IFAC

WARSZAWA 1969

INTERNATIONAL FEDERATION OF AUTOMATIC CONTROL

Optimization of Large-Scale Systems

Fourth Congress of the International Federation of Automatic Control
Warszawa 16–21 June 1969

SURVEY PAPER

16

Organized by
Naczelna Organizacja Techniczna w Polsce

OPTIMIZATION OF LARGE-SCALE SYSTEMS

by

Roman Kulikowski

1. Introduction

Problems of optimum control of complex industrial systems, such as e.g. an integrated power system, steel mill, chemical factory, etc., create a new branch of optimization theory.

In this theory two basic trends may be distinguished. The first one which can be referred to as analytic, aims at decomposing the original large-scale problem into a number of smaller and simpler sub-problems which can be solved effectively by the existing mathematical tools. The second trend, which can be referred to as synthetic, starts with simple controlled sub-processes having known performance properties, and by a process called aggregation creates a complex system with the desirable property. In other words: the first trend stems from the desire for better knowledge of the complex nature of large-scale problems by breaking them down to simple sub-problems, the second trend tries to synthesize the large scale project from the well known sub-systems or operations.

It should be noted that the intuitive idea of decomposition as well as aggregation is not new and it is frequently used in the design of complex industrial systems. However, for the purpose of optimization of large-scale systems formal notions of the decomposition and aggregation is needed. The papers by Dantzig and Wolfe [11, 12] constitute an important contribution in this respect. These authors formulate the decomposition problem for the complex linear-programming problem and give an effective algorithm for the solution of complex problem in terms of solution of sub-problems. A similar method was also applied to nonlinear programming problems [36].

In control theory and its applications, the decomposition methods for dynamic processes constitute the most important and interesting problem which can be formulated in the following way. Let the functionals $F_i(x_i)$, $i = 1, \ldots, n$, and the operators $G_i(x_i)$, $H_i(x_i)$, $i = 1, \ldots, n$, be given, where $x_i \in X_i$, $G_i : X_i \to Y_i$, $H_i : X_i \to Z$, $i = 1, \ldots, n$, and X_i, Y_i, Z be generally speaking, Banach spaces

The local optimization problems consist in finding elements $x_i = \bar{x}_i \in X_i$ such that the functionals $F_i(x_i)$ attain their conditional maximum subject to the inequality constraints $G_i(x_i) \geqslant 0$, viz.,

$$F_i(\bar{x}_i) = \max_{G_i(x_i) \geqslant 0} F_i(x_i), \quad i = 1, \ldots, n \tag{1}$$

The global optimization problem consist in finding such elements $x_i = \bar{\bar{x}}_i \in X_i$, for which the functional

$$F(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} F_i(x_i) \tag{2}$$

attains its conditional maximum subject to the inequality constraints

$$G_i(x_i) \geqslant 0, \quad i = 1, \ldots, n \tag{3}$$

$$\sum_{i=1}^{n} H_i(x_i) \geqslant h \tag{4}$$

where h is a given element of Z.

Relation (4) represents the interactions between the n individual sub-systems specified by F_i and G_i. If (4) is absent or inactive (i.e. if $\sum_{1}^{n} H_i(x_i) > h$ for all $x_i \in X_i$, $i = 1, \ldots, n$) the global problem decomposes into n local problems.

Let us assume that the solutions for the local and global problems exist and that it is much easier to derive the local than the global optimum solution. Then, assuming that the local solutions are know, we search for the global solution in terms of local ones. In the sections to follow several methods of this type shall be considered.

A typical optimization problem of the type considered is the optimization of integrated power system [15] which consists of m thermal and n hydro stations generating electrical power for common load.

Assuming that the instantaneous cost of generating electric power P_{ci} in the i-th thermal station is $F_i(P_{ci})$, the resulting cost in the interval $[0,T]$ becomes

$$C = \sum_{i=1}^{m} \int_{0}^{T} F_{ci}\left[P_{ci}\right] dt \qquad (5)$$

The hydrostations are characterized by functions $P_{Hi}\left[q_i\right]$, where P_{Hi} is the power generated by the i-th hydrostation and q_i is the rate of water-flow through the hydro-turbine.

Denoting by $r(\tau)$ the rate of water inflow and by $v(t)$ the instantaneous water storage in the reservoir we get the relation

$$v(t) = V_0 - V_{min} + \int_{0}^{t} r(\tau) d\tau$$

where: V_0 - amount of water at $t = 0$, V_{min} - minimum admissible amount of water in the reservoir. If now we assume that all the hydrostations are being supplied by the same reservoir we obtain

$$\sum_{i=1}^{n} \int_{0}^{t} q_i(\tau) d\tau \leqslant v(t) \qquad (6)$$

Denoting the power demand by $P(t)$ and neglecting transmission losses we get

$$\sum_{i=1}^{m} P_{Hi}\left[q_i\right] + \sum_{i=1}^{n} P_{ci} = P(t) \qquad (7)$$

The problem consists in finding such strategies $q(t) = \bar{q}_i(t)$, $t \in [0,T]$, which minimize (5) subject to the set of local constraints

$$\begin{aligned} P_{i\,min} \leqslant P_{ci}(t) \leqslant P_{i\,max} \\ 0 \leqslant q_i(t) \leqslant Q_i, \quad i = 1, \ldots, n \end{aligned} \qquad (8)$$

and global constraints (6), (7).

A similar optimization problem exists in the case of integrated utility gas system , certain dynamic inventory problems, etc.

The optimum strategy for a single system where $m = n = 1$ can be derived relatively easily [24]. However, when the system consists of many stations (m , $n > 1$) and the interactions (6) (7) occur the effective computation of optimum strategies poses a difficult optimization problem.

It should be observed that simpler optimization problems also exist when there are no local constraints or when equality signs appear in (3), (4).

A class of control problems called autonomous control is also known in which the interactions appearing between the coordinates of a dynamic system can be compensated in the controller. However, in systems of this type the processes controled are, generally speaking, neither optimum nor sometimes even realizable.

The purpose of the present paper is to give a short review of optimization methods based on the decomposition or aggregation of large-scale systems, which can be implemented in the form of a two- or multi-level structure including local and higher level controllers. The problem of optimum organization of the multilevel structure will be also considered.

The limited space, however, will not allow to present all the methods known, and greatest stress will be laid on the optimization of dynamic systems. The studies made in Poland will also be emphasized.

2. Two-level Control of Linear Systems with Interactions [20]

Let us consider linear system shown in Fig. 1 with n controlled inputs $u_1, u_2, \ldots, u_n$, and n output terminals $y_1, \ldots, y_n$.

The input-output relations are specified by the formula

$$y_i = \sum_{j=1}^{n} A_{ij}(u_j), \quad i = 1, \ldots, n \tag{9}$$

where A_{ij} are linear continuous operators in Hilbert space H. The performance measure is assumed to be

$$F(u) = \sum_{i=1}^{n} \left\{ \|u_i\|^2 + \lambda_i \|y_{pi} - \sum_{j=1}^{n} A_{ij}(u_j)\|^2 \right\} \tag{10}$$

where: λ_i - given positive numbers, y_{pi} - given elements of Hilbert space.

The space of square integrable functions $L^2[0,T]$ and the integral operator of Volterra type (11) are concrete examples of H and A_{ij}, respectively,

$$A_{ij}(u_j) = \int_0^t k_{ij}(t-\tau)u_j(\tau)d\tau \tag{11}$$

where $u_j(\tau)$ and $k_{ij}(t-\tau)$ are square-integrable for $t,\tau \in$ $\in [0,T]$.

The norm

$$\|u_j\|^2 = \int_0^T |u_j(\tau)|^2 d\tau , \; j = 1, \ldots, n$$

represents here the cost of control-energy whereas

$$\|y_{pi} - \sum_{j=1}^{n} A_{ij}(u_j)\|^2$$

represents the square-error between the outputs desired (y_{pi}) and actual (y_i) of the system .

Using variational methods it is possible to derive the optimum controls $u_i = \bar{u}_i$, $i = 1, \ldots, n$, which minimize the functional (10), which become [20] :

$$u_i = -\sum_{k=1}^{n} \lambda_k A^*_{ki} \left[\sum_{j=1}^{n} A_{kj}(u_j) \right] + \sum_{k=1}^{n} \lambda_k A^*_{ki}(y_{pk}) \tag{12}$$

$$i = 1, \ldots, n$$

where A^*_{ki} - linear operator , adjoint to A_{ki} .When A_{ki} has the form like in (11), the adjoint operator becomes

$$A^*_{ki}(u_i) = \int_t^T k_{ki}(\tau - t)u_i(\tau)d\tau \qquad (13)$$

For physically realizable operators A^*_{ki} can be realized by analogue devices in an approximate manner only. By analysing the form of the optimum solution (12) it is possible to observe that the analogue synthesis of the optimum controller assumes an "adjoint" form shown for $n = 2$ in Fig. 2.

That property can be also expressed in the form of the following principle of reflected images.

The optimum structure of the controller, minimizing the measure of the quadratic performance (10), should be a reflected image of the system structure controlled.

Using this principle it is possible to synthesize the structure of an optimum controller for complicated multidimensional processes in a simple manner.

Using the terminology already introduced the controllers specified by the operators A^*_{ii}, $i = 1, \ldots, n$, can be referred to as local (or 1st-level) controllers, and the controller which realizes the operators A^*_{ij}, $i \neq j$, $i, j = 1, 2, \ldots, n$, can be referred to as coordinating (or 2nd-level) controller.

It should be stressed here that the two-level control problems play an important role not only when planning and designing complex controlled systems, but are also when a system is being reconstructed and supplied with new controlling devices. In the latter situation it is sometimes convenient to apply simple 2nd-level controllers only instead of replacing all controllers by a multidimensional and expensive global controller. The decentralized system operates with relatively simple e.g. onedimensional controllers only.

Now we can consider the problem of implementing the optimum solutions (12) by means of digital controllers. In that case it will be convenient to write Eq. (12) in a vector form

$$\underline{u} = \underline{A}(\underline{u}) + \underline{y} \qquad (14)$$

where

$$\underline{u} = (u_1, u_2, \ldots, u_n)$$

and the components of $\underline{A}$ and $\underline{y}$ are

$$-\sum_{k=1}^{n} \lambda_k A_{ki}^* \left[\sum_{j=1}^{n} A_{kj}(u_j)\right] \quad \text{and} \quad \sum_{k=1}^{n} \lambda_k A_{ki}^*(y_{pk})$$

$$i = 1, \dots, n$$

respectively. $\underline{A}$ is a linear matrix selfadjoint operator. We assume that A is a contracting operator, i. e. for arbitrary elements $\underline{u}_1, \underline{u}_2 \in H$ we get

$$\|\underline{A}(\underline{u}_1) - \underline{A}(\underline{u}_2)\| \leqslant \beta \|\underline{u}_1 - \underline{u}_2\| \tag{15}$$

where $\beta < 1$.

Then the optimum solution of (14) $u = \bar{u}$ can be approximated by iteration

$$\underline{u}^{(k+1)} = \underline{A}(\underline{u}^{(k)}) + \underline{y}\ , \quad k = 0, 1, \dots \tag{16}$$

where $\underline{u}^{(o)} \in H$ is an arbitrary element, and $\lim_{k \to \infty} \underline{u}^{(k)} = \bar{\underline{u}}$.

When n = 2, the iteration (16) can be written in the following form

$$u_1^{(k+1)} = -\lambda_1 A_{11}^*\left[A_{11}(u_1^{(k)}) - y_{p1}\right] + \delta_{1k} \tag{17}$$

$$u_2^{(k+1)} = -\lambda_2 A_{22}^*\left[A_{22}(u_2^{(k)}) - y_{p2}\right] + \delta_{2k} \tag{18}$$

where

$$\delta_{1k} = -\lambda_2 A_{21}^*\left[A_{21}(u_1^{(k)}) - y_{p2}\right] + \\ -\left[\lambda_1 A_{11}^* A_{12} + \lambda_2 A_{21}^* A_{22}\right](u_2^{(k)})$$

$$\delta_{2k} = -\lambda_1 A_{12}^*\left[A_{12}(u_2^{(k)}) - y_{p1}\right] + \\ -\left[\lambda_2 A_{22}^* A_{21} + \lambda_1 A_{12}^* A_{11}\right](u_1^{(k)})$$

It can be observed that the algorithms (17), (18) and the one dimensional algorithms (without interactions) differ by additive terms δ_{1m}, δ_{2m} only. Then the organization of the computations can be changed to that shown in Fig. 3, where the 1st-level controllers C_1, C_2 compute $u_1^{(k+1)}$, $u_2^{(k+1)}$ by formulae 17 , 18 ,whereas the 2nd-level controller C computes δ_{1m}, δ_{2m} using the results $u_1^{(k)}$, $u_2^{(k)}$ obtained from C_1,C_2. The optimization process requires then an exchange of information between the 1st- and 2nd-level controllers.

The optimization process of this type can be easily extended to the multidimensional case ($n > 2$). The main advantage of the two-level optimization is that one can use local sub-programmes of the type (17), (18) which are only slightly modified by the additive terms δ_{1m}, δ_{2m} supplied by the coordi - nating (supervisory) controller C .

Is should be noted that the idea of using two-level implementation of iterational solutions of optimization problems for linear and nonlinear processes was used by many authors.

In the case considered so far, the interactions take place among the inputs and outputs of the optimized system. Another kind of interactions is obtained when the controllers are supplied from the same source of energy, and consequently

$$\sum_{i=1}^{n} \|u_i\|^2 \leqslant U \qquad (19)$$

where U is a given number. In this case we shall also neglect the input-output interactions seting $A_{ij} = 0$, $i \neq j$, $i, j = 1, 2, \dots, n$, and denoting A_{ii} by A_i, $i = 1,2,\dots,n$, The performance measure shall be given the following form

$$F(u) = \sum_{i=1}^{n} \|y_{pi} - A_i(u_i)\|^2 \qquad (20)$$

Using variational methods it is possible to show[20],that the optimum $u = \bar{u}$ which minimizes the functional (20) subject to the condition (19) can be derived from the equations

$$\bar{u}_i = \frac{1}{\lambda_i} A_i^*[y_{pi} - A_i(\bar{u}_i)], \; i = 1, 2, \dots, n \qquad (21)$$

yielding

$$\bar{u}_i = R_{i\lambda}\left[A_i^*(y_{pi})\right], \; i = 1, \ldots, n \tag{22}$$

where $R_{i\lambda} = (\lambda I + A_i^* A_i)^{-1}$ is called the resolvent operator, and the parameter λ can be determined from the equation

$$\sum_{i=1}^{n} \left\| R_{i\lambda}\left[A_i^*(y_{pi})\right]\right\|^2 = U \tag{23}$$

Using this procedure it is possible to synthesize the two--level optimum control-system. A system for the case of two sub-processe, n = 2, is given in Fig. 4 as an example.The optimum control strategies (21) can be realized in the form of feedback systems S_1, S_2, where A_i, i = 1, 2,represents plant operators, A_i^* - the correcting systems,and A_{mi} - amplifiers with an amplification factor $\mu_i = 1/\lambda_i$. The systems described by $\mu_i A_i^*$ can be referred to as 1st-level controllers.

The 2nd-level controller observes y_{pi} and finds the value of λ which is the solution of Eq. (23). This value of λ is transmitted to the first-level controllers, where it readjusts the amplification factors of the amplifiers A_{mi} in such a way that $\mu_i = 1/\lambda$.. If the inputs y_{pi} do not vary in a certain number of consecutive optimization intervals, it is possible to construct a simpler analogue 2nd-level controller, which observes the allowed (U) and actual $\left(\sum_{i=1}^{n} \|u_i\|^2\right)$ energy consumptions, and by readjusting the amplification factor $\mu = 1/\lambda$ in the 2nd-level controllers tries to satisfy Eq. (23). This method may be particularly advantageous when the characteristics of the sub-systems are not completely known to the 2nd-level controller, and an adaptive optimization approach is needed.

Since λ may be regarded to be a Lagrange multiplier for Lagrangian

$$\Phi(\underline{u}) = \lambda\left[\sum_{i=1}^{n} \|u_i\|^2 - U\right] + \sum_{i=1}^{n} \|y_{pi} - A_i(u_i)\|^2$$

it can be also regarded to be the price asigned to the control energy. This price is derived by the 2nd-level controller and communicated to the 1st-level ones.

It should be observed that the goal of the 2nd-level controller is to find such a price strategy which makes the loss of the unemployed resources, $U - \sum_{i=1}^{n} \| \bar{u}_i \|^2$, equal zero;while the goal of the 1st-level controllers is to minimize the respective performance factors for every value of λ dictated by the 2nd-level [22] .

It should be also noted that the decomposition methods and optimum two-level control is possible also for other performance criteria, such as minimum time, minimum magnitude, etc.

3. Multistage Optimization

A characteristic feature of large-scale optimization problems is the large number of decision variables which should be determined in such a way that the given performance factors reach their minimum or maksimum value.In many cases it is convenient to realize the optimization in the form of a multistage process, when at each stage the optimization is performed with respect to certain variables, whereas the remaining variables are kept constant.

The main problem connected with this procedure, may be formulated as follows: what are the conditions for the multistage optimization process being optimum overall?In the case of continuous performance functions $f(\underline{x}, \underline{y})$, $\underline{x} \in X$, $\underline{y} \in Y$,where X, Y are compact sets in vector space E^n, E^m , respectively, it is possible to show that

$$\max_{\underline{x}} \left[\max_{\underline{y}} f(\underline{x}, \underline{y}) \right] = \max_{\underline{y}} \left[\max_{\underline{x}} f(\underline{x}, \underline{y}) \right]$$

and

$$\min_{\underline{x}} \left[\min_{\underline{y}} f(\underline{x}, \underline{y}) \right] = \min_{\underline{y}} \left[\min_{\underline{x}} f(\underline{x}, \underline{y}) \right]$$

However

$$\min_{\underline{y}} \left[\max_{\underline{x}} f(\underline{x}, \underline{y}) \right] \geqslant \max_{\underline{x}} \left[\min_{\underline{y}} f(\underline{x}, \underline{y}) \right] \qquad (24)$$

According to the well known minimax theorem we have the equality sign in (24) if X, Y are convex and $f(\underline{x}, \underline{y})$ is continuous and convex in $\underline{y}$ for each $\underline{x}$, and concave in $\underline{x}$ for each $\underline{y}$.

Let us consider the system consisting of N controlled sub-systems with variables x_i, y_i, $i = 1, 2, \ldots, N$, as a typical example of multistage optimization [6]. The performance of each sub-system can be evaluated using the functions $F_i(x_i, y_i)$, $i = 1, 2, \ldots, N$.

The sub-system constraints take the form

$$R_i(x_i, y_i) \leqslant 0, \quad i = 1, 2, \ldots, N \tag{25}$$

It is required to find such values $x_i = \bar{x}_i$, $y_i = \bar{y}_i$, $i = 1, 2, \ldots, N$, which minimize the global performance

$$F(\underline{x}, \underline{y}) = \sum_{i=1}^{N} F_i(x_i, y_i) \tag{26}$$

subject to the global constraints

$$\sum_{j=1}^{N} a_{ij}x_j + \alpha_i \leqslant 0, \quad i = 1, 2, \ldots, N \tag{27}$$

where a_{ij} and α_i are given real numbers.

Let us assume that F_i, R_i are convex functions of real variables x_i, y_i. Then, the Lagrangian

$$\Phi(\underline{x}, \underline{y}, \underline{\lambda}, \underline{\mu}) = \sum_{i=1}^{N}\Big[F_i(x_i, y_i) + \lambda_i R_i(x_i, y_i)\Big] + \sum_{i=1}^{n} \mu_i \left[\sum_{j=1}^{n} a_{ij}x_j - \alpha_i\right] \tag{28}$$

has the saddle-point $(\bar{\underline{x}}, \bar{\underline{y}}, \bar{\underline{\lambda}}, \bar{\underline{\mu}})$ which represents the global solution, viz.,

$$\Phi(\bar{\underline{x}}, \bar{\underline{y}}, \bar{\underline{\lambda}}, \bar{\underline{\mu}}) = \max_{\underline{\lambda}, \underline{\mu}} \left\{ \min_{\bar{x}, \bar{y}} \Phi(\underline{x}, \underline{y}, \underline{\lambda}, \underline{\mu}) \right\} \tag{29}$$

By interchanging the order of summation in Eq. (29) we get

$$\Phi(\underline{\bar{x}}, \underline{\bar{y}}, \underline{\bar{\lambda}}, \underline{\bar{u}}) = \max_{\underline{\mu}} \left\{ \max_{\underline{\lambda}} \left[\min_{\underline{x},\underline{y}} \sum_{i=1}^{N} (F_i(x_i,y_i) + \lambda_i R_i(x_i,y_i) + x_i \sum_{j=1}^{n} a_{ij}\mu_j - \mu_i \alpha_i) \right] \right\} \tag{30}$$

As it may be seen, the sub-system variables in Eq. (30) are grouped in such a manner that we have the sum of N independent functions depending only on x_i, y_i, $i = 1,2,\dots,N$. The optimization problem may than be performed in following stages:

1. Local problems (1st-level): minimize functions

$$f_i(x_i, y_i) = F_i(x_i, y_i) + x_i \sum_{j=1}^{n} a_{ij}\,\mu_j - \mu_i\,\alpha_i \tag{31}$$

$$i = 1, 2, \dots, N$$

subject to the constraints

$$R_i(x_i, x_i) \leqslant 0, \quad i = 1, 2, \dots, N \tag{32}$$

and fixed numbers $\mu_i \geqslant 0$, $i = 1, \dots, n$.

When it is possible to solve these problems and find $x_i = \bar{x}_i$, $y_i = \bar{y}_i$, as explicit functions of $\underline{\mu}$, i.e. $\bar{x}_i(\underline{\mu})$ and $\bar{y}_i(\underline{\mu})$, it is also possible to derive the functions

$$\varphi_i(\underline{\mu}) = f_i\left[\bar{x}_i(\underline{\mu})\right], \quad i = 1, 2, \dots, N \tag{33}$$

2. Coordination problem (2nd-level):find the values $\mu_i = \bar{\mu}_i$, $i = 1, \dots, n$, such that the function

$$\sum_{i=1}^{N} \varphi_i(\underline{\mu}) \tag{34}$$

reaches its maximum value. It is possible then to derive also $\bar{x}_i(\underline{\bar{\mu}})$, $\bar{y}_i(\underline{\bar{\mu}})$, $i = 1, \dots, N$, which represent the solution of the global problem

Another popular example of multistage optimization is a water distribution system. Let us consider for instance the system shown in Fig. 5, which consists of two reservoirs Z_1, Z_2 containing V_1, V_2 $[m^3]$ of water, respectively. Besides, the quantity q_{12} $[m^3]$ of water may be delivered from Z_1 to Z_2. The water volume $V_1 - q_{12}$, contained in Z_1, should be distributed among n receivers, demanding $a_1, \ldots, a_n$ $[m^3]$ of water, respectively. Since

$$\sum_{i=1}^{n} a_i \geqslant V_1 - q_{12} \tag{35}$$

the receivers obtain $x_i \leqslant a_i$ cubic meters of water only and they suffer the losses estimated by

$$S_1(\underline{x}) = \sum_{i=1}^{n} (a_i - x_i)^2 \tag{36}$$

In a similar way for the reservoir Z_2 we obtain

$$\sum_{i=1}^{n} b_i \geqslant V_2 + q_{12} \tag{37}$$

where $b_1, b_2, \ldots, b_m$ are water demands of the receivers supplied by Z_2.

The losses connected with Z_2 are

$$S_2(\underline{y}) = \sum_{i=1}^{m} (b_i - y_i)^2$$

where $y_1, \ldots, y_m$ are the quantities of water supplied to the receivers from Z_2.

The problem consists in finding such values of $x_i = \bar{x}_i$, $q_{12} = \bar{q}_{12}$, $y_i = \bar{y}_i$, which minimize the global losses

$$S(\underline{x}, \underline{y}) = S_1(\underline{x}) + S_2(\underline{y}) \tag{38}$$

subject to constraints (35), (37), and $x_i \geqslant 0$, $y_i \geqslant 0$, $q_{12} \geqslant 0$.

Instead of solving the global problem it is possible to fix

q_{12}, and find at the first stage; the optimum water distribution $\bar{x}_i$, $\bar{y}_j$, i = 1, 2, ..., n, j = 1, ..., m, as functions of q_{12}. It is then possible to compute the functions

$$\varphi_1(q_{12}) = S_1\left[\underline{x}(q_{12})\right] \quad , \quad \varphi_2(q_{12}) = S_2\left[\underline{y}(q_{12})\right] \tag{39}$$

and at the 2nd stage to determin the optimum value $q_{12} = \bar{q}_{12}$. For further details on this procedure and several extenssions cf. Ref. [13, 30].

It should be also observed that many examples of the multi-stage optimization procedure may be found in the Bellman dynamic programming.

When this procedure is used, the problem that poses greatest difficulties is the derivation of the resulting function (such as φ_i in (33) and (39)).

Many examples of problems are known when this function cannot be derived in an explicit manner. However, methods exist which help to overcome this drawback. We shall describe such a method [37] recurring to the formulation of local and global problems given in section 1. We assume that F_i, G_i, H_i are concave differentiable functions of real variables. It is also assumed that a solution of m auxiliary 1st-level problems exists i.e. it is possible to derive the solution $x_i = \tilde{x}_i$ of

$$\max F_i(x_i) \tag{40}$$

subject to

$$G_i(x_i) \geqslant 0 \tag{41}$$
$$H_i(x_i) = y_i, \quad i = 1, 2, \ldots, m$$

where y_i - given real numbers.

This solution is a function of y_i, i.e. $\tilde{x}_i(y_i)$.

Let us define

$$f_i(y_i) = F_i\left[\tilde{x}_i(y_i)\right], \; i = 1, \ldots, m$$

By the 2nd-level optimization problem we shall understand the problem of finding $y = \tilde{y}_i$, i = 1, ..., m, such that the function

$$\sum_{i=1}^{m} f_i(y_i) \tag{42}$$

attains the maximum value subject to the constraint

$$\sum_{i=1}^{n} y_i \geqslant h \qquad (43)$$

It is obvious that (42),(43) represent a nonlinear programming problem which may be solved by known iterative procedures when the gradient of the function (42) is known. Since for the sub-problem Lagrangians

$$\phi_i = F_i(x_i) + \lambda_i G_i(x_i) + \mu_i\left[H_i(x_i) - y_i\right]$$

the well known property

$$\frac{dF_i(\tilde{x}_i)}{dH_i(\tilde{x}_i)} = \mu_i, \quad i = 1, \ldots, m \qquad (44)$$

holds, and the Lagrange multipliers μ_i may be derived by the 1st-level controllers, the gradient of (42) may also be determined. The 1st-level controllers derive then components of the gradient for the 2nd-level, and an iterational optimization procedure can be realized. As shown in Ref. [18], this procedure may be extended to the case when F_i, H_i are functionals and G_i - operators in the Banach spaces.

Another interesting approach, used by economists [16], has been based on the iterational solution of a fictitious game which is being played between level 1 and 2. To explain this approach we shall consider the dual problems of linear programming:

Problem 1	Problem 2
$\max_{\underline{x}} (\underline{c}, \underline{x})$	$\min_{\underline{y}} (\underline{y}, \underline{b})$
$\underline{Ax} \leqslant \underline{b}$	$\underline{yA} \geqslant \underline{c}$
$\underline{x} \geqslant 0$	$\underline{y} \geqslant 0$

where $\underline{b}$ is the magnitude of the resources used in the given technological processes with the intensity $\underline{x}$. The intensity vector $\underline{x}$ should be chosen in Problem 1 in such a manner that the global production income $(\underline{c}, \underline{x})$ be maximum. In Problem 2 the cost of the resources consumed $(\underline{y}, \underline{b})$ should be minimized by a proper choice of the price-vector $\underline{y}$. The constraints $\underline{Ax} \leqslant \underline{b}$, $\underline{yA} \geqslant \underline{c}$ represent the restrictions imposed on the magnitude of resources consumed and production costs,respectively. Let us assume that the solutions $\bar{\underline{x}} \neq 0$, $\bar{\underline{y}} \neq 0$ of problem 1 and 2, respectively, exist and that the saddle-point relation

$$\max_{\underline{x} \in X} (\underline{c}, \underline{x}) = \max_{\underline{y} \in Y} \underline{y}, \underline{b} = (\underline{c}, \bar{\underline{x}}) = (\bar{\underline{y}}, \underline{b}) = K \qquad (45)$$

holds, X, $\bar{X}$, Y, $\bar{Y}$ representing the sets of admissible and optimum solutions for problems 1 and 2, respectively.Let us also assume that the matrix $\underline{A}$ consists of n sub-matrices, viz.,

$$\underline{A} = \|\underline{A}_1, \underline{A}_2, \ldots, \underline{A}_n\|$$

and the vectors $\underline{x}$, $\underline{c}$ consists of n sub-vectors

$$\underline{x} = [\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n], \quad \underline{c} = [\underline{c}_1, \underline{c}_2, \ldots, \underline{c}_n]$$

We now introduce the vector $\underline{u}$ with components $\underline{u}_i$, $i = 1, \ldots, n$, which have the same dimensions as $\underline{b}$ and $\sum_{i=1}^{n} \underline{u}_i = \underline{b}$. This vector shall be called the central strategy, and the linear programming sub-problems

$$\max_{x_i} (\underline{c}_i, \underline{x}_i), \quad i = 1, \ldots, n \qquad (46)$$

$$\underline{A}_i\underline{x}_i \leqslant \underline{u}_i$$

$$\underline{x}_i \geqslant 0$$

shall be referred to as sector optimization.

Here $\underline{x}_i$ represent the production-intensities of sector i. Vectors $\underline{y}_i$, $i = 1, \ldots, n$, appearing in the corresponding dual of (46) represent the prices in the sector i , when the magnitude of resources in that sector is $\underline{u}_i$. Hence the central strategy consists in finding the optimum distribution of the given magnitude of resources among n sectors,and the optimum sector strategies $\bar{\underline{x}}_i$, $i = 1, \ldots, n$,consist in finding the corresponding optimum production intensities.

Let us assume further that for each admissible central strategy $\underline{u}_i$ there exist vector functions

$$\underline{\varphi}_i(\underline{u}_i) = \max_{\underline{x}_i \in X_i(\underline{u}_i)} (\underline{c}_i, \underline{x}_i) = \max_{\underline{y}_i \in Y_i(\underline{u}_i)} (\underline{y}_i, \underline{u}_i) \qquad (47)$$

and denote

$$\underline{\varphi}(\underline{u}) = \sum_{i=1}^{n} \underline{\varphi}_i(\underline{u}_i)$$

The two-level optimization consists in:

(a) Finding the admissible central strategies $\underline{u} = \left[\underline{u}_1, \underline{u}_2, \ldots, \underline{u}_n\right]$ which ensure the global maximum, i.e.solving the concave programming problem

$$\max_{\underline{u} \in \underline{U}} \underline{\varphi}(\underline{u}) \qquad (48)$$

(b) Finding the optimum production intensities in each sector $\bar{\underline{x}}_i$, i.e. solving the linear programming problem

$$\max_{\underline{x}_i} (\underline{c}_i, \underline{x}_i) \qquad (49)$$

$$\underline{A}_i\underline{x}_i \leqslant \bar{\underline{u}}_i, \quad i = 1, 2, \ldots, n$$

Since the effective determination of $\underline{\varphi}(\underline{u})$ is not easy, this problem has been reduced to the two person polyhedral game.

The first (maximizing) player strategy is the vector $\bar{\underline{u}} = \left[\bar{\underline{u}}_1, \ldots, \bar{\underline{u}}_n\right] \in \underline{U}$, and the second (minimizing) player strategy is the vector $\underline{v} = \left[\underline{y}_1, \underline{y}_2, \ldots, \underline{y}_n\right] \in \underline{V}$.

The game value is

$$K = \max_{\underline{u} \in \underline{U}} \min_{\underline{v} \in \underline{V}} (\underline{v}, \underline{u})$$

The iteration process, known as the Brown fictitious play has been used to find the best game strategies $\bar{\underline{u}} \in \underline{U}$, $\bar{\underline{v}} \in \underline{V}$. This process has an interesting economic interpretation.As already shown, the initial optimization problem has been reduced to the two-level fictitious play between sectors (1st-level) and center (2nd-level). The 2nd-level strategy is the admissible distribution of resources and the 1st-level strategy - the admissible prices in dual problems. In the process of iterational solution each sector optimizes its own strategy according to the resources received from the 2nd-level, and after solving the dual problem it sends the result of optimization to the 2nd-level. The 2nd-level solves the problem of optimum distribution of resources and sends a new distribution strategy to 1st-level, etc.

The decentralized optimization process derived in this way has proved to be usefull in the planning of socialist economy[16].

4. Two-level Adaptive Optimization of Interacting Systems

So far the assumption was made that the optimized processes were deterministic and completely known to the controllers. However, in many practical systems the information on plant characteristics may be incomplete. In these cases one may use the known adaptive control methods in which the controller identifies the plant characteristics during the control actions by proper organization of the control actions and observation of output reactions. We shall consider a simple example of an iterational procedure, based on the so called stochastic approximations, which can be realized in the form of a decentralized, two-level control system.

Let us now consider a simple regulator system, which consists of n sub-systems including processes P_i and local controllers C_i, $i = 1, 2, \ldots, n$, which is shown in Fig. 6 for $n = 2$.

The input-output relation for the sub-systems

$$y_i = f_i(x_i)$$

and the additive interactions

$$z_{ji} = \varphi_{ji}(x_j),\ i,j = 1, 2, \ldots, n$$

are given continuous functions of x_j.

It is desired to obtain the resulting outputs

$$f_i(x_i) + \sum_{\substack{j=1 \\ j\neq i}}^{n} \varphi_{ji}(x_j), \; i = 1, 2, \ldots, n$$

equal to the given numbers Y_i.

If no interactions were present, each controller C_i could determine the required control values $x_i = \bar{x}_i$ by solving the equation

$$f_i(x_i) = Y_i, \quad i = 1, 2, \ldots, n$$

For this purpose it is convenient to solve the equivalent equation

$$x_i = x_i + a_i\left[Y_i - f_i(x_i)\right] = F_i(x_i), \quad i = 1, \ldots, n \qquad (50)$$

where the numbers a_i are chosen in such a way that the functions $F_i(x_i)$ satisfy the contraction conditions in the intervals X_i including $\bar{x}_i$:

$$\left|F_i(x_i') - F_i(x_i'')\right| \leqslant \alpha \left|x_i' - x''\right|, \quad \alpha < 1, \; i = 1, 2, \ldots, n$$

where x_i', x_i'' - arbitrary points in X_i .

The values $\bar{x}_i$ can then be derived by iterations

$$x_i^{(k+1)} = F_i(x_i^{(k)}), \; k = 0, 1, 2, \ldots, n, \; i = 1, 2, \ldots, n \quad (51)$$

starting with the arbitrary values $x_i^{(0)} \in X_i$, $i = 1,2, \ldots,n$. It is well known that $\lim_{k \to \infty} x_i^{(k)} \to \bar{x}_i$, and the solution obtained is unique.

The iterations can be also used when the explicit form of the input-output relations is unknown, but the controllers can observe the outputs $y_i^{(k)}$, which correspond to the fixed input $x_i^{(k)}$, using feedback loops (denoted by the dashed line in Fig. 6). Since these observations are frequently influenced by random noise, in the present case we are interested in the expected values of $y_i^{(k)}(\omega)$, i.e.

$$E\left\{y_i^{(k)}(\omega) \middle| x_i^{(k)}(\omega)\right\} = f_i(x_i^{(k)}), \; i = 1, 2, \ldots,n, \; k = 0,1,\ldots$$

where ω is a random variable.

The functions $f_i(x_i)$ should be now treated as regression functions, and the problem which faces us is the solution of the regression equations

$$f_i(x_i) - Y_i = 0, \quad i = 1, 2, \ldots, n$$

by iterations, using values $y_i^{(k)}(\omega)$ taken from observations.

This may be done by the so called stochastic approximations having the form

$$x_i^{(k+1)}(\omega) = x_i^{(k)}(\omega) + a_n\left[Y_i - y_i^{(k)}(\omega)\right], \quad k = 0,1,\ldots, \quad i = 1, \ldots, n$$

which, as shown by Robbins, Monro [35], will converge stochastically to the values $\bar{x}_i$, $i = 1, \ldots, n$, i.e.

$$\lim_{k \to \infty} E\left\{ \left\|x_i^{(k)}(\omega) - \bar{x}_i\right\| \right\} = 0$$

if certain regularity conditions hold. The regularity conditions include apart of the contractions the requirement that the numbers C_1, C_2, C_3 exist such that:

(a) the probability $P\left\{|y(x)| < C_1\right\} = 1$,

(b) $\frac{C_2}{k} \leqslant a_{k-1} \leqslant \frac{C_3}{k}$, $k = 1, 2, \ldots$ and the dispersion of $x_i^{(0)}$ is finite.

The iterations can be also used when interactions are present. In that case we get

$$x_i^{(k+1)} = x_i^{(k)} + a_i\left[Y_i' - f_i(x_i^{(k)})\right] = F_i(x_1^{(k)}, \ldots, x_n^{(k)}) \qquad (52)$$

$$Y_i' = Y_i - \sum_{\substack{j=1 \\ j \neq 1}}^{n} \varphi_{ji}(x_j^{(k)}) \qquad (53)$$

instead of (50) and (51), or when vector notation is used (52), (53) can be written:

$$\underline{x}^{(k+1)} = \underline{F}\left[\underline{x}^{(k)}\right], \quad k = 0, 1, \ldots \qquad (54)$$

where $\underline{x} \equiv \left[x_1, x_2, \ldots, x_n\right]$, $\underline{F} \equiv \left[F_1, F_2, \ldots, F_n\right]$ is a nonlinear continuous operator in n dimensional space E^n. If $\underline{F}$ is a contracting operator in a set $X \subset E^n$, and $\underline{x}^0 \in X$, the iterations (54) will converge to the unique solution $\bar{\underline{x}} \equiv [\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n] \in X$.

The calculations corresponding to (52), (53) can be implemented in the two-level form, shown in Fig. 6 for $n = 2$, where the 2nd-level controller C derives the values Y_i' by (53), and the 1st-level controller derive $x_i^{(k+1)}$ by (52). The advantage of the two-level process is that it utilizes the same control algorithms for level 1 as in the case without interactions. However, it requires the exchange of information between 1st- and 2nd-level controllers.

The control processes (52), (53) can be also realized when the values $y_i^{(k)}(\omega)$, $z_{ji}^{(k)}(\omega)$, obtained from observations, are used instead of $f_i\left[x_i^{(k)}\right]$, $\varphi_{ji}\left[x_{ji}^{(k)}\right]$. In that case, instead of (52), (53), we get the following algorithms for level 1

$$x_i^{(k+1)} = x_i^{(k)} + a_i^{(k)}\left[Y_i'^{(k)}(\omega) - y_i^{(k)}(\omega)\right], \quad i = 1,2,\ldots,n \quad k = 0,1,\ldots \tag{55}$$

and for level 2

$$Y_i'^{(k)}(\omega) = Y_i - \sum_{\substack{j=1 \\ j \neq i}}^{n} z_{ji}^{(k)}(\omega) \qquad i = 1,2,\ldots,n \tag{56}$$

If the regularity conditions for the multidimensional case hold, the two level iteration processes (55), (56) converges stochastically to the solutions $\bar{x}_i$, $i = 1, \ldots, n$, of the regression equation

$$Y_i - f_i(x_i) - \sum_{\substack{j=1 \\ j \neq i}}^{n} \varphi_{ji}(x) = 0$$

One should observe that the processes (50) - (54) can be easily extended to the case when the outputs Y_i derived, of sub-systems P_i, are given functions (F_i) of the outputs $f_i(x_i)$ and interactions $\varphi_{ji}(x_j)$, $i,j = 1, 2, \ldots, n$, i.e.

$$Y_i = F_i\left[f_i(x_i),\ \varphi_{1i}(x_1),\ \ldots,\ \varphi_{ni}(x_n)\right] \tag{57}$$

We shall also assume that it is possible to derive $f_i(x_i)$ as the unique continuous functions (ϕ_i) of Y_i and φ_{ji};

$$f_i(x_i) = \phi_i\left[Y_i,\ \varphi_{1i}(x_1),\ \ldots,\ \varphi_{ni}(x_n)\right] \tag{58}$$

$$i = 1,\ \ldots,\ n$$

Then, in order to solve the regression Eqs. (57) or the equivalent Eqs. (58), we can use the processes

$$x_i^{(k+1)} = x_i^{(k)} + a_i\left[\phi_i^{(k)} - f_i(x_i^{(k)})\right]$$

$$\phi_i^{(k)} = \phi_i\left[Y_i,\ \varphi_{1i}(x_1^{(k)}),\ \ldots,\ \varphi_{ni}(x_n^{(k)})\right]$$

$$i = 1,\ \ldots,\ n, \quad k = 1,\ 2,\ \ldots$$

It is also possible to optimize the system when the values $\varphi_{ji}(x_j^{(k)})$, $f_i(x_i^{(k)})$ are obtained by observation. The latter case is illustrated in Fig. 7, where $n = 2$ and the values $\varphi_{ji}(x_j^{(k)})$, $f_i(x_i^{(k)})$ are denoted by $z_{ji}^{(k)}(\omega)$ and $y_i^{(k)}(\omega)$ respectively.

In reference [25] it was shown that the stochastic approximations can be also used for the decomposition of complex optimization problems.

5. Aggregation and Synthesis of Optimum Organizational Structures of Multi-level Systems

Many examples of complex technical, economic, social and biological systems or organizations exist which are controlled by several cooperating or interacting decision centers or controllers. These systems are frequently organized according to the hierarchic principle, i.e. each sub-system, consisting of a controller and controlled processes, receives certain directions, information or resources from a higher-level controller and at the same time it can influence the performance of lower-level sub-systems.

An interesting feature of the hierarchic structure is that the particular sub-systems are autonomous in the sense that every controller derives his control-strategy on the basis of a limited amount of information.The higher is the control-level the smaller is the global amount of information. In other words the information is "compressed" or aggregated when it travels from the lower to the higher levels of the hierarchic structure. On the other hand the directions of the higher-levels passing to lower-level sub-systems become supplemented by information suitable for local conditions. This corresponds to the "decompression" of information. It should be also noted that in such systems there exist a decentralization of decision processes, which permits the controllers to deal with (or transform) a limited amount of information (or calculations) in a fixed time interval. This feature also permits effective control of complex processes or organizations by standard analogue and digital computers or by human operators.

Examples are also known of systems or organizations whose performance is evaluated as poor, inefficient or bureaucratic. Many authors, including Parkinson, have contributed much to a better understanding of these organizations.However,the evaluation of the quality of organization in these researches has been performed on the basis of emotions rather than strict analysis.

In the present section we shall consider a simple model of hierarchic organization, shown in Fig. 8, consisting of controlled processes (denoted by circles), controllers (denoted by rectangles), and communication or transport means linking the controllers and processes.

We shall show that the performance of a controlled process can be described by a single number,referred to as the process quality index P. Q. I. ,and that the losses due to the transmission of information or resources can be again described by numbers, referred to as loss coefficients L.C. . The resulting performance index of the whole organization can then be derived. Comparing organizations, described by different performance indices, it is also possible to choose from the given sets of controllers and processes the best organizational structure or, in other words, it is possible to solve the synthesis problem.

The main concepts which are used in this section are based on the ideas described in [19, 23].

5.1. Performance Measure of Hierarchic Organizations

Let us consider a simple hierarchic organization, shown in Fig. 8, which consists of controlled processes $P_1, \ldots, P_n$, local (1st-level), controllers $C_1, \ldots, C_n$, supervisory (2nd-level) controller C_{1n} and transmission lines $L_1, \ldots, L_n$, which link C_{1n} with $C_1, \ldots, C_n$.

The operation of controllers is specified by given objective functionals, which together with the process equations and constraints can be used for determination of the optimum control algorithms. Since in the present section we are interested mainly in the organizational aspects of complex systems, we shall not devote much attention to the derivation of the optimum control algorithms, but we shall concentrate on the notion of the so called optimum performance characteristics(O.P.C.) of optimum processes, which are essential for the evaluation of the organization quality.

For this purpose let us consider a dynamic process which is described by a given operator A :

$$y = A(x), \quad y, x \in X$$

where x is the controlled input, y - output process, and X is, generally speaking, a Banach space of functions of time t. Now let us assume that there exists a unique input $\bar{x} \in X$, which minimizes the given objective functional $F(x)$, $x \in X$ (called the control cost), subject to a number of equality or inequality constraints:

$$\Phi(x) \geqslant B, \ \ldots, \ \Psi(x) \leqslant Z$$

where Φ, ..., Ψ given functionals in X and $B, \ldots, Z$ - given nonnegative numbers which may represent the desired output production, magnitude of resources available, optimized time-interval, etc.

If $\bar{x}$ can be effectively derived as a function of time t and $B, \ldots, Z$, it is also possible to derive a function $A =$

$= F[\bar{x}(t, B, \ldots, Z)] = f(B, \ldots, Z)$, which represents the value of control cost as a function of "outer parameters" $B, \ldots, Z$ and which does not depend on the time variable t. The function $A = f(B, \ldots, Z)$ will be called the O.P.C. of processes optimized.

As an example we shall consider a transport process, using electrical motor, which should shift an inertial load to the given distance Y in the given time interval T with minimum energy consumption.

The position of the load $y(t)$ can be described by the operator

$$y(t) = A(x) = y(0) + a\int_0^t (t - \tau)\, x(\tau)\, d\tau \tag{59}$$

where a - given coefficient, $x(\tau)$ - current in the armature of the motor.

The optimization problem consists in finding such a control-current $\bar{x}(t) \in L^2[0,T]$, which minimizes the energy cost

$$F(x) = \int_0^T [x(\tau)]^2\, dt \tag{60}$$

subject to the constraints

$$\Phi(x) = y(T) - y(0) = Y$$

$$\Psi(x) = \left.\frac{dy(t)}{dt}\right|_{T=0} = 0$$

It can be shown that

$$\bar{x}(t) = \frac{3Y}{aT^3}\left(\frac{T}{2} - t\right) \tag{61}$$

and

$$A = F(x) = \frac{3Y^2}{4a^2T^3} \tag{62}$$

Relation (62) can be also written in the form

$$AT^3Y^{-2} = k^2, \quad k = \sqrt{3}/2a \tag{63}$$

and can be called the O.P.C. of the transport process.

Another example is a P.E.R.T. x)-project in which the cost C of each operation is assumed to be inversly proportional to the optimization time T . In that case we get for the O.P.C.

$$CT = k^2$$

where k is a given coefficient.

In Reference [22] the O.P.C. have been derived for many dynamic optimized processes. For many cases they assume a simple analytic form:

$$A^{\alpha} B^{\beta} \ldots Y^{\psi} Z^{\omega} = (k)^q , \quad q = \alpha + \beta + \ldots + \omega \qquad (64)$$

where $A, B, Y, Z, \alpha, \beta, \ldots, k$ are positive numbers and $\omega, \psi, \ldots$ negative numbers xx). Since the smaller is k the better the properties of the optimized processes (e.g. in the case of (63): $AT^3Y^{-2} = \frac{3}{4a^2}$ and fixed T, Y the value of A is small when $k = \frac{\sqrt{3}}{2a}$ is a small number), k can be called the quality index.

Assuming that the O.P.C. of the sub-systems K_i including processes P_i and local controllers C_i , $i = 1,2,\ldots,n$ (see Fig. 8) are given, $A_i = f_i(B_i, \ldots, Z_i)$, we can concentrate on the derivation of the O.P.C. for the aggregated system K_{1n}, which apart from the sub-systems K_i includes a supervisory controller C_{1n} and transmission lines $L_1, \ldots, L_n$.

We shall take $\sum_{i=1}^{n} \alpha_i A_i$ as the objective function for an aggregated system and determine such values of $B_i, \ldots, Z_i$ which minimize

$$\sum_{i=1}^{n} \alpha_i A_i = \sum_{i=1}^{n} \alpha_i f_i(B_i, \ldots, Z_i) \qquad (65)$$

subject to the set of aggregated constraints

x) Process Evaluation and Review Technique.

xx) The well known economic model of Cobb-Douglas is a special case of the model described by (64).

$$\sum_{i=1}^{n} \beta_i B_i \leqslant B, \ldots, \sum_{i=1}^{n} \omega_i Z_i \geqslant Z \qquad (66)$$

where $\alpha_i, \beta_i, \ldots \geqslant 1$, $\omega_i, \ldots \leqslant 1$ and $B, \ldots, Z$ are given positive numbers. We have here a nonlinear programming problem. We now assume that there exists a unique solution $B_i^o, \ldots, Z_i^o$, $i = 1, 2, \ldots, n$, and that it is possible to compute the function

$$A = \sum_{i=1}^{n} \alpha_i f_i(B_i^o, \ldots, Z_i^o) = f(B, \ldots, Z)$$

which will be called the O.P.C. of the aggregated system. There exist many industrial and economic systems which are aggregated and optimized according to (65), (66).

As an example we may consider the integrated electric power system which consists of n power stations with given performance functions $F_i = f_i(P_i)$, $i = 1, 2, \ldots, n$, relating the fuel cost F_i and the amount of power production P_i. The global production $\sum_1^n P_i \eta_i$ (where η_i are the so called penalty factors, which represent power losses in transmission lines) should be at least equal to the power demand P and the global fuel cost $\sum_{i=1}^{n} \alpha_i f_i(P_i)$ (where α_i represent fuel losses during transport) should be minimized by proper dispatching of the power production P_i.

It is possible to show [22] that for certain types of O.P.C. the derivation of the aggregated O.P.C. is relatively simple. For example, in the case of processes with the O.P.C. in the form $A_i^\alpha B_i^\beta \ldots Z_i^\omega = (k_i)^q$, $i = 1, \ldots, n$, the aggregated O.P.C. becomes $A^\alpha B^\beta \ldots Z^\omega = (k)^q$, where

$$K = \sum_{i=1}^{n} k_i \lambda_i, \quad \lambda_i = \left[\alpha_i^\alpha \beta_i^\beta \ldots \omega_i^\omega \right]^{1/q} \qquad (67)$$

and A_i^o, B_i^o, ..., Z_i^o can be determined from linear equations. A similar property is characteristic of the functions

$$A_i = \beta_i f\left[\frac{B_i + \bar{B}_i}{\beta_i}\right] + a_i, \tag{68}$$

where β_i, a_i, $\bar{B}_i$ - given numbers, and f is a monotonic differentiated function having a unique inverse $[f']^{-1}$ [22]. It is possible also to show that a continuous O.P.C. can be piecewise approximated by functions (68) with the desired degree of accuracy.

It should be observed that when the O.P.C.-s are described by a function of type (67) or (68), the aggregation and optimization processes can be applied to multi-level structures yielding at each stage the same form of O.P.C. with quality indexes which can be derived from simple relations of the type (67).

In other words, the amount of variables or information which is to be considered at each control level is strictly limited.

It is also possible to evaluate qualities of different organizational structures. Let us assume, for example, that three different processes, described by (64), with performance k_1, k_2, k_3 and three different organizations shown in Fig. 9, are given. The corresponding quality indexes, derived by (67), become

$$k_a = \lambda_1^1 k_1^1 + \lambda_2^1 k_2^1 + \lambda_3^1 k_3^1 \tag{69}$$

$$k_b = \lambda_1^{11} \lambda_1^1 k_1^{11} + \lambda_2^{11} \lambda_1^1 k_2^{11} + \lambda_2^1 k_2^1 \tag{70}$$

$$k_c = \lambda_1^1 k_1^1 + \lambda_1^{11} \lambda_2^1 k_1^{11} + \lambda_2^{11} \lambda_2^1 k_2^{11} \tag{71}$$

where λ_1, λ_2, λ_3, λ_{12}, λ_{23} represent losses introduced by transmission lines which link the respective controllers.

Now we are able to compare different organizations what will be done in the next section.

5.2. Synthesis and Optimum Control or Organizational Structures

As it follows from (67) (compare also (69) - (71)), the resulting quality index of an organization which consists of n controlled processes with given performance indexes $k_1, k_2 \ldots,$ $\ldots, k_n$ can be written in the form

$$k = \sum_{i=1}^{n} k_i l_i \tag{72}$$

where l_i - loss indices depending on the organization structure. It is also obvious that the smaller is k the better the global system performance.

The minimum value of k can be obtained by:

(a) assigning processes to the given fixed structure, i. e. to the given, ordered set $\left\{l_i\right\}_1^n$:

$$l_1 \leqslant l_2 \leqslant \ldots \leqslant l_n$$

the indexes ν in the set $\left\{k_\nu\right\}_1^n$ should be assigned in such a way that (72) is minimum;

(b) permitted reorganization of the structure, by changing the position of the controllers and transmission lines, which decrease the value of (72).

As far as the assignment problem is concerned the following two theorems may present certain interest.

Theorem 1. Let two sets $\left\{l_\nu\right\}_1^n$, $\left\{k_j\right\}_1^n$ of positive numbers be given. The set of $K = \sum_1^n k_i l_i$ corresponding to any possible assignment of indexes ν , j, is contained in the interval

$$\left[\frac{2\,\bar{l}\,\bar{k}}{\sqrt{lK/Lk} + \sqrt{Lk/lK}}, \bar{l}\,\bar{k}\right] \tag{73}$$

where

$$\bar{l} = \left\{\sum_{i=1}^{n} l_i^2\right\}^{1/2}, \quad \bar{k} = \left\{\sum_{i=1}^{n} k_i^2\right\}^{1/2}$$

$l = \min_i l_i$, $k = \min_i k_i$, $L = \max_i l_i$, $K = \max_i k_i$.

The upper limit $(\bar{l}\,\bar{k})$ is reached if and only if $k_i = \alpha l_i$ $i = 1, \ldots, n$, α = const. The lower limit in (73) is reached if and only if $p = \frac{L/l}{L/l + K/k}$ is an integer and

$$\left.\begin{matrix} k_i = k \\ l_i = L \end{matrix}\right\} i = 1, 2, \ldots, p, \qquad \left.\begin{matrix} k_i = K \\ l_i = l \end{matrix}\right\} i = p+1, \ldots, n$$

The proof of this theorem is based on the known Cauchy and G. Polya and C. Szego inequalities [23].

Theorem 2. Let two sets $\{l\}_1^n$, $\{k\}_1^n$ of positive numbers be given. The value of $K = \sum_1^n l_i k_i$ is minimum, if

$$k_1 \leqslant k_2 \leqslant \cdots \leqslant k_n, \quad l_1 \geqslant l_2 \geqslant \cdots \geqslant l_n \tag{74}$$

or if

$$k_1 \geqslant k_2 \geqslant \cdots \geqslant k_n, \quad l_1 \leqslant l_2 \leqslant \cdots \leqslant l_n \tag{75}$$

These conditions become also necessary in the case of strict inequalities in (74), (75).

The validity of this theorem for $n = 2$ is obvious. For $n > 2$ it can be proved by induction [23].

Example. Let us consider two organizations shown in Fig. 9b and 9c and assume that $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_{12} = \lambda_{23} = \lambda > 1$. In the case of the system 9b we have $l_1 = \lambda < l_2 = \lambda^2$ and $l_2 = l_3 = \lambda^2$. Then, according to theorem 2 this organization is optimum if $k_1 \geqslant k_2 \geqslant k_3$. For the same reason the organization shown in Fig. 9c becomes optimum if $k_1 \leqslant k_2 \leqslant k_3$. These structures become equivalent when $\lambda = 1$.

The last example indicates that in order to get best results it is necessary, apart from optimum control of the processes, to reorganize the structure when the quality indexes of the sub-systems change with time. In other words, the higher level controllers should reorganize the system structure if necessary.

Theorem 2 can be used also for synthesis of multilevel structures. As an example we assume that n controllers C_i,

$i = 1, 2, \ldots, n$, and N processes, equiped with local controllers so that they can be completely described by the indexes

$$k_1 \leqslant k_2 \leqslant \cdots \leqslant k_N$$

are given. Transmission losses are assumed to be the same for each interconnection and $\lambda > 1$. The maximum amount of processes which can be controlled by controllers C_i is m_i, $i = 1, 2, \ldots, n$, respectively, and

$$m_1 \leqslant m_2 \leqslant \cdots \leqslant m_n, \quad \sum_{i=1}^{n} m_i = N$$

Besides, we assume that each controller can also optimize one sub-system of controllers and processes. The problem consists in determining the best organization of the controllers and processes.

Let us observe that the numbers

$$K_i = \sum_{j=m_i'}^{m_i} k_j, \quad m_i' = m_{i-1} + 1, \quad i = 1, 2, \ldots, n$$

satisfy the condition: $K_1 \leqslant K_2 \leqslant \cdots \leqslant K_n$, and the possible organizations will give loss coefficients of the form λ^k, $k = 1, 2, \ldots, n$.

Then, using theorem 2, we can obtain the structure shown in Fig. 10, with the quality index

$$K = \sum_{i=1}^{n} K_i l_i \tag{76}$$

where $l_i = \lambda^{n-1+1}$, and $l_1 > l_2 > \cdots > l_n > 1$.

This organization is optimum in the sense that no allowed reorganization (consisting in exchanging processes and sub-systems) exists which would decrease the values of K given by (76).

Other examples of synthesis of organizations and some extenssions are given in Ref. [23]. For example, one can assume that the numbers of processes m_i are not fixed and the loss coefficients λ_i are increasing functions of m_i . In that case the optimum number of control-levels depends, generally speaking, on the global number of processes N .

As an example we can assume $K_i = k$, $\lambda_i = \lambda$, $i = 1,2,\ldots,\ldots, N$, $N = m^2 = \text{const.}$, and compare the resulting losses for the single-level (l_I) and two-level (l_{II}) organizations of the type shown in Fig. 9.

We obtain

$$l_I = m^2 \lambda(m^2), \quad l_{II} = m^2 \left[\lambda(m)\right]^2$$

When m increases there exists, generally speaking, such a number $m = m_o$ that $l_{II} < l_I$. Let us assume for instance $\lambda(m) = 1 + \delta m$, then

$$l_I(m) = m^2(1 + \delta m^2) > l_{II}(m) = m^2\left[1 + \delta m\right]^2$$

when $m > 2/(1 - \delta)$.

So far it has been assumed that the systems under consideration were deterministic and stationary in time.However,in many systems the coefficients of O.P.C. as well as the loss coefficients may change at random with time as a result of environment changes, noise, etc. Optimum control of the systems of this type becomes more complicated. First of all it is necessary to consider all the performance functionals as expected values. Then it is expedient to observe the optimized processes in the past, and utilize the information about the process-parameters obtained in this way, for a better control-action in the future. It is well known that the systems acting in this fashion are called adaptive. Since during the observations (or in other words - identification of the processes) one cannot control the process effectively, the observation time should be as short as possible. On the other hand the accuracy of identification is an increasing function of observation time. It is therefore necessary to coordinate the identification and control action in such a way that the resulting performance is optimum. This can be done by applying the general theory of statistical decision functions.

It should be also observed that in the adaptive hierarchic systems the optimization process should be, generally speaking, accompanied by a process of reorganization of system structure. The convergence of these processes and the stabilization of the structure represent difficult theoretical problems.

Beside the random variation of processes and structures, in many large-scale cybernetic systems one can discover processes of structural evolution. In the industrial and management systems these processes depend on the scientific and technological progress, which creates new production branches. It depends also on the capital investments. Evolutionary processes may be accompanied, in turn, by the reorganization processes. If, for example, it is necessary in the given branch of industry to create a new technological process, the management may decide to change the existing organization by forming new departments and appoint new directors.

One should observe that the simple model of organization which was described in the previous sections may be also used for investigation of the processes in adaptive and evolutionary systems. However, due to limited space we shall not pursue these interesting considerations.

References

1. Abadie F., Williams A., Dual and parametric methods in decomposition. In: Recent advances in matehmatical programming. New York 1963.
2. Aoki M., Control of large-scale dynamic systems by aggregation. IEEE Trans. on Automatic Control, June 1968.
3. Balas E., An infeasibility - pricing decomposition method for linear programs. Operations Research 1966 No. 5.
4. Baumol W.J., Fabian T., Decomposition, pricing for decentralization and external economics. Management Sci. 1964 No. 1.
5. Benders J.F., Partitioning procedures for solving mixed-variables programming problems. Numerische Mathematik 1962 S. 238-252.
6. Brosilow C.B., Lasdon L.S., Pearson J.D., Feasible optimization methods for interconnected systems. Rep. No. SRC-70-A-

-65-25 Case Institute of Technology 1965.

7. Бурков В.Н., Распределение ресурсов как задача оптимального быстродействия. Автоматика и Телемеханика 1966 № 7.
8. Бусленко Н.П., Юркевич О.М., Об операциях над агрегатами в сложных системах. Техническая Кибернетика 1964 № 2.
9. Бусленко Н.П., К теории сложных систем. Техническая Кибернетика 1963 № 5.
10. Coviello G.J., An organizational approach to the optimization of multi-level systems. Joint Automatic Control Conf. June 1964.
11. Dantzig G.B., Wolfe P., The decomposition algorithm for linear programs. Econometrica 1961 No. 4.
12. Dantzig G.B., Wolfe P., Decomposition principle for linear programs. Operations Research 1960 No. 1.
13. Findeisen W., Parametric optimization by primal method in a two-level system. Archiwum Automatyki i Telemechaniki 1967 No. 4.
14. Голштейн Ю.Г., Об одном подходе к решении задач линейного программирования большой структуры. Доклады АН СССР 1966 Том 166. Вып. 5.
15. Kirchmayer L.K.: Economic operation of power systems. New York 1958.
16. Kornai J., Liptak T.: Two-level planning. Econometrica 1965 pp. 141-163.
17. Кухтенко А.Л., О теории систем с иерархической структурой управления. Киев 1966.
18. Kulikowski R., Decentralized optimalization of large scale systems. Paper submitted to IV IFAC Congr.
19. Kulikowski R., Optimum control of multidimensional and multilevel systems. In: Advances in control systems. Vol. 4. New York 1966.
20. Kulikowski R., Optimum control of dynamic interacting systems. Archiwum Automatyki i Telemechaniki 1966 No. 2.
21. Kulikowski R., Optimum control of aggregated multi-level systems. Proc. III IFAC Congr. London 1966.
22. Kulikowski R., Optimization of aggregated dynamic systems. Archiwum Automatyki i Telemechaniki 1966 No. 3.

23. Kulikowski R., Synthesis and optimum control of organization in large scale systems. Archiwum Automatyki i Telemechaniki 1967 No. 3.
24. Kulikowski R., On optimum control of nonlinear,dynamic industrial processes. Archiwum Automatyki i Telemechaniki 1967 No. 1.
25. Kulikowski R., Multilevel, adaptive optimalization of interacting systems. Bull. Ac. Pol. Sci. Cl. IV 1968 No. 2.
26. Kulikowski R., Synthesis and optimum control of organization in large scale cybernetic systems. Proc. V Congr. of Cybernetics, Namur 1967.
27. Küntzi H.P., Tan S.T.,Lineare Optimierung grosser Systeme. Berlin 1966.
28. Lasdon L.S., Schoeffler I.D., Decentralized plant control. Proc. 19th Annual ISA Conf., New York 1964.
29. Lasdon L.S., Schoeffler I.D., A multi-level technique for optimization. Proc. Joint Automatic Control Conf. 1965 pp. 85-93.
30. Mańczak K., Agregacja w pewnych wielopoziomowych zagadnieniach kwadratowego programowania. Archiwum Automatyki i Telemechaniki 1966.
31. Mesarovic M., Pearson J., Macko D., Takahara Y., On the synthesis of dynamic multilevel systems. Proc. III IFAC Congr., London 1966.
32. Nemhauser G.L., Decomposition of linear programs by dynamic programming. Naval Research Logistics Quart. June 1964.
33. Pearson J.D., Multilevel programming.Rep. No. SRC-70-A-65-25 Case Institute of Technology 1965.
34. Pearson J., Decomposition, coordination and multi-level systems. IEEE Trans. on Systems, Science, Cybernetics 1966 No. 1.
35. Robbins H., Monro S., A stochastic approximation method. Ann. Math. Stat. 1951 pp. 400-407.
36. Rosen J.B., Convex partition programming. In: Recent advances in mathematical programming. New York 1963.
37. Sanders I.L., A nonlinear decomposition principle. Operations Research 1965 No. 2.

38. Sekine Y., Decentralized optimization of an interconnected system. Trans. IEEE 1960 vol. CT-10 No. 2.

39. Sprague C.G., On the reticulation problem in multi-variable control systems. Joint Automatic Control Conf. June 1964.

40. Straszak A., On the structure synthesis problem in multi-level control systems. Preprints IFAC Symp. on System Engineering for Control Systems Design, Tokyo 1965.

41. Varaiya P., Decomposition of large-scale systems. SIAM J. on Control 1966 No. 1.

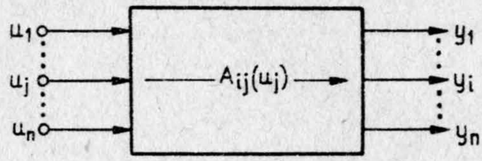

Fig. 1

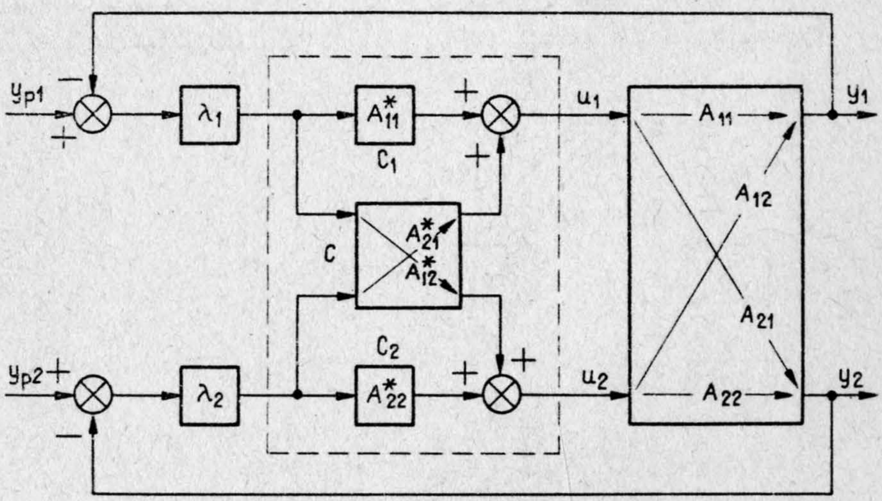

Fig. 2

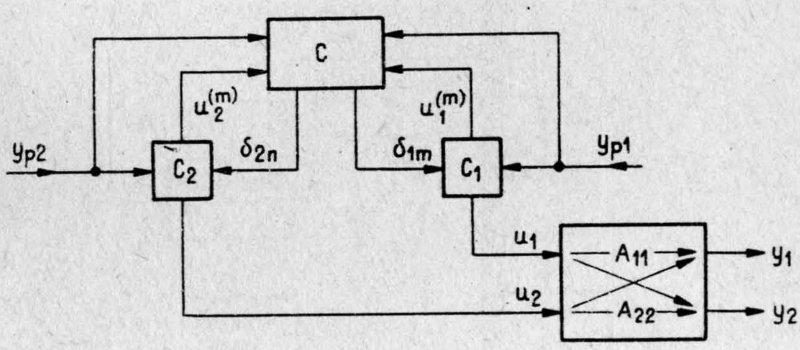

Fig. 3

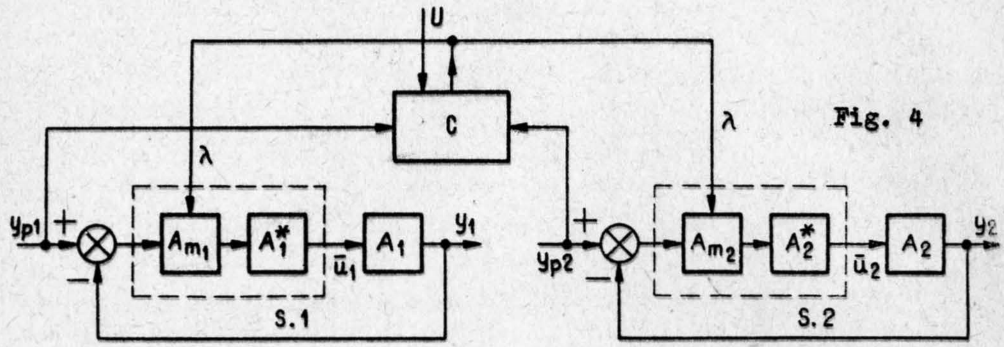

Fig. 4

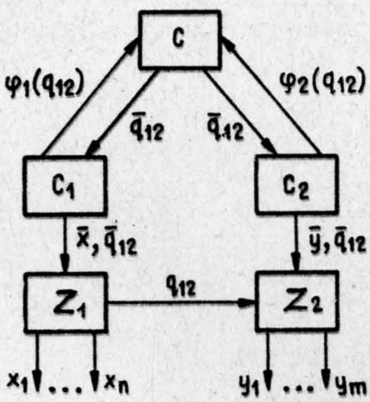

Fig. 5

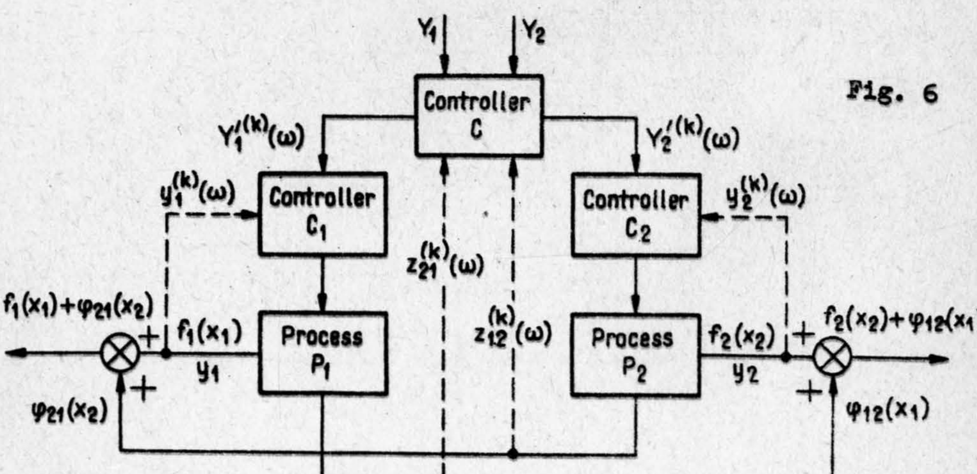

Fig. 6

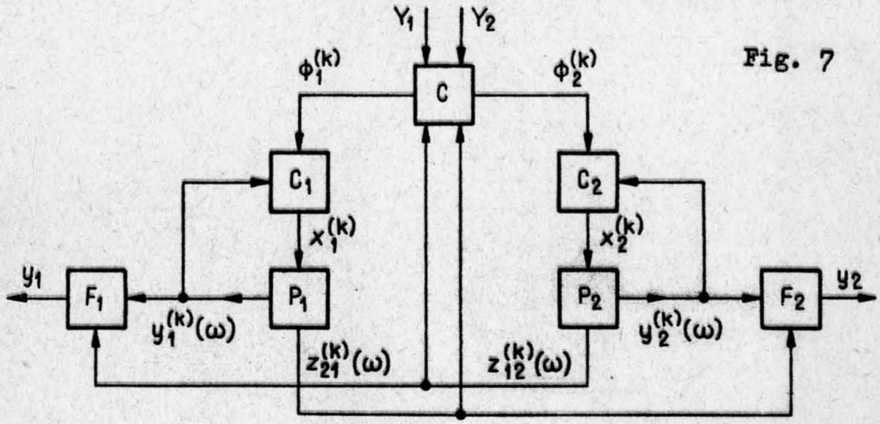

Fig. 7

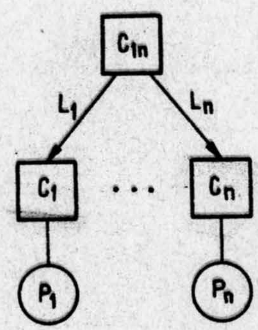

Fig. 8

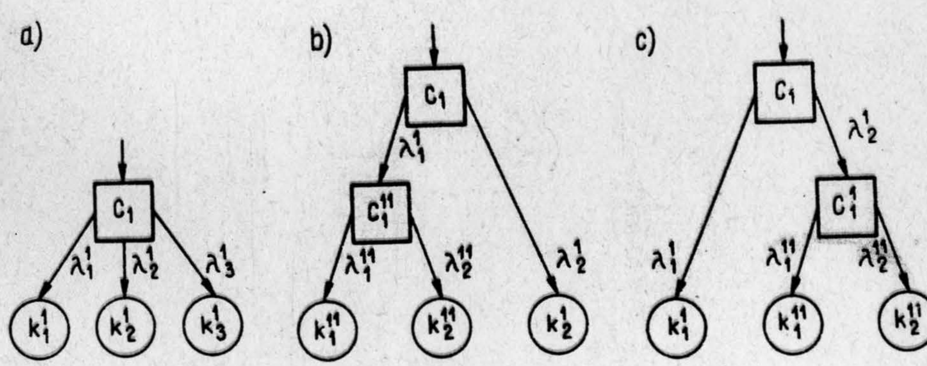

Fig. 9

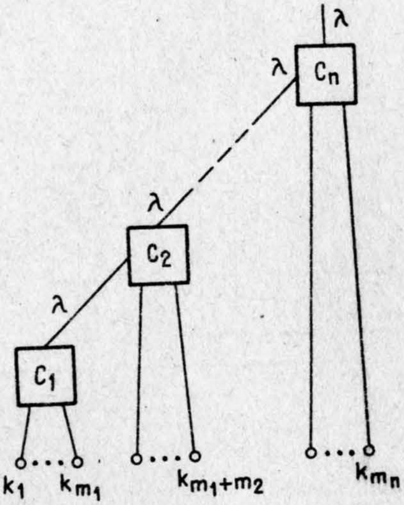

Fig. 10