Control and Cybernetics
VOL. 4 (1975) No. 2

Application of Interaction Balance Method to Real Process Coordination

by

KRZYSZTOF MALINOWSKI,
ANDRZEJ RUSZCZYŃSKI

Technical University of Warsaw
Institute of Automatic Control

In the paper the coordination of the large real system is discussed. It is assumed that the available model of the system is not precise. A new two-lewel method (Interaction Balance Method with Feedback — IBMF) is proposed.

The coordinability by IBMF is defined and proved in some cases. This is followed by a short discussion of suboptimality of IBMF. Next the coordination strategies are dissussed in a general case and the detailed presentation of them is given for the linear-quadratic static optimization problems. The method is illustrated by a computational example.

1. Introduction

The interaction balance method of coordination (see Section 2 of this paper) finds the optimal controls for the system upon the base of the mathematical models of all subsystems. Such model-optimal controls applied to real system will yield the truly optimal value of the system perfomance generally only in the case, when the mathematical models are the ideal description of reality.

In practice however the models are always less or more precise, but never the ideal image of reality. There are numerous reasons of such situation.

First when mathematical model is being built some simplifying assumptions are alwys made. It is well known too that many of model coefficients are taken from experiments. At last the real process is influenced by various disturbances, which cannot be exactly foreseen. In successive sections one possibility of coupling the interaction balance method using mathematical models with the real system will be analysed. The aim is to find some suboptimal controls. This method was first sugested in [3], [4].

2. Presentation of IBMF

Assume that there are given: the mathematical models of the objects (subsystems)

$$y^i = f^i(c^i, u^i), \quad i = 1, \ldots, N \tag{1}$$

the system structure equations

$$u^i = M^i y = \sum_{j=1}^{N} M_j^i y^j, \quad i=1, \ldots, N \tag{2}$$

and the system perfomance

$$Q(c, u) = \sum_{i=1}^{N} Q^i(c^i, u^i). \tag{3}$$

We assume that $c^i \in C^{i'}$, $u^i \in U^{i'}$, $y^i \in Y^{i'}$, where $C^{i'}$, $U^{i'}$, $Y^{i'}$ are locally convex topological spaces and f^i: $C^{i'} \times U^{i'} \to Y^i$, Q^i: $C^{i'} \times U^{i'} \to R$, M_j^i: $Y^{j'} \to U^{i'}$ (M_j^i — linear continuous). For compactness of notation we denote $c = (c^1, \ldots, c^N)$, similiarly with u and y.

Also for convenience we will write, when needed, the equations (1) and (2) in the compact form

$$y = f(c, u)$$

and

$$u = My$$

The real subsystems are assumed to be described by operator equations which are different from (1) (and of course unknown)

$$y^i = f_*^i(c^i, u^i), \quad i=1, \ldots, N \tag{1'}$$

where

$$f_*^i: C^{i'} \times U^{i'} \to Y^{i'}.$$

We assume that the structure equations in the real system are the same as in the mathematical description (eq. (2)). Controls c^i and interactions u^i (inputs), y^i (outputs) should also satisfy the constraints

$$H_0^i(c^i, u^i, y^i) \in Z^i \subset Z^{i'}, \quad i=1, \ldots, N \tag{4}$$

where H_0^i: $C^{i'} \times U^{i'} \times Y^{i'} \to Z^{i'}$ ($Z^{i'}$ — linear space). The set of (c^i, u^i, y^i) satisfying (4) we denote by CUY^i. We assume that for every control c, where $c^i \in P_{c^{i'}}(CUY^i)$ ($P_{c^{i'}}$ — projection on $C^{i'}$), $i=1, \ldots, N$, there exist unique u, and unique u_* such that

$$u = Mf(c, u) \quad \text{and} \quad u_* = Mf_*(c, u_*).$$

If we substitute equations (1) into (4) we will obtain constraints for (c^i, u^i), i.e.

$$H^i(c^i, u^i) = H_0^i(c^i, u^i, f^i(c^i, u^i)) \in Z^i. \tag{5}$$

The set of (c^i, u^i) satisfying (5) we denote by CU^i. Similiary the set of (c^i, u^i) satisfying

$$H_*^i(c^i, u^i) = H_0^i(c^i, u^i, f_*^i(c^i, u^i)) \in Z^i$$

we denote by CU_*^i. Of course in a general case $CU_*^i \neq CU^i$, but if the constraints H_0^i in (4) do not depend on y^i, then $CU_*^i = CU^i$.

DEFINITION 1. We say that $\hat{c}=(\hat{c}^1, ..., \hat{c}^N)$ is a model-optimal control if $(\hat{c}^i, \hat{u}^i) \in CU^i$, where $\hat{u}^i = Mf(\hat{c}, \hat{u})$, and $Q(\hat{c}, \hat{u}) \leqslant Q(c, u)$ for every $(c, u) \in CU = CU^1 \times ... \times CU^N$ such that $u = Mf(c, u)$.

DEFINITION 2. We say that $\hat{c}_* = (\hat{c}_*^1, ..., \hat{c}_*^N)$ is a real-optimal control if $(\hat{c}_*, \hat{u}_*) \in CU_* = CU^1 \times ... \times CU^N$ where $\hat{u}_* = Mf_*(\hat{c}_*, \hat{u}_*)$ and $Q(\hat{c}_*, \hat{u}_*) \leqslant Q(c, u)$ for every $(c, u) \in CU$ such that $u = Mf_*(c, u)$.

Since the operators f_*^i are unknown the real-optimal control cannot be found unless experimental optimization on the real system is applied. In the case of a complex system with many controls the extremal regulation can hardly be supposed to be successful. On the other hand the model-optimal control can be found through application of a suitable minimization procedure. But if this model-optimal control $\hat{c}$ is applied to the real process, then it will be nonoptimal in general. It may also very easily happen that the pair $(\hat{c}, \check{u})$, where $\check{u} = Mf_*(\hat{c}, \check{u})$, does not belong to CU_* even in the case, when $CU_* = CU$. In ref. [3], [4] it had been suggested to look for another control $\hat{c}$ through the specific coordination method. To explain this idea we must remind that thanks to the special structure of the system model the model-optimal control $\hat{c}$ can be found through the so called interaction balance method (ref. [7]). This is a two-level method: on the lower (infimal) level the infimal decision units solve independently their local problems:

I) given $p = (p^1, ..., p^N)$, $p^i \in U^{i\prime *}$ find

$$\min_{(c^i, u^i)} [Q^i(c^i, u^i) + \langle p^i, u^i \rangle - \sum_{j=1}^{N} \langle p^j, M_j^i f^i(c^i, u^i) \rangle]\ ^{1)}$$

where $(c^i, u^i) \in CU^i$.

The set of solutions of all infimal problems we denote by $\overline{CU}(p)$. Of course $\overline{CU}(p) \subset CU$ for every p. On the upper (supremal) level the coordinator task is:

II M) find $\hat{p} \in U'^*$ such that $\overline{CU}(p)$ is nonempty and

$$u - Mf(c, u) = 0 \tag{6}$$

for every $(c, u) \in P^M(\overline{CU}(p))$, where a selection mapping $P^M: 2^{C' \times U'} \to 2^{C' \times U'}$ (ref. [6]] has the property, that for every $A \subset C' \times U'$: $P^M(A) \subset A$ and $A \neq \emptyset \Rightarrow P^M(A) \neq \emptyset$.

If there exists a solution $\hat{p}$ of the supremal problem, then every c such that $(c, u) \in P^M(\overline{CU}(\hat{p}))$ is model-optimal and we say, that the model of the system is coordinable by the interaction balance method.

Now the Interaction Balance Method with Feedback (IBMF) can be proposed. Let the infimal decision problems be the same as above (i.e. of the form I). But the coordinator task on the supremal level is modified:

II R) find $\tilde{p} \in U'^*$ such that $\overline{CU}(\tilde{p})$ is nonempty and

$$u - u_*(c) = 0 \tag{7}$$

---

1) To distinguish continuous linear functionals we denote the value of $y \in X^*$ (X — linear top. space) on $x \in X$ by $\langle y, x \rangle$.

for every $(c, u) \in P^R(\overline{CU}(\tilde{p}))$ where $u_*(c)=Mf_*(c, u_*(c))$ and a selection mapping $P^R\colon 2^{C'\times U'}\to 2^{C'\times U'}$ has the similar properties as P^M. $u_*(c)$ is the value of interactions in the real system when control c is applied. According to previously made uniqueness assumption $u_*(\cdot)$ is well defined mapping. We assume of course that the values of interactions in the real system can be measured. It is evident that a point $\tilde{c}\big((\tilde{c}, \tilde{u})\subset \subset P^R(\overline{CU}(\tilde{p}))\big)$ is in general case neither model-optimal nor real-optimal. This point may also not belong to the set CU_*. However, in the case, when $CU_*=CU$ (i.e. when the constraints H_0^i do not depend on y^i) we know that $(\tilde{c}, u_*(\tilde{c}))\in CU_*$, which is very important.

The IBMF is applicable in the sense that when the models are ideal, i.e. $f=f_*$, then every $\tilde{c}\big((\tilde{c}, u_*(\tilde{c}))\in P^R(\overline{CU}(\tilde{p}))\big)$ is real-optimal.

So the first question that arises is when there exists the solution $\tilde{p}$ of IBMF supremal problem. We shall try to answer this question in the next section.

The second queation that may be asked concerns the suboptimality of $\tilde{c}$.

This problem will be briefly discussed in section 4.

Finally, one should know the way to look for $\tilde{p}$. This question about possible coordination strategies is considered in sections 5 and 6.

3. Coordinability by IBMF

Now we shall state and prove the basic coordinability theorem.

THEOREM 1. Let the set $CU=CU^1\times\ldots\times CU^N$ beconvex and compact in $C'\times U'$. Let f^i and f_*^i be continuous on CU^i. Denote by $\mathfrak{U}$ the set of such $s\in U'$ that for some $(c, u)\in CU$ we have $Mf_*(c, u) - Mf(c, u)=s$. Suppose that for every $s\in\mathfrak{U}$ the model of the system with equations (1), (2) replaced by

$$u=Mf(c, u)+s \tag{8}$$

is cooridnable by the interaction balance method. Suppose also that model-optimal control $\hat{c}_s$ for the above "s-shifted" system is unique for every $s\in\mathfrak{U}$, so that the mapping $\hat{c}(\cdot)\colon \mathfrak{U}\ni s\to\hat{c}(s)=\hat{c}_s$ is well defined. We assume that this mapping is continuous on $\mathfrak{U}$ as well as the mapping $\hat{u}(\cdot)\colon \mathfrak{U}\to U'$ defined by $\hat{u}(s)=Mf\big(\hat{c}(s), \hat{u}(s)\big)+s$. The set CU is assumed to be independent on s (rel. (5)).

Then there exists the solution $\tilde{p}$ of IBMF (supremal problem IIR) with $P^R=P^M$.

Proof. We shall construct a certain mapping $W\colon CU\to CU$.

For this let us take any point $(c, u)\in CU$ and define $s\in U'$ as

$$s=Mf_*(c, u)-Mf(c, u).$$

Then $s\in\mathfrak{U}$ and according to the assumptions made there exists $\hat{p}_s\in U'^*$ such that

$$\big(\hat{c}(s), \hat{u}(s)\big)=P^M\big(\overline{CU}(\hat{p}_s)\big)$$

and

$$\hat{u}(s)-Mf(\hat{c}(s),\ \hat{u}(s))=s=Mf_{*}(c,u)-Mf(c,u).$$

Let us define the mapping W as $W=W_2\circ W_1$
where

$$W_1\colon CU\ni(c,u)\to s=Mf_{*}(c,u)-Mf(c,u)\in\mathfrak{U}$$

$$W_2\colon \mathfrak{U}\ni s\to(\hat{c}(s),\ \hat{u}(s))\in CU.$$

Since M is a licar continuous operator and f_{*}, f are continuous then W_1 is continuous on CU. W_2 is continuous from the assumption and so W is continuous. Now we can apply the following theorem (ref. [2]):

If X is a compact convex subset of a locally convex topological space X and $F\colon X\to$ $\to X$ is continuous mapping then there exists at least one point $x_0\in X$, such that $F(x_0)=x_0$.

In our case the set CU and the mapping W fulfil the assumptions of the above theorem. So there exists at least one point $(c_0,u_0)\in CU$ such that $W(c_0,u_0)=$ $=(c_0,u_0)$. It means that

$$(c_0,u_0)=(\hat{c}(s_0),\ \hat{u}(s_0)),$$

where

$$(\hat{c}(s_0),\hat{u}(s_0))=P^M(\overline{CU}(\hat{p}_{s_0})),$$

$$s_0=Mf_{*}(c_0,u_0)-Mf(c_0,u_0).$$

If we denote $\hat{p}_{s_0}$ by $\tilde{p}$ then

$$(c_0,u_0)=P^M(\overline{CU}(\tilde{p}))$$

and

$$u_0-Mf(c_0,u_0)=s=Mf_{*}(c_0,u_0)-Mf(c_0,u_0).$$

From this we have

$$u_0=Mf_{*}(c_0,u_0).$$

So $c_0=\tilde{c}$ and the theorem is proved. Q.E.D.

The above theorem is founded on two basic assumptions. One of them is the coordinability of "s-shifted" model (8) by the interaction balance method. The question of coordinability by this method is exhaustively analysed in [5], [6], [7]. In [5] the following theorem has been proved.

THEOREM 2. Assume that $C^{i'}$, $U^{i'}$ are reflexive Banach spaces, the set $CU=$ $=CU^1\times\ldots\times CU^N\subset C'\times U'$ is bounded, closed and convex, operators f^i and f^i_{*} are weakly continuous on CU^i, $i=1,\ldots,N$, Q is weakly lower semicontinuous and bounded from above on CU and

(i) there exists $K_1>0$, $K_1\in R$, such that for every $h\in U'$, $\|h\|\leqslant K_1$ exists $(c,u)\in CU$ such that $u=Mf(c,u)+h$;

(ii) for every $(c, u) \in CU$ we have

$$\|Mf_*(c, u) - Mf(c, u)\| \leqslant K_2, \text{ where } 0 < K_2 < K_1;$$

(iii) for every $p \in U'^*$, $\|p\| \leqslant r$, where

$$r \geqslant \frac{K_0' - K_0''}{K_1 - K_2} \left(K_0' = \sup_{CU} Q(c, u),\ K_0'' = \min_{CU} Q(c, u)\right)$$

the infimal problems I have unique solutions;

(iv) the mappings $\hat{c}(\cdot)$, $\hat{u}(\cdot)$ (see Theorem 1) are well defined and weakly continuous on the set $\mathfrak{U} = \{s \in U' : \|s\| \leqslant K_2\}$.

Then there exists the solution $\tilde{p}$ of IBMF supremal problem IIR with $P^R = I$ (I — identity mapping) and $\|\tilde{p}\| \leqslant r$.

The second vital assumption made in theorem 1 is about existence and continuity of $\hat{c}(\cdot)$ and $\hat{u}(\cdot)$. It is difficult to give general conditions on which this assumption is fulfilled. We shall consider here two particular cases.

Assume that $C^{i'}$, $U^{i'}$, $i = 1, \ldots, N$, are Banach spaces, operators f^i are continuous and affine and the assumptions (i) and (ii) of Theorem 2 are fulfilled.

Now we can formulate two lemmas:

Lemma 1. Suppose that the above assumptions hold, the set CU is convex, closed and bounded, spaces U' and C' are finite dimensional and Q is strictly convex and continuous on CU. Then the mappings $\hat{c}(\cdot)$ and $\hat{u}(\cdot)$ are well defined and continuous on the set $\mathfrak{U} = \{s \in U' : \|s\| \leqslant K_2\}$.

Lemma 2. Let the assumption of Lemma 1 hold with the change that U', C' are any reflexive Banach spaces, Q is strongly[2]) convex, weakly lower semicontinuous and bounded from above on CU. Then the mappings $\hat{c}(\cdot)$ and $\hat{u}(\cdot)$ are well defined and norm-continuous on the set $\mathfrak{U} = \{s \in U' : \|s\| \leqslant K_2\}$.

We shall omit the proofs of these lemmas because they are rather lenghty. The proof of Lemma 1 can be found in [5]. The proof of Lemma 2 is similar and only a bit more complicated.

4. Suboptimality of IBMF

Suppose now, that $CU = CU_*$ and the assumptions of Theorem 2 are fulfilled. Then the solution $\tilde{p}$ of problem IIR exists and yields some control $\tilde{c}$. Of course we can write

$$Q(\tilde{c}, \tilde{u}) + \langle \tilde{p}, \tilde{u} - Mf(\tilde{c}, \tilde{u}) \rangle \leqslant Q(\hat{c}_*, \hat{u}_*) + \langle \tilde{p}, \hat{u}_* - Mf(\hat{c}_*, \hat{u}_*) \rangle$$

where $\tilde{u} = u_*(\tilde{c})$, $\hat{c}_*$ is the real-optimal control and $\hat{u}_* = u_*(\hat{c}_*)$.

---

[2]) Functional g defined on a convex subset X of Banach space X' is strongly convex on X if for every $x_1, x_2 \in X$ and for every $\lambda \in [0, 1]$ we have

$$\lambda g(x_1) + (1-\lambda) g(x_2) \geqslant g(\lambda x_1 + (1-\lambda) x_2) + \alpha\lambda(1-\lambda)\|x_1 - x_2\|, \quad \alpha > 0.$$

Because $CU = CU^*$ then evidently

$$Q(\tilde{c}, \tilde{u}) - Q(\hat{c}_*, \hat{u}_*) \geq 0.$$

From the above inequalities and from Theorem 2 it follows, that

$$0 \leq Q(\tilde{c}, \tilde{u}) - Q(\hat{c}_*, \hat{u}_*) \leq \langle \tilde{p}, (\hat{u}_* - Mf(\hat{c}_*, \hat{u}_*)) + (\tilde{u} - Mf(\tilde{c}, \tilde{u})) \rangle =$$

$$= \langle \tilde{p}, (Mf_*(\hat{c}_*, \hat{u}_*) - Mf(\hat{c}_*, \hat{u}_*)) + (Mf_*(\tilde{c}, \tilde{u}) - Mf(\tilde{c}, \tilde{u})) \rangle \leq 2K_2 \frac{K_0' - K_0''}{K_1 - K_2}. \quad (9)$$

In the above upper bound for the performance loss in the real system all coefficients except K_2 do not depend on the differences between the model and the real system. So it is evident that if $K_2 \to 0$ then performance loss $Q(\tilde{c}, \tilde{u}) - Q(\hat{c}_*, \hat{u}_*) \to 0$. Of course the upper bound (9) is rather excessive. In particular we can easily see from inequalities (9), that in the case, when for every $(c, u) \in CU$: $Mf_*(c, u) - Mf(c, u) = \beta$ (β does not depend on (c, u)), $\|\beta\| \leq K_2$, we hawe

$$Q(\tilde{c}, \tilde{u}) - Q(\hat{c}_*, \hat{u}_*) = 0.$$

It means that in such case $\tilde{c}$ is the real-optimal control generated by IBMF. It should be noted also, that in this situation the model-optimal control $\hat{c}$ is not real-optimal in general. In ref. [5] a comparison between the control $\tilde{c}$ generated by IBMF and model-optimal control $\hat{c}$ was made for a special class of differences between f^i and f_*^i. It was assumed, that

$$f^i(c^i, u^i) = f_I^i(c_s, u^i, \alpha)$$

and

$$f_*^i(c^i, u^i) = f_I^i(c^i, u^i, \alpha + \delta\alpha).$$

Under rather restictive smoothness assumptions the comparison was in favour of IBMF.

We do not present here full considerations because they would requier to much spa e but state only the final conclusions:

1. Linear with respect to $\delta\alpha$ increment of optimal value [3]) of the real system performance is the same as the linear with respect to $\delta\alpha$ increment of the real system performance when this system is coordinated by IBMF.

2. When the model-optimal control $\hat{c}$ is applied to the real system then either the constraints on (c, u) are forced or the linear with respect to $\delta\alpha$ increment of the real system performance is notless than the linear with respect to $\delta\alpha$ increment of this performance when the system is coordinated by IBMF.

5. Coordination strategies, a general discussion

In this section we shall give an indication of the problems connected with the coordination strategies for the coordinator of IBMF.

[3]) It is the difference between the value of Q in the real system when $\delta\alpha \neq 0$ and the value of Q in the real system when $\delta\alpha = 0$ (the model-optimal value).

Assume, that for every $p \in U'^*$ the infimal problems I have unique solutions. Then as in the Theorem 2, we can take the identity mapping I as the selection mapping P^R in IIR). If we denote the infimal solutions by $(\bar{c}(p), \bar{u}(p)) = \overline{CU}(p)$, then the task of the coordinator will be the following: find the solution $\tilde{p}$ of the operator equation

$$F_*(p) = \bar{u}(p) - u_*(\bar{c}(p)) = 0. \tag{10}$$

If can be easily proved that if the set CU is a compact subset of Banach space, operators f^i are continuous and Q is lower semicontinuous, then the mappings $\bar{c}(\cdot)$ and $\bar{u}(\cdot)$ are continuous. If in addition $u_*(\cdot)$ is continuous (which is a quite natural assumption in almost every real-system), then the operator F_* defined above is continuous.

Since a very little more can be said about F_* in a general case without making explicit assumptions on $\bar{c}(\cdot)$ and $\bar{u}(\cdot)$ it is difficult to propose any general strategy for the solution of (10) except possibly, the following one:

$$\min_{p \in U'^*} \|F_*(p)\|. \tag{10a}$$

Here we assume, that U' is a Banach space. There exists the solution of the above minimization problem if the solution of (10) exists. But we can not say more about this minimization problem than about (10) in general. There still remains the question of possible minimization procedures for solving (10a). It should be noted, that local extrema of $\|F_*(\cdot)\|$ can be the cause of serious troubles. The task of finding the solution $\tilde{p}$ of (10) or (10a) may be simplified by starting the coordination strategy from the functional $\hat{p}$ being the solution of IIM.

Assume now, that:

1) $\bar{c}(\cdot)$, $\bar{u}(\cdot)$ are Fréchet differentiable mappings (C', U' — Banach spaces) on some closed convex set $P \subset U'^*$ with nonempty interior and derivatives $\bar{c}'(\cdot)$, $\bar{u}'(\cdot)$ are continuous on P;

2) the operators f and f_* have continuous Fréchet derivatives f'_c, f'_u, f'_{*c}, f'_{*u} with respect to c and u;

3) the operators $[I - Mf'_{*u}(c, u_*(c))]$ and $[I - Mf'_u(c, u_*(c))]$ are invertible and the inverses are continuous on C'; under the above assumptions F_* is continuously Fréchet differentiable operator on P and

$$F'_*(p) = \bar{u}'(p) - [I - Mf'_{*u}(\bar{c}(p), u_*(\bar{c}(p)))]^{-1} \circ Mf'_{*c}(\bar{c}(p), u_*(\bar{c}(p))) \circ \bar{c}'(p).$$

Since we cannot calculate f'_{*u} and f'_{*u} it seems reasonable to approximate $F'_*(p)$ by $A(p)$, where

$$A(p) = \bar{u}'(p) - [I - Mf'_u(\bar{c}(p), u_*(\bar{c}(p)))]^{-1} \circ Mf'_c(c(p), u_*(\bar{c}(p))) \circ \bar{c}'(p)$$

and then propose the coordination strategy:

$$p_{n+1} = p_n - [A(p_n)]^{-1} F_*(p_n), \quad n = 0, 1, 2, \ldots .$$

This strategy is well defined if the operators $A(p_n)$ are invertible.

To show that the proposed strategy can be succesful in some cases we give the following theorem

THEOREM 3. Suppose that the above assumptions 1, 2, 3 hold and let

(i) for every $p_1, p_2 \in P$

$$\|F'_*(p_1)-F'_*(p_2)\| \leqslant \alpha \, \|p_1-p_2\|;$$

(ii) $p_0 \in P$ and $A(p_0)$ is invertible with

$$\|[A(p_0)]^{-1} F_*(p_0)\| \leqslant \eta, \quad \|[A(p_0)]^{-1}\| < \beta;$$

(iii) for every $p \in P$

$$\|\{[I-Mf'_{*u}(\bar{c}(p), u_*(\bar{c}(p)))]^{-1} \circ Mf'_{*c}(\bar{c}(p), u_*(\bar{c}(p)))+ \\ -[I-Mf'_u(\bar{c}(p), u_*(\bar{c}(p)))]^{-1} \circ Mf'_c(\bar{c}(p), u_*(\bar{c}(p)))\} \circ \bar{c}'(p)\| \leqslant \delta,$$

where $\delta \geqslant 0$;

(iv) $3\beta\delta < 1$,

$$\gamma = \frac{\alpha\beta\eta}{(1-3\beta\delta)^2} \leqslant \frac{1}{2}$$

and

$$B(p_0, \xi) \in P, \text{ where } B(p_0, \xi) = \{p \in U'^*: \|p-p_0\| < \xi\}$$

where $\xi = \dfrac{1-\sqrt{1-2\gamma}}{\alpha\beta}(1-3\beta\delta)$.

Then the sequence $\{p_n\}$ generated by the proposed strategy starting from the point p_0 exists in $B(p_0, \xi)$ and converges to the solution $\tilde{p}$ of (10). This theorem is proved immediately by application of theorem 3.2 of ref. [1].

Assumptions of the above theorem can be comparatively easily satisfied for the class of linear-quadratic optimization problems. Since it seems that IBMF can find practical applications mainly in the coordination of steady state process we shall not discuss here the linear-quadratic processes in general. In the next section the coordination strategies for the linear-quadratic static optimization problems are presented in a detailed fashion.

6. Strategies of coordination. The linear-quadratic static case

Consider a linear system given by

$$y = Y_{*c} + Y_{*u}\, u + y^o_*, \qquad u = My \tag{11}$$

where inputs $u \in R_u$, outputs $y \in R_y$, controls $c \in R_c$ and R_u, R_y, R_c are finite dimensional spaces.

Assume that we have the approximate model

$$y = Y_c\, c + Y_u\, u + y^o, \qquad u = My \tag{12}$$

parameters Y_c, Y_u, y_0 of which may differ from the real Y_{*c}, Y_{*u}, y^o_* but the structure M is determined property. We introduce the notation $x = \begin{bmatrix} c \\ u \end{bmatrix}$. The cost function Q is supposed to be of the form $Q=\frac{1}{2}x'Wx+q'x$, where the matrix W is symetric and positive definite (superscript' is used in this section to denote conjugation). Let the set of admissible solutions be the non void polyhedron $X=\{x: H_1 x=$ $=h_1, H_2 x\leqslant h_2\}$. In such situation the costructive method of solving (10) can be developed.

Basing upon the model (12) we can foresee the left-hand sides of (6) and (7) in the form $\bar{u}(p)-My(\bar{c}(p), \bar{u}(p))=G\bar{x}(p)-g$ and $\bar{u}(p)-u(\bar{c}(p))=D\bar{x}(p)-$ $-d$, where G and D are certain linear transformations, $u(\bar{c}(p))$ satisfies the equation $u=MY_c\bar{c}(p)+MY_u u+My^o$ and $\bar{x}(p)$ denotes the solution of the lower level optimization problem

$$\min_{x\in X}[L(x,p)=\tfrac{1}{2}x'Wx+q'x+p(Gx-g)]. \tag{13}$$

In the real system there will be $\bar{u}(p)-u_*(\bar{c}(p))=D_*\bar{x}(p)-d_*$ where $D_*\neq D$, $d_*\neq d$ in general, and $\bar{u}(p)-My_*(\bar{c}(p),\bar{u}(p))=G_*\bar{x}(p)-g_*$.

The equation (10) can be written now in the form

$$D_*\bar{x}(p)-d_*=0. \tag{14}$$

We shall prove that on certain conditions the iterative procedure

$$p^{k+1}=p^k+\varepsilon E(D_*\bar{x}(p^k)-d_*) \tag{15}$$

in which

$$E=(DAG')^{-1} \tag{16}$$

A is hermitian, nonnegative definite such that $\mathcal{N}(A)\cap\mathcal{R}(G')=\{0\}$ ($\mathcal{N}(\cdot)$, $R(\cdot)$ denote the null space and the range of a linear transformation) and ε is a sufficiently small constant, converges to the solution $\bar{p}$ of (14).

We assume that the equations of the interconnected system are linearly independent (i.e. $\mathcal{R}(G)=R_u$).

Before passing to the investigations of the convergence of this algorithm we shall verify that DAG' is nonsingular. Indeed, let for $z\neq 0$ be $DAG'z=0$. Thus $AG'z\in\mathcal{N}(D)$ whence, since the conditions $Dx-d=0$ and $Gx-g=0$ are equivalent, $AG'z\in\mathcal{N}(G)$ $\perp\mathcal{R}(G')$. From this if follows that $AG'z\perp G'z$, wghich contradicts the assumption $\mathcal{N}(A)\cap\mathcal{R}(G')=\{0\}$.

We assume that the real system satisfies the condition (A): exists x^o such that $G_*x^o-g_*=0$, $H_1x^o=h_1$ and $H_2x^o-h_2<0$.

Then the real system is coordinable by the Interaction Balance Method and the function $L_*(x,p)=Q(x)+p'(G_*x-g_*)$ has a saddle point.

THEOREM 4. If $Y_c=Y_{*c}$ and $Y_u=Y_{*u}$ then there exists a constant ε_0 such that for $0<\varepsilon<\varepsilon_0$ the algorithm (15) converges to the solution $\hat{p}$ of (14) and $\bar{x}(\hat{p})$ minimizes Q in the real system (11), no matter how great the difference $y^o-y^o_*$ is.

Proof. Consider the assumptive Lagrange function for the precise model $L_*(x,p)=\frac{1}{2}x'Wx+q'x+p'(Gx-g_*)$. The solution $\bar{x}(p)$ of (13) does not depend on g and is the same as it would be if the models were precise. Thus the gradient of $\bar{L}_*(p)=\min_{x\in X} L_*(x, p)$ is equal to

$$\nabla\bar{L}_*(p)=G\bar{x}(p)-g_*.$$

Let $\tilde{p}$ be any solution of (14). For brevity we introduce the following notations:

$$\bar{x}_d(p)=\bar{x}(p)-\bar{x}(\tilde{p}),\ L^k=\bar{L}_*(p^k),\ s^k=ED\bar{x}_d(p^k).$$

According to the mean value theorem thesre exists $0\leqslant\tau\leqslant 1$ such that

$$L^{k+1}-L^k=\varepsilon\bar{x}_d'(p^k+\tau\varepsilon s^k)\ G'\ ED\bar{x}(p^k)=\varepsilon\bar{x}_d'(p^k)\ G'\ ED\ \bar{x}_d(p^k)+ +\varepsilon\left(\bar{x}_d'(p^k+\tau\varepsilon s^k)-\bar{x}_d'(p^k)\right)G'ED\ \bar{x}_d(p^k). \quad (17)$$

We note that for any $\bar{x}_d$ we can find z such that

$$D\bar{x}_d=DAG'z \quad (18)$$

and for the same $\bar{x}_d$, z there is

$$G\bar{x}_d=GAG'z. \quad (19)$$

Indeed, from (18) it follows that the projections of $AG'z$ and $\bar{x}_d$ on the hyperplane $\mathcal{N}(D)^{\perp}$ are equal. But $\mathcal{N}(D)=\mathcal{N}(G)$ and, consequently (19) holds. Thus the first summand of (17) is equal to

$$\alpha=\varepsilon\bar{x}_d'(p^k)\ G'\ ED\ \bar{x}_d(p^k)=\varepsilon z'\ GAG'(DAG)^{-1}\ DAG'\ z=\varepsilon z'\ GAG'z.$$

Since A is positive semi definite and $\mathcal{N}(A)\cap\mathcal{R}(G')=\{0\}$, there exist a constant ν such that

$$\alpha\geqslant\varepsilon\nu\ \|z\|^2\geqslant\frac{\varepsilon\nu}{\|GAD'\|^2}\|D\bar{x}_d(p^k)\|^2. \quad (20)$$

We shall estimate the second summand of (17)

$$|\beta|=\varepsilon|\left(\bar{x}_d(p^k+\tau\varepsilon s^k)-\bar{x}_d(p^k)\right)'\ G'\ ED\bar{x}_d(p^k)|\leqslant \leqslant\varepsilon\ \|G'\ E\|\cdot\|\ D\bar{x}_d(p^k)\|\cdot\|\bar{x}_d(p^k+\tau\varepsilon s^k)-\bar{x}_d(p^k)\|. \quad (21)$$

It is easy to see that the function $\bar{x}(p)$ satisfies the Lipschitz's inequality. Indeed, $\bar{x}(p)$ is the orthogonal projection of $\hat{\bar{x}}(p)$ on X in the sense of inner product $\langle x, x\rangle_W= =x'\ Wx$, where $\hat{\bar{x}}(p)$ minimizes $L(x, p)$ on $R_c\times R_u$. It is easy to show that $\hat{\bar{x}}(p)= =-W^{-1}(q+G'p)$ and so

$$\|\bar{x}(p_1)-\bar{x}(p_2)\|\leqslant\theta\ \|p_1-p_2\|,$$

where $\theta\leqslant\|W^{-1}\ G'\|_W$.

Therefore from (21) we obtain

$$|\beta|\leqslant\varepsilon\ \|G'\ E\|\cdot\|\ D\bar{x}_d(p^k)\|\cdot\tau\varepsilon\theta\ \|s^k\|\leqslant\varepsilon^2\ \theta\ \|G'\ E\|\cdot\|E\|\cdot\|\ D\bar{x}_d(p^k)\|^2. \quad (22)$$

Form (17), taking into account (20) and (22) we conclude that

$$L^{k+1}-L^k\geqslant\alpha-|\beta|\geqslant\frac{\varepsilon\nu}{\|DAG'\|^2}\|D\bar{x}_d(p^k)\|^2-\varepsilon^2\ \theta\ \|G'\ E\|\cdot\|E\|\cdot\|D\bar{x}_d(p^k)\|^2.$$

Set

$$\varepsilon_0=\frac{\nu}{2\theta\,\|G'E\|\cdot\|E\|\cdot\|DAG'\|^2}$$

For $0<\varepsilon\leqslant\varepsilon_0$ we obtain

$$L^{k+1}-L^k\geqslant\frac{\varepsilon\nu}{2\cdot\|DAG'\|^2}\|D\bar{x}_a(p^k)\|^2.$$

Since $\bar{L}_*(p)$ is bounded from above, it follows from the last inequality that the sequence $\{L^k\}$ converges, and

$$\{\|D\bar{x}_a(p^k)\|\}\to 0 \text{ as } k\to\infty.$$

From (15) we have

$$\|p^{k+l}-p^k\|\leqslant\varepsilon\|E\|\sum_{i=k}^{k+l-1}\|D\bar{x}(p^i)-d_*\|\leqslant\varepsilon\|E\|\sum_{i=k}^{k+l-1}\frac{2\|DAG'\|^2}{\varepsilon\nu}(L^{i+1}-L^i)\leqslant$$

$$\leqslant\frac{2\|E\|\cdot\|DAG'\|^2}{\nu}(L^{k+l}-L^k).$$

Thus $\{p^k\}$ is the Cauchy sequence. Hence $p^k\to\hat{p}$ and $D\bar{x}(\hat{p})-d_*=0$, since $\bar{x}(p)$ is continuous. According to the Theorem 2 we have $\bar{x}(\hat{p})=\hat{x}_*$ what completes the proof.

Having established the convergence of the method for $D=D_*$, our next step is to investigate the behaviour of the new algorithm in the general case with both D and d being approximate. Before proceeding to this problem we shall make some preliminary remarks. We denote by H the matrix whose rows are the rows of H_1 and those rows of H_2 which correspond to the inequalities active in $\bar{x}(\tilde{p})$. If there were no other constraints in the problem it would be

$$\bar{x}(p)=-BG'p+x_0 \tag{23}$$

where

$$B=W^{-1}-W^{-1}H'(HW^{-1}H')^{-1}HW^{-1}. \tag{24}$$

We shall prove that transformation B defined by (24) is nonnegative definite and $\mathcal{N}(B)=\mathcal{R}(H')$. For any $z,\nu\in R_x$ we have $\alpha=(z+\nu)'W^{-1}(z+\nu)\geqslant 0$ and equality holds only for $z+\nu=0$. Set $z=-H(HW^{-1}H')^{-1}HW^{-1}\nu$. We obtain

$$\alpha=z'W^{-1}z+\nu'W^{-1}\nu+z'W^{-1}\nu+\nu'W^{-1}z=$$

$$=\nu'W^{-1}\nu-\nu'W^{-1}H'(HW^{-1}H')^{-1}HW^{-1}\nu=\nu'B\nu.$$

Suppose that $\nu\in\mathcal{R}(H')$. Thus there exists ν_1 such that $\nu=H'\nu_1$ and $B\nu=0$, since $BH'=0$. On the other hand, if $\nu\notin\mathcal{R}(H')$ then $z+\nu\neq 0$ and $\nu'B\nu=\alpha=(z+\nu)'W^{-1}(z+\nu)>0$.

THEOREM 5. Suppose that

(i) The solution $\tilde{p}$ of (14) has the neighbourhood $\mathfrak{U}(\tilde{p})$ such that if $p\in\mathfrak{U}(\tilde{p})$ then in $\bar{x}(p)$ the same constraints are active as in $\bar{x}(\tilde{p})$.

(ii) The rows of H and G form the linear independent set.

Then there exists a positive constant ε_0 which has the following property: for any $0<\varepsilon<\varepsilon_0$ we can choose $\delta_\varepsilon>0$ and the neighbourhood $\mathfrak{U}_\varepsilon(\tilde{p})$ such that if $\|D-D_*\|\leqslant\delta_\varepsilon$ and $p^1\in\mathfrak{U}_\varepsilon(\tilde{p})$ then the algorithm (15) with the step length ε converges to $\tilde{p}$.

Proof. We consider the operator $V(p)=p+\varepsilon E(D\bar{x}(p)-d^*)$. For $p\in\mathfrak{U}(\tilde{p})$ equations (23) and (24) are valid and $V(p)=(I-\varepsilon(DAG')^{-1}(DBG'))p+v$. We first establish that $C=(DAG')^{-1}DBG'$ has positive eigenvalues. Let $p\in R_p$ (dim $R_p=m$) and $x\in R_x$ (dim $R_x=n, n>m$). We introduce in R_x the ortonormal base $x_1, x_2, ..., x_n$ such that for $i\leqslant m$ $x_i\in\mathscr{R}(G')$. We define in the subspace $\mathscr{R}(G')$ of the space R_x the transformation B_1 as follows

$$B_1 x_i=\sum_{j=1}^{m}\langle B_{xi}, x_j\rangle x_j, \qquad i=1,2,..,m.$$

Take any $x\in\mathscr{R}(G')$, $x=\sum_{i=1}^{m}\xi_i x_i$. We have

$$B_1 x=\sum_{i=1}^{m}\xi_i B_1 x_i=\sum_{i=1}^{m}\xi_i\sum_{j=1}^{m}\langle Bx_i, x_j\rangle x_j=\sum_{j=1}^{m}\sum_{i=1}^{m}\xi_i\langle Bx_i, x_j\rangle x_j.$$

For $y=\sum_{j=1}^{m}\eta_j x_j$ we obtain

$$\langle y, B_1 x\rangle=\sum_{j=1}^{m}\eta_j\left(\sum_{i=1}^{m}\xi_i\langle Bx_i, x_j\rangle\right)=\sum_{i=1}^{m}\sum_{j=1}^{m}\xi_i\eta_j\langle Bx_i, x_j\rangle=$$
$$=\langle y, Bx\rangle=\langle By, x\rangle=\langle B_1 y, x\rangle.$$

Therefore the transformation B_1 is hermitian and positive definite, since according to (ii) $(\mathscr{N}(B)=\mathscr{R}(H'))\cap\mathscr{R}(G')=\{0\}$. In a similar fashion we introduce the operator A_1

$$A_1 x_i=\sum_{j=1}^{m}\langle Ax_i, x_j\rangle, \qquad i=1,2,...,m$$

which is in $\mathscr{R}(G')$ hermitian and positive definite. We define nonsingular operators D_1 and G_1 in $\mathscr{R}(G')$ as follows: $D_1x=Dx$, $G_1x=Gx$ if $x\in\mathscr{R}(G')$. Thus $C=$ $=(D_1A_1G_1')^{-1}D_1B_1G_1'=(G_1')^{-1}A_1^{-1}B_1G_1'$. The eigenvalues λ_i of C are equal to the eigenvalues of $A_1^{-1}B_1$ which are positive. We choose $\varepsilon_0=\frac{2}{\max\lambda_i}$. Then for $0<\varepsilon<\varepsilon_0$ all eigenvalues of $I-\varepsilon C$ will lie in the interval $(-1, 1)$. Thus there exist a norm $\|\cdot\|_\varepsilon$ and a constant $\rho_\varepsilon<1$ such that for any $p_1, p_2\in\mathfrak{U}(\tilde{p})$ $\|V(p_1)-V(p_2)\|_\varepsilon\leqslant$ $\leqslant\rho_\varepsilon\|p_1-p_2\|_\varepsilon$. Let us turn to the algorithm (15). It can be written in the form $p^{k+1}=$ $V_*(p^k)$, where $V_*(p)=p+\varepsilon E(D_*\bar{x}(p)-d_*)$. We have for any p_1, p_2

$$\|V_*(p_1)-V_*(p_2)\|_\varepsilon=\|V(p_1)-V(p_2)+\varepsilon E(D-D_*)(\bar{x}(p_1)-\bar{x}(p_2))\|_\varepsilon.$$

As it was shown in the proof of the Theorem 4. $\bar{x}(p)$ satisfies the Lipschitz's inequality. Thus there exists a constant Θ_ε such that $\|\bar{x}(p_1)-\bar{x}(p_2)\|_\varepsilon\leqslant\Theta_\varepsilon\|p_1-p_2\|_\varepsilon$. Therefore

$$\|V_*(p_1)-V_*(p_2)\|_\varepsilon\leqslant(\rho_\varepsilon+\varepsilon\Theta_\varepsilon\|E\|_\varepsilon\|D-D_*\|_\varepsilon)\|p_1-p_2\|_\varepsilon$$

Set $\delta_{1\varepsilon}=\frac{1-\rho_\varepsilon}{\varepsilon\Theta_\varepsilon \|E\|_\varepsilon}$. Then for $\|D-D_*\|_\varepsilon<\delta_1 \varepsilon$ we obtain

$$\|V_*(p_1)-V_*(p_2)\|\varepsilon\leqslant\alpha\|p_1-p_2\|_\varepsilon \tag{25}$$

and

$$\alpha=\rho_\varepsilon+\varepsilon\Theta_\varepsilon\|E\|_\varepsilon\cdot\|D-D_*\|_\varepsilon<1.$$

Since all norms in a finite dimensional space are equivalent, there exists $\delta_\varepsilon>0$ such that $(\|D-D_*\|<\delta_\varepsilon)\Rightarrow(\|D-D_*\|_\varepsilon<\delta_{1\varepsilon})$. Also, it is posible to choose $\beta>0$ such that

$\mathfrak{U}_\varepsilon(\tilde{p})=\{p: \|p-\tilde{p}\|_\varepsilon\leqslant\beta\}\subset\mathfrak{U}(\tilde{p})$. Therefore, if $p^1\in\mathfrak{U}_\varepsilon(\tilde{p})$ then by virtue of the formula (25)

$\|p^k-\tilde{p}\|_\varepsilon\leqslant(\alpha)^{k-1}\cdot\|p^1-\tilde{p}\|_\varepsilon$ and the sequence $\{p^k\}$ converges to $\tilde{p}$. Q.E.D.

The theorem has been proved.

Corollary. If there are no inequality constraints in the problem (i.e. $X=\{x: Hx=h\}$), then the whole space R_p can be taken as $U_\varepsilon(\tilde{p})$ in the above theorem.

7. Some computational results

The following example of the system has been considered.

Subsystem 1:

$$y_1=c_{11}-c_{12}+2u_1+y_1^0,$$
$$Q_1=(u_1-1)^2+(c_{11})^2+(c_{12}-2)^2.$$

Subsystem 2:

$$y_{21}=c_{21}-c_{22}+u_{21}-3u_{22}+y_{21}^0,$$
$$y_{22}=2c_{22}-c_{23}-u_{21}+2u_{22}+y_{22}^0,$$
$$Q_2=2(c_{21}-2)^2+(c_{22})^2+3(c_{23})^2+4(u_{21})^2+(u_{22})^2.$$

Subsystem 3:

$$y_3=z_c c_3+4z_u u_3+y_3^0,$$
$$Q_3=(c_3+1)^2+(u_3-1)^2,$$
$$CU_3=\{(c_3,u_3): c_3+u_3\geqslant-0.5\}.$$

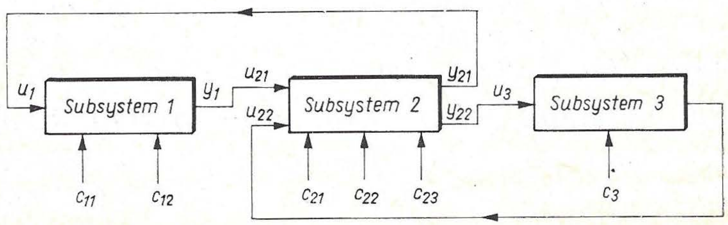

Fig. 1

z_c and z_u are certain variable parameters. The structure of the above system is shown on Fig. 1. Three real systems have been tested:

I. The nondisturbed system with $y_1^o = y_{21}^o = y_{22}^o = y_3^o = 0$. and $z_c = z_u = 1.0$.

II. The shifted system with $y_1^o = -1.0$, $y_{21}^o = 1.0$, $y_{22}^o = -3.0$, $y_3^o = 2.0$. Still $z_c = z_u = 1.0$.

III. The disturbed system with y_i^o the same as in the system II and $z_c = 1.0$, $z_u = 0.5$.

A FORTRAN subroutine based upon the algorithm (15) and the model I has been written and tested extensively. In order to reduce the number of iterations in the interior of X the choice of $A = W^{-1}$ has been made. Indeed, if there were no constrainsts in the problem, $A = W^{-1}$, $\varepsilon = 1.0$ and $D = D_*$ then the algorithm (15) would terminate in 1 iteration, no matter how great the difference $d - d_*$ would be. Three values of the step length ε have been tested for each system: 0.80, 1.00, 1.25. The discoordination norm $\|\bar{u}(p) - u_*(\bar{c}(p))\| = N$ was calculated at each iteration and the algorithm terminated when $N \leqslant 10^{-6}$. The results of the tests indicate, that the advantages of the new method which had been hoped for have been realized. The real-optimal controls for the system II have been found by the new algorithm (basing upon the model I) and no difference between the speed of convergence for

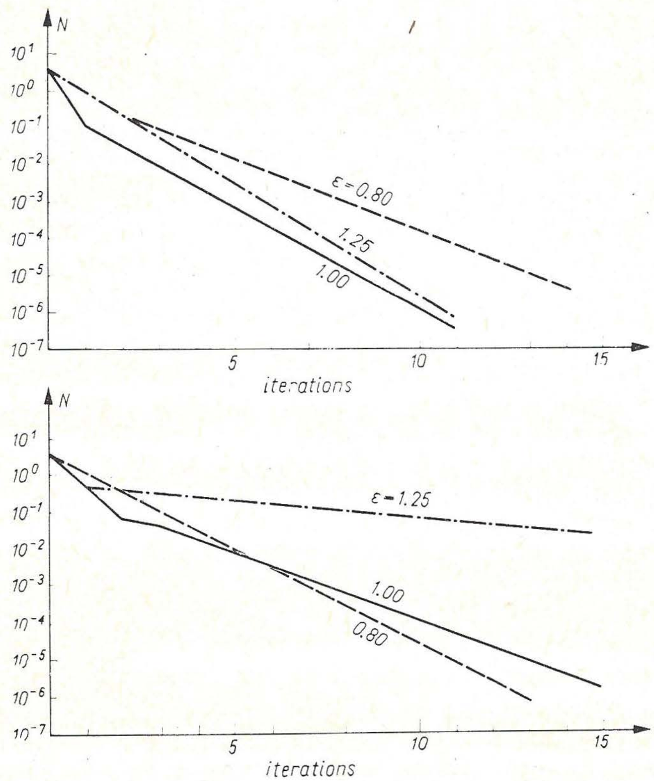

Fig. 2

the system II and for the system I has been observed. The new method proved succesful with the system III too, although considerable difference between D and D_* had been introduced. The process of the discoordination norm minimization for the disturbed systems is illustrated with Fig. 2 — system II above, system III below.

Next the comparison between various methods of controlling the real system III has been drawn. The results are shown in Table 1.

Table 1

Control	Cost in the real system III
Model I — optimal control	129.27918
Model II — optimal control	2.82109*)
IBMF based upon model I	2.65420
Real-optimal control	2.65415
*) Constraints forced in the real system	

Acknowledgment. The authors would like to express their gratitude to Professor W. Findeisen for the inspiration and voluable advices. Thanks are also due to Professor J. Szymanowski and Dr. M. Brdyś for helpfull discussions.

References

1. DENNIS I. E. jr., Toward a unified convergence theory for Newton-like methods, in: Nonlinear functional analysis and applications. Ed. L.B. Rall. New York 1971.
2. EDWARDS R. E., Functional analysis. New York 1965.
3. FINDEISEN W., Multilevel control systems (in Polish). Warszawa 1974.
4. FINDEISEN W., Control and coordination in multilevel systems. Tech. Rep. No. 5/1974. Warsaw 1974 Institute of Automatic Control, Technical University of Warsaw.
5. MALINOWSKI K., Multilevel methods of mathematical programming and complex system optimization (in Polish). Ph. D. Diss. Warsaw 1974 Institute of Automatic Control, Technical University of Warsaw.
6. MALINOWSKI K., Properties of two balance methods of coordination. To appear in Bull. Pol. Acad. Sci.
7. MESAROVIC M., MACKO D., TAKAHARA Y., Theory of hierarchical, multilevel systems. New York 1970.
8. RUSZCZYŃSKI A., On coordination of real processes by the interaction balance method (in Polish). M. Sc. Thesis. Warsaw 1974 Institute of Automatic Control, Technical University of Warsaw.

Zastosowanie metody równoważenia interakcji do koordynowania procesów rzeczywistych

Omówiono koordynację wielkich systemów rzeczywistych przy założeniu, że istniejące modele matematyczne systemów nie są dokładne. Zaproponowano nową dwupoziomową metodę koordynowania: metodę równoważenia interakcji ze sprzężeniem zwrotnym (IBMF). Określono warunki koordynowalności w sensie metody IBMF i dowiedziono ich spełnienia dla kilku konkretnych przypadków. Następnie omówiono suboptymalne własności zaproponowanej metody. Rozważono również strategie koordynacji dla przypadku ogólnego a dla zadań liniowo-kwadratowych optymalizacji statycznej przedstawiono ich dokładną postać. Metoda została zilustrowana przykładem numerycznym.

Применение метода уравновешивания взаимодействий для координации реальных процессов

Рассматривается координация реальных больших систем при условии, что существующие математические модели яаляются не точными. Предложено новый двухуровневой метод координации — метод уровновешивания зваимодействих с обратной связью (IBMF).

Определены условия для координации в смысле метода IBMF и доказано их исполнение для нескольких конкретных случаев. Затем рассмотрены субоптимальные свойства предложенного метода. Рассмотрены также стратегии координации для общего случая, а для линейно-квадратичных задач статической оптимизации предложена их точная форма.

Метод иллюстрируется численными примерами.