A SURVEY OF PREVIOUS AND RECENT TRENDS IN CONTROL ENGINEERING

By

F. Csáki

Department of Automation. Technical University, Budapest

(Received September 22, 1972)

The aim of this paper is to give a brief survey of previous and recent trends of control theory and practice. The principles of control are widely used not only in the various fields of engineering practice but also in the sphere of certain other sciences such as economics, physiology, biology, social sciences and so on. Under such circumstances the character of the theory of control is becoming more and more interdisciplinary. Therefore, it could perhaps be of some interest to analyse the past and future trends in control science and technology.

Historical Development of Control Engineering

In the history of automation, by and large, three epochs can be distinguished. The first era begins in the antiquity by the activity of the Alexandrian Ktesibios (Fig. 1) and Heron (Fig. 2), the Byzantian Philon (Fig. 3). The connecting link, as in other fields of sciences, was created by Arab scientists and craftsmen such as Benu Musa (Fig. 4), Ridwan, Al Gazhari. In the early times of the first industrial revolution some new developments can be mentioned e. g. the level controllers of Polsunov (Fig. 5) and Wood, the temperature controller of Conti and Réaumur, the steam-pressure controllers of Mead and Watt (Fig. 6) or the power-loom of Jacquard with punch-cards. This first period of automation, which extends over about two thousand years, might be characterized by the heuristic approach, i.e. that the control systems were relatively simple in structure and could be realized without deeper insight of the theory.

The *second era* of control engineering begins with the thirties of the twentieth century and lasts on in our days. This epoch is closely related with the development and progress of *linear control theory* as well as the application of Boole algebra in open-loop control. A lot of control systems have been realized; many servomechanisms and process control systems have been projected and taken into application. Numerous open-loop control systems have

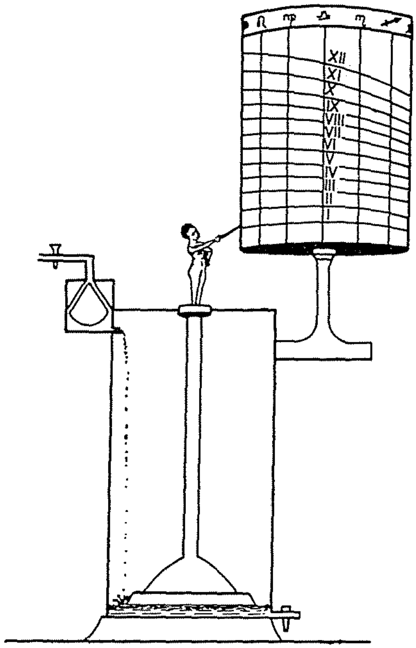

Fig. 1. KTESIBIOS' water clock

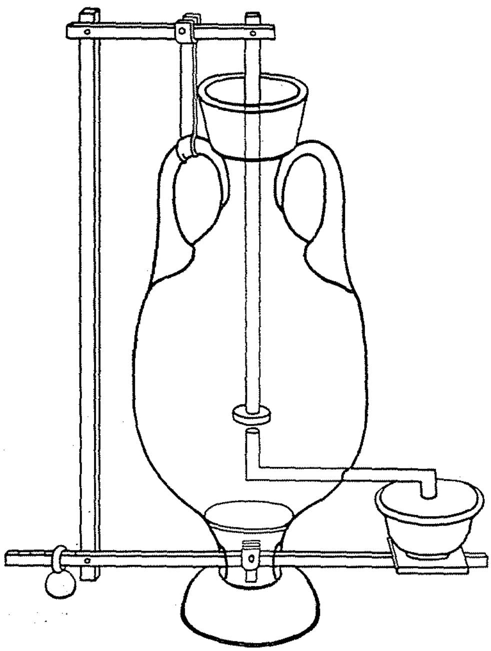

Fig. 2. Level controller of HERON

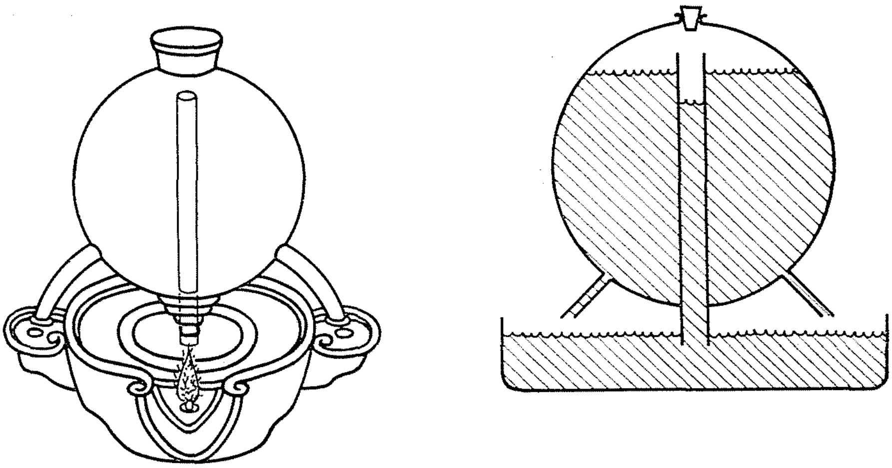

Fig. 3. Oil lamp of PHILON

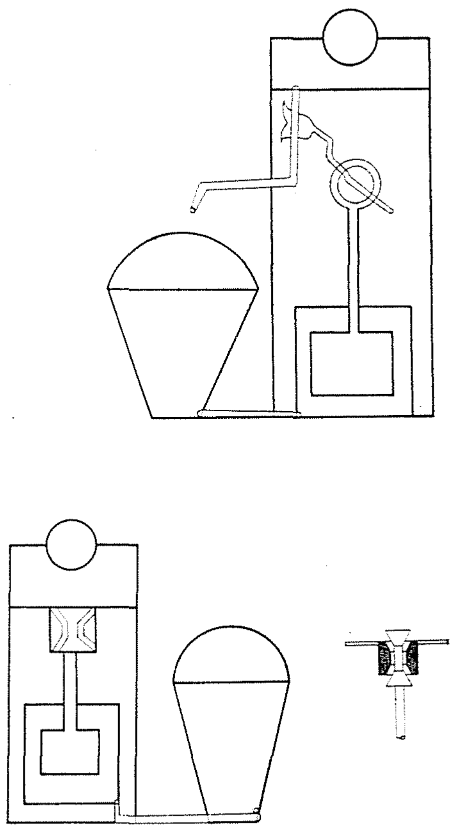

Fig. 4. Level controllers of BENU MUSA

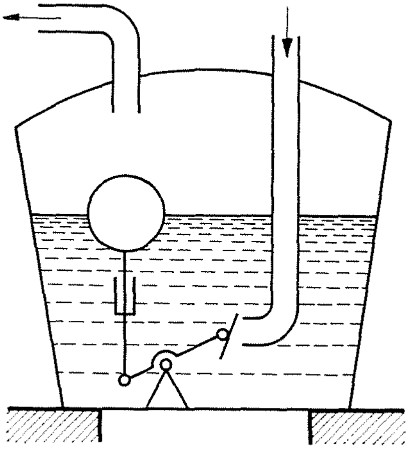

Fig. 5. Level controller of POLSUNOV

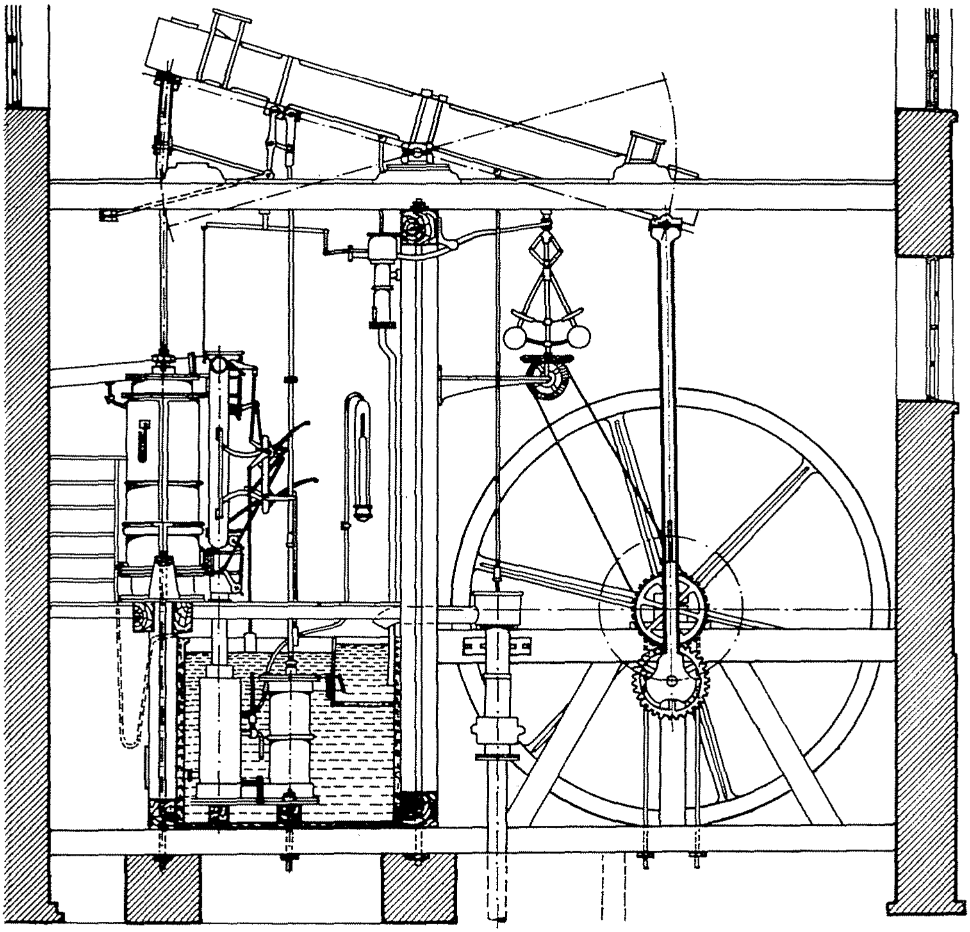

Fig. 6. WATT'S steam engine with fly-wheell regulator

been put into practice, e. g. the automatic telephone centrals, the short-circuit protection equipment of electric power systems and others. In this period the *closed-loop* and the *open-loop* controls develop seperately, and relatively few are of common application. In closed-loop control, complicated problems are attempted to be reduced, by allowable simplifications and neglects, to single-loop, time-invariant, linear problems. At the same time decoupled control loops are sought for. Initially the control systems are designed for *deterministic signals* and it is only during World War II that the investigation of *stochastic systems* is begun with. In this epoch the *transformation methods* are widely used: the LAPLACE, the two-sided LAPLACE, the FOURIER, the z and modified z transforms are applied. The simple control systems are analyzed and synthetized in the *frequency domain*, mainly by the aid of NYQUIST diagrams (Fig. 7), BODE plots (Fig. 8a and 8b) and NICHOLS curves (Fig. 9).

Instead of these trial and error methods some analytical methods, based on *integral criteria*, are being introduced. The WIENER-KOLMOGOROV theory and the *statistical design* of optimal filters, predictors and controllers are essentially also based on integral criteria. By the end of this second period of control engineering, analog and digital computers come into prominence. They are mainly used as design aids because their application facilitates to simulate linear and nonlinear, single-loop and multiloop, single-variable and

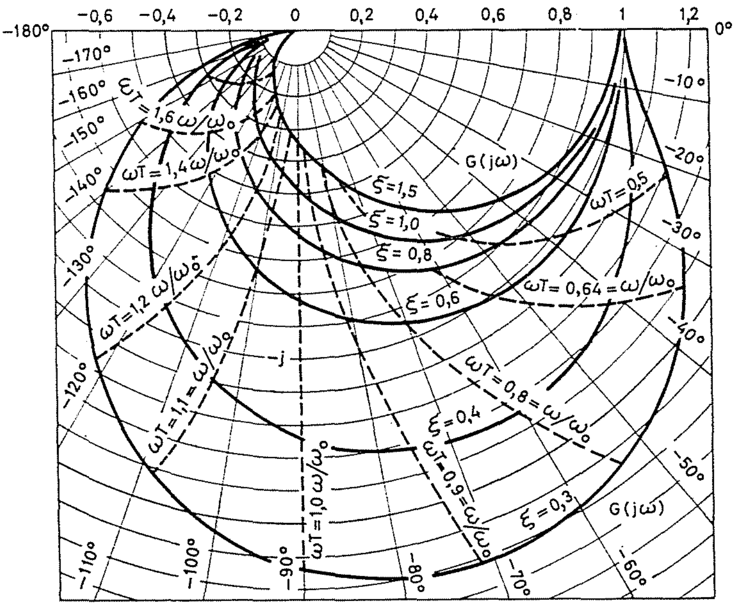

Fig. 7. NYQUIST plot of the second-order system

$$G(s) = \frac{1}{1 + 2\zeta Ts + T^2 s} = \frac{\omega_0^2}{s^2 + 2\zeta \omega_0 s + \omega_0^2}$$

multivariable systems. The principles of analog computers are general enough, whereas for each family of digital computers special simulation languages have to be developed such as MIMIC, MIDAS, DYSAC, MADBLOC, DSL, CSMP, CSSL. In some cases hybrid simulation gives the fastest results.

Finally, the *third era* of automation has begun about two decades ago. It is impossible to draw a sharp line but, broadly speaking, this third epoch can be characterized by the wide-spread application of digital computers not only in the projects of control systems but also in the operation of complicated processes. The application spectrum extends from the simple *off-line open-loop control systems* to the complicated *on-line closed-loop control systems.*

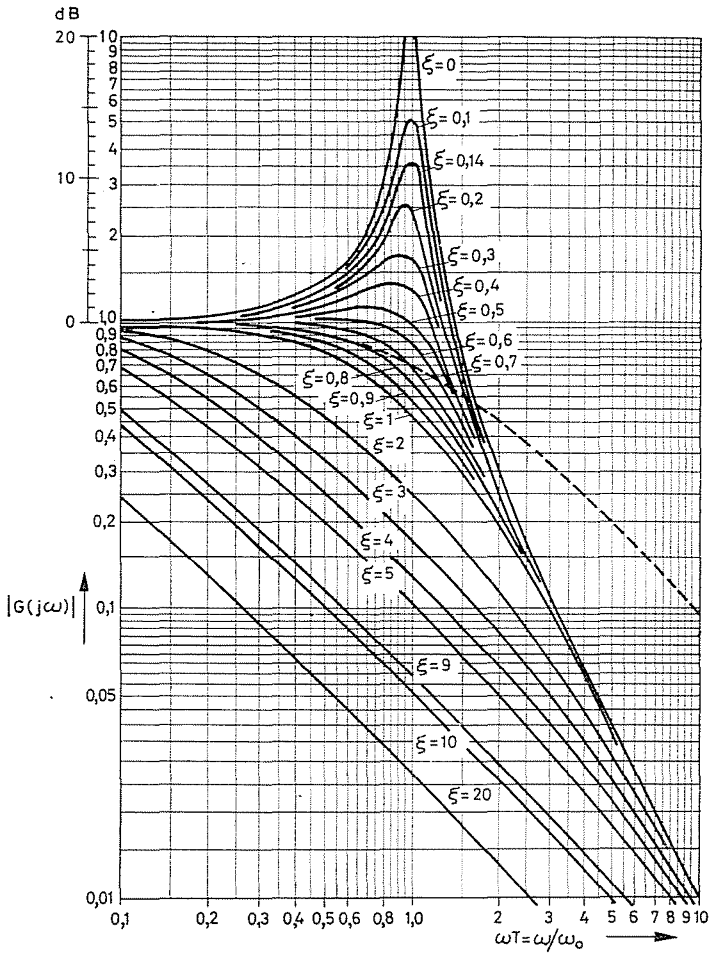

Fig. 8a. Bode magnitude plot of the second-order system

$$G(s) = \frac{1}{1 + 2\xi Ts + T^2 s} = \frac{\omega_0^2}{s^2 + 2\xi\omega_0 s + \omega_0^2}$$

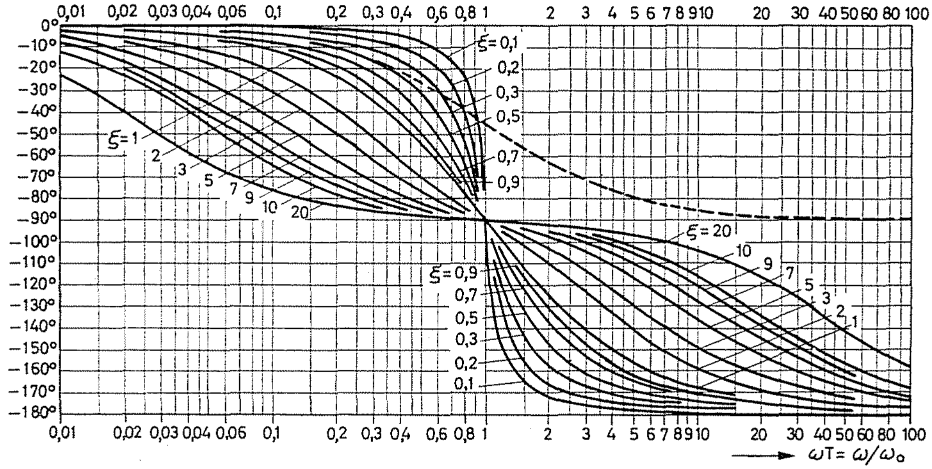

Fig. 8b. Phase plot of the second-order system $G(s)$

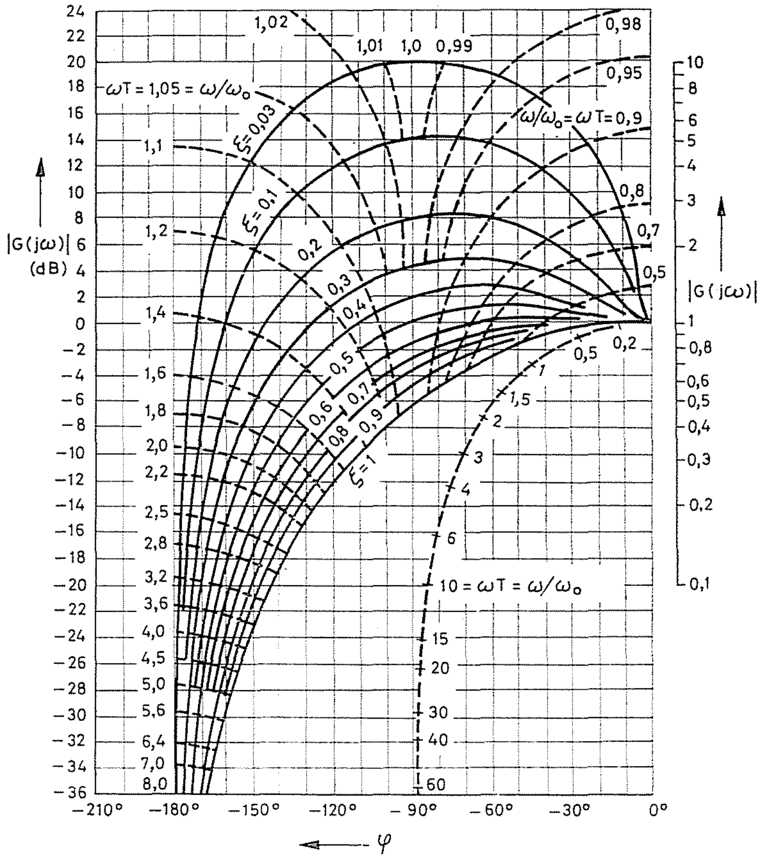

Fig. 9. Nichols plot of the second-order system

$$G(s) = \frac{1}{1 + 2\zeta Ts + T^2 s^2} = \frac{\omega_0^2}{s^2 + 2\zeta\omega_0 s + \omega_0^2}$$

At the same time the development of control theory is so vigorous that a gap between practice and theory is often spoken of. Despite all efforts, this gap seems to have widened in the last few years. The theory of control moved enough far from the adjustment rules of PID (proportional integral and derivative or rate) controllers, although in practice the latter are still prevailing.

Recent Trends in Control Theory

In the theory of control systems many recent trends can be observed. Nonlinear methods have come more and more into prominence. Some nonlinear techniques, such as the *tangential linearization*, *statistical linearization*, *harmonic balance* and *describing function* techniques have just begun in the second period but some other methods such as the *numerical methods* in the solution of nonlinear differential equations, the application of the LYAPUNOV *methods* in investigating the stability of nonlinear systems characterize the third period. All these methods, except the POPOV stability-test method, are *time-domain* methods. One of them, the phase-plane or state-plane method is very suitable for the analysis of nonlinear second-order systems. Although the efforts to generalize the geometry of the phase-plane or state-plane method into phase-space or state-space method, resp., failed, the analytic form of the state-plane method itself served as a basis for the contemporary development of the common state-space methods.

As it is well known, under fairly general conditions almost all concentrated-parameter dynamic systems can be described by the state-space vector differential equation and the algebraic auxiliary equation as follows:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}, t) \tag{1}$$

$$\mathbf{y} = \mathbf{g}(\mathbf{x}, \mathbf{u}, t) \tag{2}$$

Here $\mathbf{u}$ is the input vector or control vector, $\mathbf{y}$ the output vector whereas $\mathbf{x}$ is the state vector of the system, $\mathbf{f}$ and $\mathbf{g}$ are specified nonlinear vector functions with vector arguments. This formulation is widely used in the modern control theory, and serves as basis for the theory of optimum systems, as well as for the LYAPUNOV theory, not speaking of the adaptive or learning systems, differential games and others. For linear, time-variable concentrated-parameter systems, Eqs (1) and (2) are replaced by state-space equations

$$\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x} + \mathbf{B}(t)\mathbf{u} \tag{3}$$

$$\mathbf{y} = \mathbf{C}(t)\mathbf{x} + \mathbf{D}(t)\mathbf{u} \tag{4}$$

where $\mathbf{A}(t)$, $\mathbf{B}(t)$, $\mathbf{C}(t)$, $\mathbf{D}(t)$ are appropriate time-invariant systems, the latter matrices become constant and the state-space equations get the form

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \tag{5}$$

$$\mathbf{y} = \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{u} \tag{6}$$

In this case it is easy to show the link between the time-domain and the frequency-domain techniques. Applying the LAPLACE transformation the input-output relationship can be obtained as

$$\mathbf{y}(s) = \{\mathbf{C}[s\mathbf{I} - \mathbf{A}]^{-1}\mathbf{B} + \mathbf{D}\}\,\mathbf{u}(s) \tag{7}$$

where $\mathbf{y}(s)$ and $\mathbf{u}(s)$ denote the LAPLACE transform of $\mathbf{y} = \mathbf{y}(t)$ and $\mathbf{u} = \mathbf{u}(t)$, respectively. By the way

$$\mathbf{G}(s) = \mathbf{C}[s\mathbf{I} - \mathbf{A}]^{-1}\mathbf{B} + \mathbf{D} \tag{8}$$

is the transfer matrix of the multivariable system.

Further we are going to show some applications of the state-space method restricting ourselves in the first line on optimal systems.

Static Optimizations

In the static case the systems are characterized by algebraic equations. Well known problems arise from the operational research, such as transportation problems, allocation of resources, load distribution of electric power systems etc.

In the static time-invariant case Eq. (1) reduces to the algebraic equation

$$\mathbf{f}(\mathbf{x}, \mathbf{u}) = \mathbf{0} \tag{9}$$

If the state to be kept at the constant steady-state value $\mathbf{x} = \mathbf{x}^\circ$ then, from Eq. (9), the steady-state control vector $\mathbf{u}^\circ$ can be determined. With both $\mathbf{x} = \mathbf{x}^\circ$ and $\mathbf{u} = \mathbf{u}^\circ$ constant, the goal or objective function of the problem

$$f_0 = f_0(\mathbf{x}^\circ, \mathbf{u}^\circ) \tag{10}$$

becomes also a constant. Both $\mathbf{f}$ and f_0 are assumed as continuous functions. One of the most simple cases of static optimization is that where both Eqs. (9) and (10) are linear. This is the so-called linear programming problem. It can be stated as follows: Find non-negative values for the coordinates x_i of vector $\mathbf{x}$ minimizing (or maximizing) the linear objective function

$$f_0 = \mathbf{c}^T \mathbf{x} \tag{11}$$

subject to the linear inequality constraints

$$\mathbf{A}\mathbf{x} \leqq \mathbf{b} \tag{12}$$

where the latter inequality must be fulfilled for each coordinate of vector $\mathbf{b}$ and $\mathbf{Ax}$. (The superscript T in (11) denotes a transposition.) As it is well known, the optimal solution must lie on at least one of the corner points of the polyhedron surface formed by Eq. (12) and $\mathbf{x} \geqq \mathbf{0}$. Linear programming implies therefore a straightforward search by exchanging the extremum points, i.e. moving from one feasible solution to another. For high-dimensional problems this procedure becomes uneconomic even when using the high-speed bulk storage of modern digital computers. A more economic procedure is the well-known simplex method. Recently also decomposing techniques were developed. In this way the large problem is reduced to many small problems which could independently optimized. After the optimization of the subsystems, the total system could be optimized.

In most practical cases the goal function (11) and/or the constraint (12) depends also on stochastic disturbances (parameters). Numerous methods have been developed in order to find the optimum expectation and optimum variance of the goal function.

An interesting extension of the above formulation leads to nonlinear programming. This problem is given here only for the unforced autonomous case. A nonlinear objective function

$$f_0 = f_0(\mathbf{x}) \tag{13}$$

subject to the nonlinear inequality constraints

$$\mathbf{f}(\mathbf{x}) \leqq \mathbf{0} \tag{14}$$

and non-negativity condition $\mathbf{x} \geqq \mathbf{0}$, both applying for each coordinate, has to be minimized or maximized.

The rigorous analysis of this general optimization problem is rather questionable, therefore optimum seeking methods are often applied. If, however, $f_0(\mathbf{x})$ and $f_i(\mathbf{x})$, that is, the coordinates of $\mathbf{f}(\mathbf{x})$, are convex functions, then the so-called convex programming methods can be introduced.

Let us define the Lagrangian function

$$F(\mathbf{x}, \boldsymbol{\lambda}) = f_0(x) + \boldsymbol{\lambda}^T \mathbf{f}(\mathbf{x}) \tag{15}$$

where $\boldsymbol{\lambda}$ is the Lagrangian multiplier vector. The necessary and sufficient condition, that $\mathbf{x}^\circ$ be a solution of the convex programming problem, is that there exists a vector $\boldsymbol{\lambda}^\circ$ such that

$$\mathbf{x}^\circ \geqq \mathbf{0}, \qquad \boldsymbol{\lambda}^\circ \geqq \mathbf{0} \tag{16}$$

and

$$F(\mathbf{x}^\circ, \boldsymbol{\lambda}) \leqq F(\mathbf{x}^\circ, \boldsymbol{\lambda}^\circ) \leqq F(\mathbf{x}, \boldsymbol{\lambda}^\circ) \tag{17}$$

for all $\mathbf{x} \geqq \mathbf{0}$, $\boldsymbol{\lambda} \geqq \mathbf{0}$. Eq. (17) expresses the so-called saddle-point condition.

The arguments concerning the linear programming of large-scale systems concern even more the convex programming problems. Here decomposition and multilevel optimization techniques are also useful.

In practice the optimum search techniques have some importance. They all involve some comparison. The objective function $f_0(\mathbf{x})$ is evaluated for different allowed vectors $\mathbf{x}$ and the values of $f_0(\mathbf{x})$ so obtained are compared in order to find the optimum value. The effectivity of the various searching techniques is strongly influenced by the global and local properties of the objective and constraint functions. Sometimes penality functions are employed to transform the constraint problem into a sequence of unconstrained optimizations.

In the case of stochastic disturbances it is necessary to make experiments with the system in real time. In such circumstances hill-climbing techniques, adaptive control techniques and, generally speaking, computer process control methods have to be employed.

Dynamic Optimization

As most automatic control systems are inherently of transient character, the problem of dynamic optimization seems to be very important for the control engineer. In this case state equation (1) describes the dynamic behaviour of the system and, instead of a goal or objective function, a performance functional is to be minimized or maximized. This functional is of the general form

$$J = \int_{t_0}^{t_f} f_0(\mathbf{x}, \mathbf{u}, t)\, \mathrm{d}t + f_{00}(\mathbf{x}(t_f), t_f) \tag{18}$$

where t_0 and t_f are the initial and terminal time, respectively, f_0 is the goal or objective function whereas f_{00} is an auxiliary function. In many cases, f_{00}

can be neglected without loss of generality. In final value problems, however, f_{00} occurs, whereas the first term in Eq. (18) is neglected. Let us consider some special cases of Eq. (18) assuming $f_{00} = 0$. If $f_0 = 1$ then we have a time-optimal control problem, if $f_0 = \mathbf{u}^T \mathbf{sgn}\, \mathbf{u}$, where $\mathbf{sgn}\, \mathbf{u} = [\text{sgn}\, u_1, \ldots, \text{sgn}\, u_r]^T$ then the fuel-optimal problem arises, if $f_0 = \mathbf{u}^T\mathbf{u}$ or more generally $f_0 = \mathbf{u}^T\mathbf{R}\mathbf{u}$, then the energy-optimal control problem is stated.

One of the most frequently applied performance indices is

$$J = \frac{1}{2}\mathbf{x}^T(t_f)\,\mathbf{S}\mathbf{x}(t_f) + \frac{1}{2}\int_{t_0}^{t_f} (\mathbf{x}^T\mathbf{Q}\mathbf{x} + \mathbf{u}^T\mathbf{R}\mathbf{u})\,\mathrm{d}t \qquad (19)$$

a quadratic performance index of the form of Eq. (18). In practical systems one of the most serious problems arises from the formulation of the appropriate performance index. For this purpose no general rules can be indicated.

Frequently some (or all) of the initial and/or terminal states are not fixed *a priori* but can move along a specified starting or target set. In such cases so-called transversality (orthogonality) conditions supply the necessary boundary conditions.

In order to make the mathematical model of the optimization problem more appropriate, that is, to approach the model to physical reality, some constraints must also be taken into account. One of the most frequent constraints refers to the control vector $\mathbf{u}$ in the general form $\mathbf{u} \in U$. that is, $\mathbf{u}$ is constrained in some subspace U of the Euclidean r-dimensional space. Most often this general constraint is reduced to $|u_j| \leq 1$, $(j = 1, 2, \ldots, r)$, where u_j are the coordinates of vector $\mathbf{u}$.

Let us proceed now towards the solution of the dynamic optimization problems.

The Calculus of Variations

If the problem is unconstrained then the classical calculus of variations can be applied to the solution procedure. In this case a so-called isoperimetric problem arises and the functional to be minimized is

$$J = \int_{t_0}^{t_f} F(\mathbf{x}, \dot{\mathbf{x}}, \mathbf{u}, \boldsymbol{\lambda}, t)\,\mathrm{d}t \qquad (20)$$

where the generalized objective function F is given in the form

$$F(\mathbf{x},\dot{\mathbf{x}},\mathbf{u},\boldsymbol{\lambda},t) \triangleq f_0(\mathbf{x},\mathbf{u},t) + \boldsymbol{\lambda}^T[\mathbf{f}(\mathbf{x},\mathbf{u},t) - \dot{\mathbf{x}}] = f_0(\mathbf{x},\mathbf{u},t) + [\mathbf{f}(\mathbf{x},\mathbf{u},t) - \dot{\mathbf{x}}]^T\boldsymbol{\lambda} \quad (21)$$

with $\boldsymbol{\lambda}$ a Lagrangian multiplier vector. The *necessary* conditions of extremum are expressed by the well-known EULER—LAGRANGE equations given here in vector form:

$$\frac{\partial F}{\partial \mathbf{x}} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial F}{\partial \dot{\mathbf{x}}} = \frac{\partial f_0}{\partial \mathbf{x}} + \frac{\partial \mathbf{f}^{\mathrm{T}}}{\partial \mathbf{x}}\boldsymbol{\lambda} + \dot{\boldsymbol{\lambda}} = \mathbf{0} \tag{22}$$

$$\frac{\partial F}{\partial \mathbf{u}} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial F}{\partial \dot{\mathbf{u}}} = \frac{\partial f_0}{\partial \mathbf{u}} + \frac{\partial \mathbf{f}^{\mathrm{T}}}{\partial \mathbf{u}}\boldsymbol{\lambda} = \mathbf{0} \tag{23}$$

$$\frac{\partial F}{\partial \boldsymbol{\lambda}} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial F}{\partial \dot{\boldsymbol{\lambda}}} = \mathbf{f}(\mathbf{x}, \mathbf{u}, t) - \dot{\mathbf{x}} = \mathbf{0} \tag{24}$$

Here $d\mathbf{f}^{\mathrm{T}}/d\mathbf{x}$ and $d\mathbf{f}^{\mathrm{T}}/d\mathbf{u}$ denote Jacobian matrices. The extremal trajectory $\mathbf{x}(t)$, the extremal control vector $\mathbf{u}(t)$ and the extremal multiplicator vector $\boldsymbol{\lambda}(t)$ must fulfil Eqs. (22), (23), (24). The latter is the original state-space differential equation, whereas (22) can be regarded as an adjoint differential equation, that is, a costate equation. Introducing the Hamiltonian state function

$$H(\mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}, t) \triangleq f_0(\mathbf{x}, \mathbf{u}, t) + \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{f}(\mathbf{x}, \mathbf{u}, t) = f_0(\mathbf{x}, \mathbf{u}, t) + \mathbf{f}^{\mathrm{T}}(\mathbf{x}, \mathbf{u}, t)\boldsymbol{\lambda} \tag{25}$$

Eq. (23) can also be expressed as

$$\frac{\partial H}{\partial \mathbf{u}} = \mathbf{0} \tag{26}$$

the common *necessary* condition of the extremum of function H. It must be emphasized that even Eq. (26) will lose its validity in case of a constraint on vector $\mathbf{u}$.

The Maximum and Minimum Principles

The maximum and minimum principles can be regarded as an extension of the classical calculus of variations. Since by introducing a new coordinate $x_{n+1} = t$ with differential equation $\dot{x}_{n+1} = 1$ and zero initial condition all time-variable problems can be reduced to time-invariant ones, only the latter will be discussed here.

Let the state-space differential equation of the system be

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}) \tag{27}$$

subject to the initial condition $\mathbf{x}(t_0) = \mathbf{x}_0$ and terminal condition $x(t_f) \in C$, where C is a specified target set. The functional

$$J = \int_{t_0}^{t_f} f_0\big(\mathbf{x}(t), \mathbf{u}(t)\big)\, \mathrm{d}t \tag{28}$$

has to be minimized. Let us find the control vector $\mathbf{u} \in U$, which transfers the state of the system from $\mathbf{x}(t_0) = \mathbf{x}_0$ to $\mathbf{x}(t_f) \in C$ and minimizes the cost functional (28).

Let us now introduce the Hamiltonians

$$H_\varphi(\mathbf{x},\mathbf{u},\psi) = -f_0(\mathbf{x},\mathbf{u}) + \psi^T\mathbf{f}(\mathbf{x},\mathbf{u}) \tag{29}$$

and

$$H_p(\mathbf{x},\mathbf{u},\mathbf{p}) = f_0(\mathbf{x},\mathbf{u}) + \mathbf{p}^T\mathbf{f}(\mathbf{x},\mathbf{u}) \tag{30}$$

Comparing (29) and (30) with (25) we can recognize that on the one hand

$$\mathbf{p}(t) = -\psi(t) = \lambda(t) \tag{31}$$

and on the other

$$H_p = -H_\varphi = H. \tag{32}$$

Let $\mathbf{x}^\circ(t)$ denote the solution of differential equation (27) for the optimal control $\mathbf{u}^\circ(t)$. Then, corresponding to $\mathbf{u}^\circ(t)$ and $\mathbf{x}^\circ(t)$ a costate vector ψ° or $\mathbf{p}^\circ$ exists such that with $H_\varphi^\circ = H_\varphi(\mathbf{x}^\circ,\mathbf{u}^\circ,\psi^\circ)$ and $H_p^\circ = H_p(\mathbf{x}^\circ,\mathbf{u}^\circ,\mathbf{p}^\circ)$ the canonical equations

$$\dot{\mathbf{x}}^\circ = \frac{\partial H_\varphi^\circ}{\partial \psi^\circ} \qquad \text{or} \qquad \dot{\mathbf{x}}^\circ = \frac{\partial H_p^\circ}{\partial \mathbf{p}^\circ} \tag{33}$$

$$\dot{\psi}^\circ = -\frac{\partial H^\circ}{\partial \mathbf{x}^\circ} \qquad \text{or} \qquad \dot{\mathbf{p}}^\circ = -\frac{\partial H_p^\circ}{\partial \mathbf{x}^\circ} \tag{34}$$

hold, subject to the boundary conditions:

$$\mathbf{x}^\circ(t_0) = \mathbf{x}_0, \ \mathbf{x}^\circ(t_f) \in C$$

or, according to the transversality condition, $\psi^\circ(t_f)$ or $\mathbf{p}^\circ(t_f)$ must be normal to C at $\mathbf{x}^\circ(t_f)$. Then the *necessary* condition of optimality can be expressed as

$$H_\varphi(\mathbf{x}^\circ,\mathbf{u}^\circ,\psi^\circ) \geqq H_\varphi(\mathbf{x}^\circ,\mathbf{u},\psi^\circ) \tag{35}$$

or

$$H_p(\mathbf{x}^\circ, \mathbf{u}^\circ, \mathbf{p}^\circ) \leqq H_p(\mathbf{x}^\circ, \mathbf{u}, \mathbf{p}^\circ) \tag{36}$$

or every t in the interval $t_0 \leqq t \leqq t_f$ and for all $\mathbf{u} \in U$.

Eq. (35) states the maximum whereas Eq. (36) the minimum principle. Both are equivalent and the choice depends on preference. PONTRYAGIN originally formulated his theory as maximum principle, the minimum principle has, however, a closer relation to other theorems, therefore we will refer to the latter in the following development.

In applying the minimum principle to the determination of the optimal control $\mathbf{u}^\circ(t)$, the following steps are to be taken. We start from relationship (36) in order to deduce a relationship

$$\mathbf{u}^\circ = \mathbf{g}(\mathbf{x}^\circ, \mathbf{p}^\circ). \tag{37}$$

If $\mathbf{x}^\circ(t)$ and $\mathbf{p}^\circ(t)$ uniquely specify $\mathbf{u}^\circ(t)$ in the whole interval $[t_0, t_f]$ then we have a so-called *normal* problem, otherwise we have a *singular* problem. Restricting ourselves to the former case and substituting Eq. (37) into (33) and (34), the latter will depend only on $\mathbf{x}^\circ(t)$ and $\mathbf{p}^\circ(t)$. Then two vector equations are at our disposal for the solution of the state vector $\mathbf{x}^\circ(t)$ and the costate vector $\mathbf{p}^\circ(t)$. The solution leads, however, generally to a two-point boundary-value problem. The initial and terminal conditions $\mathbf{x}^\circ(t_0)$ and $\mathbf{x}^\circ(t_f)$ or the transversality conditions, supplying the coordinates of $\mathbf{p}^\circ(t_f)$, give $2n$ boundary conditions for the solution of the $2n$ scalar equations (33) and (34). However, n of them refer to the initial states and n of them to the final states or costates. In any case, after finding the optimal state vector $\mathbf{x}^\circ(t)$ and costate vector $\mathbf{p}^\circ(t)$ also the optimal control vector $\mathbf{u}^\circ(t)$ is delivered by Eq. (37). From the outlined procedure it becomes obvious that no analytical solution can be expected but for the most simple cases. This is one of the main reasons why numerical solutions are of so a great significance.

The Discrete Minimum Principle

The discrete maximum or minimum principle has some advantages over the continuous ones, at least from computation aspects, as they are in a form directly suitable for digital computation. We may discretize the optimum problem from the beginning and replace the differential equations by difference equations. It must be mentioned, however, that the intuitive insight into the character and structure of the optimum problem gets somewhat lost by the discretization.

Let us describe the dynamic system by the vector difference equation

$$\mathbf{x}_{k+1} - \mathbf{x}_k = \mathbf{f}_k(\mathbf{x}_k, \mathbf{u}_k) \tag{38}$$
$$(k = 0, 1, 2, \ldots, K-1)$$

where $\mathbf{x}_k$ is the value of the state vector at the k-th sampling instant, $\mathbf{u}_k$ the value of the control vector at the same moment, and $\mathbf{f}_k$ a vector function of vector arguments $\mathbf{x}_k$ and $\mathbf{u}_k$. Given the control constraint

$$\mathbf{u}_k \in U \quad \text{for all} \quad k = 0, 1, 2, \ldots, K-1 \tag{39}$$

and the cost functional in the form

$$J = \sum_{k=0}^{K-1} f_{0k}(\mathbf{x}_k, \mathbf{u}_k) \tag{40}$$

where $f_{0k}(\mathbf{x}_k, \mathbf{u}_k)$ is a scalar objective function of the vector arguments $\mathbf{x}_k$ and $\mathbf{u}_k$, let us suppose that the boundary conditions are

$$\mathbf{x}_0 = \alpha \quad \text{and} \quad \mathbf{x}_k \in C \tag{41}$$

where C is some specified target set in the n-dimensional Euclidean state space.

Our aim is to find the optimal control sequence $\mathbf{u}_0^\circ, \mathbf{u}_1{}^\circ, \ldots, \mathbf{u}_{k-1}^\circ$ satisfying the constraint (39) such that for the generated state sequence $\mathbf{x}_0^\circ, \mathbf{x}_1{}^\circ, \ldots, \mathbf{x}_{k-1}^\circ$, subject to the boundary conditions (41), the cost functional (40) is at its minimum. Let us define the Hamiltonian

$$H_k(\mathbf{x}_k, \mathbf{u}_k, \mathbf{p}_{k+1}) = f_{0k}(\mathbf{x}_k, \mathbf{u}_k) + \mathbf{p}_{k+1}^T \mathbf{f}_k(\mathbf{x}_k, \mathbf{u}_k) \tag{42}$$

for $k = 0, 1, 2, \ldots, K-1$.

Corresponding to the optimal state sequence $\mathbf{x}_k^\circ$ and control sequence $\mathbf{u}_k^\circ$, $(k = 0, 1, 2, \ldots, K-1)$ a costate vector sequence $\mathbf{p}_k^\circ$, $(k = 0, 1, 2, \ldots, K)$ exists such that with $H_k^\circ = H_k(\mathbf{x}_k^\circ, \mathbf{u}_k^\circ, \mathbf{p}_{k+1})$ the canonical difference equations

$$\mathbf{x}_{k+1}^\circ - \mathbf{x}_k^\circ = \frac{\partial H_k^\circ}{\partial \mathbf{p}_{k+1}^\circ} \tag{43}$$

$$\mathbf{p}_{k+1}^\circ - \mathbf{p}_k^\circ = -\frac{\partial H_k^\circ}{\partial \mathbf{x}_k^\circ} \tag{44}$$

hold, subject to the boundary conditions $\mathbf{x}_0^\circ = \alpha$; $\mathbf{x}_K^\circ \in C$, $\mathbf{p}_K^\circ$ normal to C at $\mathbf{x}_K^\circ$. Then the *necessary* condition of optimality can be expressed as

$$H_k(\mathbf{x}_k^\circ, \mathbf{u}_k^\circ, \mathbf{p}_{k+1}^\circ) \leqq H_k(\mathbf{x}_k^\circ, \mathbf{u}_k, \mathbf{p}_{k+1}^\circ) \tag{45}$$

for all $\mathbf{u}_k \in U$ and all $k = 0, 1, 2, \ldots, K-1$. The comparison between the continuous and the discrete minimum principles shows a close analogy.

Dynamic Programming

According to the principle of optimality, the final part of an optimal trajectory in itself is also an optimal trajectory. Based on this principle Bellman developed the dynamic programming method. The discrete form of this method is essentially a multistage decision process, where x_k denotes the state vector and $\mathbf{u}_k$ denotes the control vector or in other words the decision or policy vector. The optimal policy minimizes the cost functional

$$J = f_{00}(\mathbf{x}_K) + \sum_{k=0}^{K-1} f_0(\mathbf{x}_0, \mathbf{u}_0). \tag{45}$$

Denoting by S_{K-k} the optimal value of the partial sum I_{K-k}, that is,

$$S_{K-k} = \min_{\mathbf{u}_{K-k} \in U} I_{K-k} \tag{46}$$

where

$$I_{K-k} = I_{K-k+1} + f_0(\mathbf{x}_{K-k}, \mathbf{u}_{K-k}) \tag{47}$$

and applying the principle of optimality, the following recurrent relationship can be obtained:

$$S_{K-k}(\mathbf{x}^\circ_{K-k}) = \min_{\mathbf{u}_{K-k} \in U} \{S_{K-k+1}(\mathbf{x}^\circ_{K-k} + \mathbf{f}(\mathbf{x}^\circ_{K-k}, \mathbf{u}_{K-k})) + f_0(\mathbf{x}^\circ_{K-k}, \mathbf{u}_{K-k})\} \tag{48}$$

As a result of the minimization procedure from (48) the optimal value $\mathbf{u}^\circ_{K-k}$ of $\mathbf{u}_{K-k}$ can be computed. By iterating (48) and the minimization procedure the whole control sequence $\mathbf{u}^\circ_{K-1}$, $\mathbf{u}^\circ_{K-2}, \ldots, \mathbf{u}_1$, $\mathbf{u}_0$ is obtained as a final result.

In the continuous, time-variable case the performance functional to be minimized can be expressed as

$$J = \int_{t_0}^{t_f} f_0(\mathbf{x}(t), \mathbf{u}(t), t)\, \mathrm{d}t. \tag{49}$$

Let us denote by $S(\mathbf{x}^\circ(t), t)$ the minimum value of the partial functional owing o the trajectory segment starting from $\mathbf{x}^\circ(t)$, that is

$$S(\mathbf{x}^\circ(t), t) = \min_{\mathbf{u} \in U} \int_{t_0}^{t_f} f_0(\mathbf{x}^\circ(\vartheta), \mathbf{u}^\circ(\vartheta), \vartheta)\, \mathrm{d}\vartheta. \tag{50}$$

Then, the fundamental equation of dynamic programming, the so-called Hamilton—Jacobi—Bellman equation can be derived as

$$-\frac{\partial S(\mathbf{x}^\circ(t), t)}{\partial t} = \min_{\mathbf{u} \in U} \left\{ \frac{\partial S(\mathbf{x}^\circ(t), t)}{\partial \mathbf{x}^{\circ T}} \mathbf{f}(\mathbf{x}^\circ(t), \mathbf{u}(t), t) + f_0(\mathbf{x}^\circ(t), \mathbf{u}(t), t) \right\}. \tag{51}$$

After the minimization procedure the partial differential equation of Hamilton—Jacobi can be obtained:

$$\frac{\partial S(\mathbf{x}^\circ(t), t)}{\partial t} + \frac{\partial S(\mathbf{x}^\circ(t), t)}{\partial \mathbf{x}^{\circ T}} f(\mathbf{x}^\circ(t), \mathbf{u}^\circ(t), t) + f_0(\mathbf{x}^\circ(t), \mathbf{u}^\circ(t), t) = 0. \tag{52}$$

In the time-invariant case the left-hand side term of (51) becomes zero and after introducing

$$\mathbf{p}^{\circ} = \left[\frac{\partial S}{\partial x_1^{\circ}}, \ \frac{\partial S}{\partial x_2^{\circ}}, \ldots, \frac{\partial S}{\partial x_n^{\circ}} \right]^{\mathrm{T}} = -\psi^{\circ} \tag{53}$$

we obtain from (51):

$$0 = \min_{\mathbf{u} \in U} \{ f_0(\mathbf{x}^{\circ}, \mathbf{u}) + \mathbf{p}^{\circ \mathrm{T}} \mathbf{f}(\mathbf{x}^{\circ}, \mathbf{u}) \} \tag{54}$$

which is just the PONTRYAGIN minimum principle (see Eqs (30) and (36)). As an additive result, in time-invariant, free-terminal-time problems the minimum value of the Hamiltonian is seen to become zero. In a quite similar way the maximum principle may also be stated. Furthermore, from Eq. (52) the EULER—LAGRANGE equation (22) of calculus of variations can also be derived. Thus, dynamic programming seems to be the most general optimization method.

Computational Difficulties

The calculus of variations, the maximum or minimum principle reduces the optimal control problem to a two-point boundary-value problem involving vector differential or difference equations. To the solution of the latter, in general, numerical methods must be applied. The amount of numerical computations require the use of relatively large-scale digital computers. That is why variational methods have found so little application in applied science, for example, in control engineering. But even in case of an appropriate computer it is a difficult problem to choose the right fast-converging algorithm for a certain type of optimization. The difficulties associated with two-point boundary-value problems appear at the first sight avoided by the discrete dynamic programming. This method, however, necessitates bulk storage capacities in the memory of computers.

The above-mentioned problems restrict the application field of optimization methods. In space research they are perhaps inevitable but in process control they are seldom used to now.

Computers in Process Control

A recent trend in control engineering is the application of relatively small-size computers for process control purposes. Without speaking of supervision, recording and reporting duties we mention here only one aspect: the direct digital control (DDC). This method facilitates to apply a digital computer to perform the control calculations and, with suitable connections to the process equipment, eliminates much of the usual analog devices. Therefore, DDC seems to have great potentialities.

As it is well known, a common PID (proportional-integral-derivative or rate) controller operates after the following rule:

$$u = K_0 + K_P e + K_P K_I \int_0^t e \, dt + K_P K_D \frac{de}{dt} . \quad (55)$$

The corresponding DDC algorithm can be expressed for the simplest case as

$$u_k = K_0 + K_P e_k + K_P K_I \sum_{j=0}^{j=k} e_j \Delta t + K_P K_D \frac{e_k - e_{k-1}}{\Delta t} . \quad (56)$$

Direct digital control offers an opportunity to apply new theoretical solutions in control problems. It must be emphasized, however, that the application of DDC depends mostly on the reliability of computer equipment and a high MTBF (mean time before failure) of about ten thousand hours or more must be guaranteed.

Some Concluding Remarks

The advent of the state-space methods and the wide-spread application of digital computers is in close relation to each other at least in the field of control engineering. On the one hand, the digital computations required the introduction and application of the state-space methods and, on the other hand, the problems formulated in this form, even for moderately complicated problems, can only be solved by the use of a digital computer.

The modern trends of control engineering, therefore, can indeed be characterized by the abundant application of digital computers and the state-space method.

In the limited scope of this paper it was impossible to treat all the topics of modern control theory and practice. For example, the *optimizing*, *adaptive* and *learning* systems become more and more important. The same can be stated in connection with the theory of large-scale systems, where the problems of *hierarchy*, *decomposition*, and *reorganization* arise. All the problems mentioned can only be solved by appropriate algorithm and structures.

These problem formulations are concerned not only with technological processes but also with economic, social, biological and physiological systems. In this way, various fields of science approach each other, control theory and practice become of interdisciplinary character. At the same time, much is learned from the various branches of science. The theory of control develops towards generalizations and integrations whereas the practice of control seems to be more specific.

Summary

This paper gives a brief survey of previous and recent trends of control theory and practice. After reviewing the three historical epochs of control engineering, some recent trends as, for example, state-space methods, static and dynamic optimization techniques (linear programming, convex programming, calculus of variations, maximum and minimum principles, dynamic programming) are treated. The fundamentals of process control computer applications and some other topics are also mentioned.

References

1. Akhiezer, N. I.: The Calculus of Variations. Blaisdell, 1962
2. Athans, M.—Falb, P.: Optimal Control. McGraw-Hill, 1966
3. Aoki, M.: Optimization of Stochastic Systems. Academic Press, 1967
4. Bellman, R.: Dynamic Programming. Princeton University Press, 1957
5. Bellman, R.—Glicksberg, I.—Gross, O.: Some Aspects of the Mathematical Theory of Control Processes. Rand Corp., 1958
6. Bellman, R.: Adaptive Control Processes: A Guided Tour. Princeton University Press, 1961
7. Bellman, R.—Dreyfus, S.: Applied Dynamic Programming. Princeton University Press, 1962.
8. Bellman, R.: Mathematical Optimization Techniques. University of California Press, 1963
9. Balakrishnan, A.—Neustadt, L.: Computing Methods in Optimization Problems. Academic Press, 1964
10. Bellman, R.—Kalaba, R.: Mathematical Trends in Control Theory. Dover, 1964
11. Bellman, R.—Kalaba, R.: Dynamic Programming and Modern Control Theory. Academic Press, 1966
12. Bellman, R.: Introduction to the Mathematical Theory of Control Processes. Academic Press, 1967
13. Bailey, J.: Nonlinear Two-Point Boundary Value Problems. Academic Press, 1968
14. Cunningham, W.: Introduction to Non-linear Analysis. McGraw-Hill, 1958
15. Coughanowr, D.—Koppel, L.: Process Systems Analysis and Control. McGraw-Hill, 1965
16. Chu, Y.: Digital Simulation of Continuous Systems. McGraw-Hill, 1969
17. Csáki, F.: Korszerű szabályozáselmélet. Nemlineáris, optimális és adaptív rendszerek. Akadémiai Kiadó, 1970
18. Csáki, F.: Szabályozások dinamikája. Lineáris szabályozáselmélet. 2. kiadás, Akadémiai Kiadó, 1970
19. Csáki, F.—Bars, R.: Automatika. 2. kiadás, Tankönyvkiadó, 1972
20. Csáki, F.: Modern Control Theories. Akadémiai Kiadó, 1972
21. Dantzig, G.: Linear Programming and Extensions. Princeton University Press, 1963
22. Dorf, R.: Time Domain Analysis and Design of Control Systems. Addison-Wesley, 1965
23. Dreyfus, S.: Dynamic Programming and the Calculus of Variations. Academic Press, 1965
24. DeRusso, P.—Roy, R.—Close, Ch.: State Variables for Engineers. John Wiley, 1966
25. Elgerd, O.: Control Systems Theory. McGraw-Hill, 1967
26. Eveleigh, V.: Adaptive Control and Optimization Techniques. McGraw-Hill, 1967
27. Fan, L.: The Discrete Maximum Principle. John Wiley, 1964
28. Feldbaum, A.: Optimal Control Systems. Academic Press, 1965
29. Fan, L.: The Continuous Maximum Principle. John Wiley, 1966
30. Forray, M.: Variational Calculus in Science and Engineering. McGraw-Hill, 1968
31. Gantmacher, F.: Theory of Matrices. Chelsea, 1959
32. Graham, D.—McRuer, D.: Analysis of Nonlinear Control Systems. John Wiley, 1961
33. Gupta, S.: Transform and State Variable Methods in Linear Systems. John Wiley, 1966
34. Goldberg, S.: Differenzengleichungen und ihre Anwendung in Wirtschaftswissenschaft. Psychologie und Soziologie. Oldenbourg, 1968
35. Higgins, T.—Naslin, P.: Progress in Control Engineering. Academic Press, 1964
36. Hestenes, M.: Calculus of Variations and Optimal Control Theory. John Wiley, 1966
37. Himmelblau, D.—Bischoff, K.: Process Analysis and Simulation. John Wiley, 1968
38. Kipiniak, W.: Dynamic Optimization and Control. John Wiley, 1961

39. Koenig, H.—Blackwell, W.: Electromechanical System Theory. McGraw-Hill, 1961
40. Liapunov, A.: On the General Problem of Stability of Motion. Ph. D. thesis, Kharkov, 1892; Reprinted in French in Ann. of Mathematics Studies, No. 17. Princeton University Press, 1949
41. Lur'e, A.: Some Nonlinear Problems in the Theory of Automatic Control. Her Majesty's Stationary Office, London, 1957
42. LaSalle, J.—Lefschetz, S.: Stability by Liapunov's Direct Method with Applications. Academic Press, 1961
43. Leitman, G.: Optimization Techniques with Applications to Aerospace Systems. Academic Press, 1962
44. Leondes, C.: Advances in Control Systems. Vol. 1—8, Academic Press, 1964—1971
45. Lapidus, L.—Luus, R.: Optimal Control of Engineering Processes. Blaisdell Publ., 1967
46. Lee, E.—Markus, L.: Foundations of Optimal Control Theory. John Wiley, 1967
47. Malkin, I.: Theory of Stability of Motion. U. S. Department of Commerce, Washington, 1958
48. Mishkin, E.—Braun, L.: Adaptive Control Systems. McGraw-Hill, 1961
49. Merriam, Ch.: Optimization Theory and the Design of Feedback Control Systems. McGraw-Hill, 1964
50. Meisel, J.: Principles of Electromechanical Energy Conversion. McGraw-Hill, 1966
51. Mayr, O.: Zur Frühgeschichte der technischen Regelungen. Oldenbourg, 1969
52. Pontryagin, L.—Boltyanski, V.—Gamkrelidze, R.—Miscsenko, E.: The Mathematical Theory of Optimal Processes. John Wiley, 1962
53. Pontryagin, L.: Ordinary Differential Equations. Addison-Wesley Publ., 1962
54. Ralston, A.—Wilf, H.: Mathematical Methods for Digital Computers. John Wiley, 1966
55. Scarborough, J.: Numerical Mathematical Analysis. John Hopkins Press, 1950
56. Saaty, Th.: Nonlinear Mathematics. McGraw-Hill, 1964
57. Savas, E.: Computer Control of Industrial Processes. McGraw-Hill, 1965
58. Schwartz, R.—Friedland, B.: Linear Systems. McGraw-Hill, 1965
59. Sworder, D.: Optimal Adaptive Control Systems. Academic Press, 1966
60. Schultz, D.—Melsa, J.: State Functions and Linear Control Systems. McGraw-Hill, 1967
61. Schwarz, H.: Mehrfachregelungen. Springer, 1967
62. Siljak, D.: Nonlinear Systems: the Parameter Analysis and Design. John Wiley, 1969
63. Tou, J.: Modern Control Theory. McGraw-Hill, 1964
64. Tomovic, R.: Introduction to Nonlinear Automatic Control Systems. John Wiley, 1966
65. Timothy, L.—Bona, B.: State Space Analysis. McGraw-Hill, 1968
66. Vajda, S.: Mathematical Programming. Addison-Wesley, 1961
67. Westcott, J.: An Exposition of Adaptive Control. Pergamon Press, 1962
68. Wilde, D.: Optimum Seeking Methods. 2. princ., Englewood Cliffs, 1965
69. Zubov, V.: The Methods of A. M. Lyapunov and their Applications. USAEC, 1961
70. Zubov, V.: Mathematical Methods of the Study of Automatic Control. Academic Press, 1962.
71. Zadeh, L.—Desoer, Ch.: Linear System Theory. The State Space Approach. McGraw-Hill, 1963
72. Zypkin, Ja.: Adaptation und Lernen in kybernetischen Systemen. Verlag Technik, 1970

Prof. Dr. Frigyes Csáki, 1502 Budapest P.O.B. 91. Hungary