8}$$

In what follows, as in [5], we assume that f_t^i is differentiable and monotone increasing in each component of the vector $\tilde{\mathbf{y}}$ and decreasing in the components of $\mathbf{x}$. Following [5], we use the notation

$$d\xi = \min_{\Omega_t} \frac{\partial}{\partial \xi}, \quad D\xi = \max_{\Omega_t} \frac{\partial}{\partial \xi}, \quad \Omega_t = \{v = \overline{t, t+T}; \quad p_\tau \in P_\tau,$$

$$\tau = \overline{t, v}; \quad \mathbf{y}_\tau^k \in Y_\tau^k(\mathbf{u}_\tau^k), k = \overline{1, N}\}, \underline{\nabla}_{\mathbf{x}} = (dx^1, \ldots, dx^N)',$$

$$\mathbf{F}_t^i = -\underline{\nabla}_{\mathbf{x}}' \varphi_v^i, \quad \mathbf{G}_{jt} = d\tilde{y}_j \nabla_{\mathbf{x}} \Phi_v, \quad H = \nabla_{\mathbf{x}} \nabla_{\mathbf{x}} \Phi_v,$$

E is the N × N identity matrix. Then we have

THEOREM 1. Let $\Sigma = (I, \pi, f)$ be defined by (7), (8) so that it satisfies the condition

$$\underline{D}_{jt}^i(\Sigma) = (d\varphi_t^i w^i)(d\tilde{y}_j f_t^i) - \mathbf{F}_t^i \sum_{\tau=1}^{T} (D\varphi_{\tau+T}^i w_t^i)(E - A\boldsymbol{\gamma}_t H)^{\tau-1} \times A\boldsymbol{\gamma}_t \mathbf{G}_{jt} \geqslant 0 \quad \text{for any } i, j, t.$$

Then Σ is a learning mechanism.

Theorems 1 and 2 are proved in the Appendix.

Substantively, this theorem establishes the progressivity conditions of an AFM [3, 4] which ensure that the divisions choose the state $y_t^* = v_t$ (i.e., do not lower their performance indexes). This permits tuning the parameters of the AE constraint model. We can similarly derive constructive progressivity conditions for other iterative procedures and thus solve analysis and design problems for AFMs with standard learning procesures as subsystems (including identification, estimation, and classification procedures) [6-8, 10].

Let us consider, for example, the problem of optimal progressive learning or, briefly, p-learning, which chooses the vectors γ_t, $t = 0, 1, \ldots$, satisfying the conditions of Theorem

1 and minimizing $J_t(a, \tilde{y}^t)$. In general, this is a fairly complex nonlinear programming problem which cannot be solved analytically. In order to obtain an approximate quasioptimal solution (γ_t^P) assume that the norm of γ_t is sufficiently small, $\|\gamma_t\| << 1$ [7]. Then the conditions of Theorem 1 take the form

$$\bar{D}_{jt}{}^i(\Sigma)=\alpha_{jt}{}^i-(\beta_{jt}{}^i, \gamma_t)\geqslant 0, \quad \alpha_{jt}{}^i=(d\varphi_t{}^i w_t{}^i)(d\tilde{y}_j f_t{}^i),$$
$$\beta_{jt}^i=\sum_{\tau=t+1}^{t+T}(D\varphi_{t+T}^i w_t^i)\left[(dx_1 f_t^i)\left(d\tilde{y}_j\frac{\partial\Phi_t}{\partial x_1}\right),\ldots,(dx_N f_t^i)\left(d\tilde{y}_j\frac{\partial\Phi_t}{\partial x_N}\right)\right] \tag{9}$$

Note that $\alpha_{jt}{}^i \geqq 0$, $\beta_{jt}{}^i \geqq 0$ due to the natural assumptions of monotonicity of $w_t{}^i$, $f_t{}^i$ used in [4, 5] and in this paper. The Kuhn–Tucker conditions for the determination of γ_t^P have the form

$$\nabla_{\gamma t} J_t(a_t, y^t)+\sum_{i,j=1}^{N}\lambda_{jt}^i\beta_{jt}^i=0, \tag{10}$$

$$(\beta_{jt}^i, \gamma_t)\leqslant\alpha_{jt}^i, \qquad \sum_{i,j=1}^{N}[(\beta_{jt}^i,\gamma_t)-\alpha_{jt}^i]\lambda_{jt}^i=0,\ \lambda_{jt}^i\geqslant 0. \tag{11}$$

From (10), using the technique developed in [7, Sec. 3.4], we can easily find that $\gamma_t^P = \gamma_{t\,qopt} - \delta_t$, where $\gamma_{t\,qopt}$ is the quasioptimal vector of gains without active behavior [7].

$$\gamma_{t\,q\,opt}=[A\nabla_a\Phi_t]^{-1}H_tA\nabla_a\Phi_t \text{ and } \delta_t=[A\nabla_a\Phi_tH_t^{-1}A\nabla_a\Phi_t]^{-1}\times$$
$$\times\sum_{i,j=1}^{N}\lambda_{jt}^i\beta_{jt}^i, \qquad H_t=\sum_{\tau=1}^{t}\nabla_a^2\Phi_\tau|_{a_{\tau-1}=a_{t-1}}.$$

Noting that $A\nabla_a\Phi_t > 0$ for $\beta_{jt}{}^i > 0$, we obtain $\gamma_t^P \leqq \gamma_{tqopt}$. By (11), for $\lambda_{jt}{}^i = 0$, $\bar{D}_{jt}{}^i(\Sigma) = \alpha_{jt}{}^i - (\beta_{jt}{}^i, \gamma_t) > 0$, i.e., when the conditions of Theorem 1 are a priori satisfied, we have $\gamma_t^P = \gamma_{tqopt}$. Conversely, for $\lambda_{jt}{}^i > 0$, $\gamma_t^P > \gamma_{tqopt}$, i.e., the gains are bounded. For instance, in the simple case when

$$m=N=1, \quad \Phi(\varepsilon_t)=\varepsilon_t{}^2 \tag{12}$$

we have $\gamma_{tqopt} = 1/t \geq \gamma_t^P = \min(1/t, \alpha_t/\beta_t)$ (the subscripts i, j ≡ 1 have been omitted). Substantively, this implies a bounded "step" in adaptive planning in (8). In other words, when the planned targets are too far from "the attained level," the AEs no longer have a motivation to disclose their reserves. This result restricts the applicability of optimal learning algorithms in automatic systems.

The p-learning algorithms may be implemented if they are convergent. Using the known sufficient conditions for convergence of the original algorithm (5) [7] and Theorem 1, we can obtain sufficient conditions for convergence of p-learning algorithms in the form of constraints on the LFM (Σ), in particular on the incentive procedure (f).

For instance, when (12) holds, they are obtained from (9) in the form $\sum_{t=0}^{\infty}\gamma_t^P\leqslant\sum_{t=0}^{\infty}(\alpha_t/\beta_t)=\infty$. If $\sum_{t=0}^{\infty}(\alpha_t/\beta_t)<\infty$, then the procedure (5) for $\gamma_t = \gamma_t^P$ in general is divergent. This implies that in case of small incentives for state improvement (small α_t, see (9)) or high penalties for plan underperformance (large β_t), a far-sighted division will find it advantangeous to choose the limiting state only for very small γ_t^P (i.e., a small adaptive planning "step"), but then the learning procedure (5) does not converge.

5. Design of a Dual Mechanism: the Problem

So far, we have assumed that the external resources $\mathbf{r}_t$ allocated to the active elements (i.e., the headquarters controls) are independent of the information accumulated in the process of learning.

With dual control, the headquarters combines learning with the control of the active system [6, 9]. In this case, the optimal control is of dual nature: it optimizes a given performance criterion while at the same time enhancing the accumulation of information about unknown parameters. The solution of the dual control problem is extremely complex even for simple automatic systems [6, 9]. For simplicity, we take N = 1 and drop the AE subscript (i ≡ 1). Here $\tilde{y}_t = y_t$, $\mathbf{u}_t = \mathbf{p}_t$ (see Sec. 2). Note that with dual control, the active element should take into account the influence of its chosen state (y_t) not only on future parameter estimators $\mathbf{a}_\tau$ and targets $\mathbf{x}_\tau$ (as in the case of learning mechanisms, see Sec. 3) but also on future feasible state sets $Y_\tau(\mathbf{p}_\tau)=[0, W_\tau(\mathbf{p}_\tau)]$ which are dependent on the resource allocation — the headquarters controls $\mathbf{r}_\tau$, $\tau > t$. In order to investigate this influence, let $\mathbf{p}_t=(\mathbf{q}_t', \xi_t)'$, where $\mathbf{q}_t=(q_{1t}, \ldots, q_{lt})'$ is a vector function of the resources $\mathbf{r}_\tau=(r_{1\tau}, \ldots, r_{n\tau})'$ allocated by the headquarters to the division in periods $\tau=\overline{0, t-1}$: $q_{it}=q_i(\mathbf{r}_0, \ldots, \mathbf{r}_{t-1})$, $i=\overline{1, l}$; ξ_t is unobservable random noise, $\xi_t \in \theta_t$, θ_t is a compactum. With dual control, the resource allocations in their turn are assumed to be functions of the divisional parameter estimators $\mathbf{a}_\tau$; $r_\tau=Q_\tau(\mathbf{a}_\tau)$.

The solution of the game is sought in the form (3), with (2) replaced by

$$\hat{w}_t(\mathbf{x}_t, y_t)=\min_{\xi_\tau\in\theta_\tau,\ \tau=\overline{t+1, t+T}}\ \max_{y_\tau\in Y_\tau(\mathbf{p}_\tau)} w_t(\varphi_t, \ldots, \varphi_{t+T}). \tag{13}$$

Comparing (13) with (2), we easily see that in this case $Y_{1\tau}^{1}(\mathbf{p}_\tau)$ is naturally replaced by the set of divisional states $\operatorname{Argmax}_{y_\tau\in Y_\tau(\mathbf{p}_\tau)} w_t$, which maximize the divisional objective function on $Y_\tau(\mathbf{p}_\tau)$. This means that, when predicting its objective function ($\hat{w}(x, y)$), the division assumes that future optimal states will be chosen under unfavorable conditions (minimax strategy of decision making under uncertainty).

A dual mechanism is the mechanism $\Sigma = (I, \pi, Q, f)$ such that the AE chooses the limiting state ($y_t = z_t$) while at the same time extremizing some control performance criterion. As the design of dual mechanisms spans a wide range of problems, we will demonstrate our approach for one example of mechanisms which use the Kalman—Bucy filter.

Suppose that the headquarters sees a divisional constraint model of the form

$$z_t=W(\mathbf{p}_t)=H\mathbf{q}_t+\xi_t,\quad \mathbf{q}_{t+1}=C\mathbf{q}_t+B\mathbf{r}_t,$$

where C, B, H are some matrices with nonnegative elements, ξ_t is a sequence of random variables with zero mean and known covariance. Let $\mathbf{q}_\tau^0$ be the desired value of the vector $\mathbf{q}_\tau$ in period τ. The headquarters minimizes the performance criterion

$$J(r)=M_\xi\left\{\sum_{\tau=0}^{b}(\mathbf{q}_{\tau+1}-\mathbf{q}_{\tau+1}^0)' K_\tau(\mathbf{q}_{\tau+1}-\mathbf{q}_{\tau+1}^0)+\mathbf{r}_\tau' N_\tau \mathbf{r}_\tau\right\}, \tag{14}$$

where K_τ, N_τ are square matrices, [0, b] is the system functioning interval. In order to obtain the estimator $\mathbf{a}_t$ of the vector $\mathbf{q}_t$ from the observations z_t, we apply a linear Kalman—Bucy filter (see, e.g., [9, Sec. 7.3]). The estimator has the form

$$\mathbf{a}_{t+1}=A_t\mathbf{a}_t+G_t z_t. \tag{15}$$

The control procedure takes the form

$$\mathbf{r}_t=-L_t(\mathbf{a}_t-\mathbf{q}_t^0). \tag{16}$$

The matrix elements of A_t, G_t, L_t are nonnegative and are determined by the problem parameters. The procedures (15) and (16) minimize the performance criterion (14) in the absence of active behavior, i.e., for $y_t = z_t$. However, active behavior, according to (3), (13), may induce the division to choose a state y_t, other than limiting (i.e., $y_t < z_t$).

Let us now consider the design of a dual functioning mechanism for a dynamic active system with estimation and control algorithms (15), (16) such that the observed state of the AE is equal to its limiting state ($y_t = z_t$, $t = 0, 1, \ldots$) and at the same time the control performance criterion J(r) (14) is minimized.

6. Design of a Dual Mechanism: the Solution

As the divisional plan vector, we take the estimator $a_t(\mathbf{x}_t=\mathbf{a}_t)$. Then by (15), the planning equations have the form

$$\mathbf{x}_{t+1}=A_t\mathbf{x}_t+G_t y_t. \tag{17}$$

Define the operators $$d\xi=\min_{\Omega_t}\frac{\partial}{\partial\xi},\quad D\xi=\max_{\Omega_t}\frac{\partial}{\partial\xi},$$

$$\Omega_t=\{\nu=\overline{t, t+T};\quad \mathbf{p}_\tau\in P_\tau,\quad \tau=\overline{t, \nu};\quad y_\tau\in Y_\tau(\mathbf{p}_\tau)\}$$

and let

$$U_t = dy\varphi_\nu,\ V_t = Dy\varphi_\nu,\ \underline{\nabla}_x = (dx_1, \ldots, dx_l)',\ \mathbf{F}_t = -\underline{\nabla}'_x\varphi_\nu.$$

We assume that $d\varphi_\tau w_t \geqq D\varphi_\nu w_t \geqq 0$, $t \leqq \tau < \nu \leqq t + T$, i.e., the value of the incentives for the AE decreases over time. Then we have

THEOREM 2. The AFM $\Sigma = (I, \pi, Q, f)$ with learning procedure (15), planning procedure (17), and control procedure (16) is dual if

$$\bar{D}(\Sigma) = U_t - \left[\mathbf{F}_t + \sum_{\tau=2}^{T}\left(\mathbf{F}_t + V_t H \sum_{\nu=1}^{\tau-1} C^{\tau-\nu} B L_t A^{\nu-\tau}\right) A_t^{\tau-1}\right] G_t \geqslant 0. \qquad (18)$$

The conditions (18) set constraints on the incentive procedure (given the divisional structure and the planning and control procedures) which ensure that the AFM is progressive and the AE does not lower its indexes. Note that this gives a solution of the design problem of an optimal adaptive mechanism with stochastic optimal feedback control algorithms for a linear system with known parameters, Gaussian noise, and quadratic performance criterion [9].

Now, since progressivity implies higher divisional incentives for higher divisional efficiency [1-3], we have to consider ways to minimize incentive costs in adaptive mechanisms.

Consider, for example, the class of progressive AFMs with given learning procedure (15), planning procedure (17), and control procedure (16), and nonzero limiting penalties ($\mathbf{F}_t > 0$): $G = \{\Sigma = (I, \pi, Q, f) \mid \bar{D}_t(\Sigma) \geqslant 0, \mathbf{F}_t > 0\}$.

The design problem of the procedure f^*, minimizing the guaranteed incentive costs in the class of AFMs G is posed in the form

$$З_t(\Sigma) = \max_{\xi_t \in \theta_t} f_t(\mathbf{x}_t, z_t) \xrightarrow[\Sigma \in G]{} \min.$$

Here we have used the fact that $y_t = z_t$ by Theorem 2.

The corresponding AFM $\Sigma^* = (I, \pi, Q, f^*)$ is called *minimally* progressive. We have the following

Proposition.† The AFM is minimally progressive if $\bar{D}_t(\Sigma^*) = 0$, $t = 0, 1, \ldots$ In this case $f_t^*(\mathbf{x}_t, y_t)$ is a linear function of its arguments.

The simple proof is omitted. The proposition establishes the minimum limiting incentives (or the maximum limiting penalties) [5] for which a dual mechanism is realizable.

7. Discussion of Results

As we have noted above, Theorems 1 and 2 are constructive and enable us to solve the analysis and design problems for AFMs with standard adaptation and learning procedures as subsystems [6-9]. Two types of results can be identified. The first type is associated with progressivity of standard learning procedures [6-8] for given planning and incentive systems (see Sec. 3). The second type is associated with construction of adaptive planning and incentive procedures which ensure progressivity of mechanisms using optimal estimation, identification, and other algorithms (as we did in Sec. 4 on the basis of Theorem 2). The results may be easily generalized to the case of learning and control by several independent observable state indexes of a vector AE (m > N), i.e., when $y_{kt}^i \leqslant z_{kt}^i = W_{kt}^i(\mathbf{p}_t^i)$, $i = \overline{1, N}$, $k = 1$, N + 1, ..., m. In this case, the total number of progressivity conditions increases from N^2 (see Theorem 1) to mN.

In conclusion note that [3-5] and the present study develop a certain direction in the theory of active systems focusing on analysis and design of adaptive functioning mechanisms in which the current information supplied by the active elements during the control process is utilized in order to alter the parameters of the planning, control, and incentive subsystems of the functioning mechanism so as to attain a certain, usually optimal, state of the active system. The trends and prospects in this area are examined in [10].

Our results on analysis and design of progressive adaptative mechanisms play a central role in development of principles and methods for improvement of applied sectoral planning

†A similar proposition holds with nonzero limiting incentives ($U_t > 0$).

and control systems [3, 10]. Indeed, the sectoral systems are usually based on essentially adaptive procedurers of planning and resource allocation starting "from the attained level." Moreover, these systems are meant to exploit the production capacity reserves, i.e., they are expected to be "progressive."

I would like to acknowledge the useful comments of V. N. Burkov.

APPENDIX

Proof of Theorem 1. Note that if $y_t^* = v_t$, then $\tilde{y}_t^* = z_t$, and by (6) we have $a_t \underset{t}{\longrightarrow} a_t^*$.

Therefore, since the AEs are friendly toward the headquarters, it suffices to prove that $R_t(\Sigma, p_t) \ni v_t$.

Note that, by (7), (8), the planning procedure has the form

$$\mathbf{x}_t = \mathbf{x}_{t-1} - A\gamma_t \nabla_{\mathbf{x}} \Phi(\tilde{\mathbf{y}}_t - \hat{\mathbf{z}}_t), \tag{A.1}$$

which is characteristic of the recursive adaptive planning procedure (1) [5]. Therefore, Lemma 2 [5] holds. Substituting (A.1) in (6), (7), [5], we obtain for t = 0, using the conditions of Theorem 1, $\hat{D}_{j0}^i(\Sigma) = \tilde{D}_{j0}^i(\Sigma) \geq 0$. But then, by Lemma 2 [5], we have

$$D_{j0}^i = \min_{p_\tau \in P_\tau,\ \tau=\overline{0,T}} \ \min_{y_\tau^k \in \mathbf{Y}_\tau^k (p_\tau^k),\ k=\overline{1,N}} \partial w_0^i / \partial y_{j0}^i \geqslant 0. \tag{A.2}$$

Noting that $Y_\tau^k(\mathbf{u}_\tau^k) \subset Y_\tau^k(\mathbf{p}_\tau^k)$ for $\mathbf{u}_\tau^k \leqslant \mathbf{p}_\tau^k$ and using the definitions (3) and (A.2), we can easily show that $\hat{w}_0^1(x_0, y_0)$ is a nondecreasing function of y_{j0}^i, $j = \overline{1, N}$. Hence, noting that $y_{j0}^i \leqslant W_0^j(\mathbf{u}_0^j) \leqslant z_0^j$, $W_0^j(\cdot) \uparrow y_{j0}^k$, $y_{j0}^k \leqslant p_{j0}^k$, $k \neq j$, $1 \leqslant k \leqslant N$, we obtain $\hat{w}_0^i(x_0, v_0) \geqslant \hat{w}_0^i(x_0, y_0)$, $y_0 = (\mathbf{y}_0^1, \ldots, \mathbf{y}_0^N)$ $\forall \mathbf{y}_0^l \in Y_0^l(\mathbf{p}_0^l) \subset Y_0^l(\mathbf{u}_0^l)$.

But then by (3), $R_0(\Sigma, p_0) \ni v_0$. We can similarly show that $R_t(\Sigma, p_t) \ni v_t$, t = 1, 2, ..., whence follows the proposition of the theorem. Q.E.D.

Proof of Theorem 2. It suffices to show that $y_t^* = z_t$, t = 0, 1, ... We will prove this by induction. Let $c_t(y_t, \ldots, y_{t+T}) \equiv w_t(\varphi_t, \ldots, \varphi_{t+T})$ and take some sequence of random noise values $\xi^t = \{\xi_t, \ldots, \xi_{t+T}\}$. For $\tau = t + T$ we have $\partial c_t / \partial y_{t+T} = (d\varphi_{t+T} w_t) U_t \geq 0$ by (18) (recall that under our assumptions all the terms in $\tilde{D}_t(\Sigma)$ except the first, are nonpositive). Therefore $c_t(y_t, \ldots y_{t+T-1}, W_{t+T}(\mathbf{p}_{t+T})) \geqslant c_t(y_t, \ldots, y_{t+T-1}, y_{t+T})$ for any $y_{t+1} \in Y_{t+T}(\mathbf{p}_{t+T}) = [0, W_{t+T}(\mathbf{p}_{t+T})]$. Now assume that for some ν, $t < \nu \leqslant t+T$ $c_t(y_1, y_\nu, W_{\nu+1}(\mathbf{p}_{\nu+1}), \ldots) \geqslant c_t(y_1, \ldots, y_\nu, y_{\nu+1}, W_{\nu+2}(\mathbf{p}_{\nu+2}), \ldots)$, $y_{\nu+1} \leqslant W_{\nu+1}(\mathbf{p}_{\nu+1})$, and show that for $y_\nu \leqslant W_\nu(\mathbf{p}_\nu)$.

$$c_t(y_1, \ldots, W_\nu(\mathbf{p}_\nu), W_{\nu+1}(\mathbf{p}_{\nu+1}), \ldots) \geqslant c_t(y_1, \ldots, y_\nu, W_{\nu+1}(\mathbf{p}_{\nu+1}), \ldots). \tag{A.3}$$

Indeed, we have $\partial c_t / \partial y_\nu \geqslant (d\varphi_\nu w_t)(dy\, \varphi_\nu) - (D\varphi_{\nu+1} w_t) \mathbf{F}_t G_t - \sum_{\tau=\nu+2}^{T} (D\varphi_\tau w_t) \times \left[\mathbf{F}_t + V_t H \sum_{\mu=1}^{\tau-1} c^{\tau-\mu} B L_t A_t^{\mu-\tau} \right] A^{\tau-1} G_t \geqslant (d\varphi_\nu w_t) \tilde{D}_t(\Sigma) \geqslant 0$. The first inequality is obtained by the technique developed in the proof of Lemma 2 [5]. We have used the property of nonpositivity of all the terms (except the first one) in the right-hand side. The second inequality holds since $d\varphi_\nu w_t \geq D\varphi_\tau w_t \geq 0$, $\tau \geq \mu$, while the third is true by assumption. We thus have (A.3) for $t \leq \nu \leq t + T$. Hence, $c_t(W_t(\mathbf{p}_t), \ldots, W_{t+T}(\mathbf{p}_{t+T})) \geqslant c_t(y_t, \ldots, y_{t+T})$ $\forall y_\tau \in Y_\tau(\mathbf{p}_\tau)$, $\tau = \overline{t, t+T}$ for any sequence ξ^t. But then by (3), (13) $R_t(\Sigma, \mathbf{p}_t) \ni W_t(\mathbf{p}_t) = z_t$, and by the friendliness principle $y_t^* = z_t$. Noting that the control (16) also minimizes (14), we conclude that this mechanism is dual. Q.E.D.

LITERATURE CITED

1. V. N. Burkov, Mathematical Fundamentals of the Theory of Active Systems [in Russian], Nauka, Moscow (1977).
2. V. N. Burkov and V. V. Kondrat'ev, Functioning Mechanisms of Organizations [in Russian], Nauka, Moscow (1971).
3. V. N. Burkov, V. V. Kondrat'ev, V. V. Tsyganov, and A. M. Cherkashin, Theory of Active Systems and Improvement of the Economic Mechanism [in Russian], Nauka, Moscow (1984).
4. V. N. Burkov and V. V. Tsyganov, "Adaptive functioning mechanisms of active systems. I," Avtomat. Telemekh., No. 9, 87-94 (1985).
5. V. V. Tsyganov, "Adaptive functioning mechanisms of active systems. II," Avtomat. Telemekh., No. 10, 116-123 (1985).
6. Ya. Z. Tsypkin, Adaptation and Learning in Automatic Systems [in Russian], Nauka, Moscow (1968).

7. Ya. Z. Tsypkin, Fundamentals of the Theory of Learning Systems [in Russian], Nauka, Moscow (1970).
8. Ya. Z. Tsypkin, Fundamentals of the Information Theory of Identification [in Russian], Nauka, Moscow (1984).
9. G. Saridis, Self-Organizing Stochastic Control Systems [Russian translation], Nauka, Moscow (1980).
10. V. N. Burkov and V. V. Tsyganov, "Trends and prospects in the construction of adaptive functioning mechanisms of active systems," Izmereniya, Kontrol', Avtomatizatsiya, No. 4, 53-60 (1985).