KYBERNETIKA — VOLUME 9 (1973), NUMBER 2

Algebraic Theory of Discrete Optimal Control for Single-variable Systems I

Preliminaries

VLADIMÍR KUČERA

A new unifying approach to the optimal control of discrete linear constant systems is proposed. The approach is based exclusively on algebraic properties of polynomials and is believed to be conceptually simpler and computationally superior to existing methods. Moreover, it applies to systems over an arbitrary field.

The whole paper is divided into three parts appearing separately. Part I (Preliminaries) establishes some basic results regarding polynomials and Diophantine equations. It also gives a rigorous definition of the systems to be investigated.

Part II (Open-Loop Control) is a systematic treatment of the theory of optimal open-loop control problems. Two time optimal criteria as well as the least squares problem are discussed. At the end the effect of disturbances upon the open-loop control and some computational aspects are briefly mentioned.

Part III (Closed-Loop Control) is devoted to the most common control scheme. The same criteria are discussed together with the pole assignment problem and a comparison is made with the open-loop control. Again, the effect of disturbances on the optimal system performance is considered.

The theory to be developed in the tripaper applies to single-variable systems only. A natural generalization to multivariable systems will be considered in a future paper.

INTRODUCTION

There are two principal schemes used to solve control problems, namely, the open-loop and the closed-loop configurations.

The open-loop optimal control problem consists in the following. Given a system $\boldsymbol{s}$ generate a control $\boldsymbol{u}$ which causes the system output $\boldsymbol{y}$ to follow a given reference signal $\boldsymbol{w}$ in a prescribed way. This configuration is shown in Fig. 1. We point out that this control is of feedforward type, i.e., no attempt is made to neutralize the effect of disturbances.

In contrast, the closed-loop optimal control problem considered here is the following. Given a system $\boldsymbol{s}$, find such a controller $\boldsymbol{r}$ feeded by the error signal $\boldsymbol{e}$ that the output y of the system follows a given reference signal w in a prescribed manner (see Fig. 2). This configuration is of feedback type, i.e., it counteracts possible disturbances in the control loop.

The open-loop control problem is the simplest and basic one and we dispose of it in the forthcomming Part II. The closed-loop control problem will be fully discussed in Part III of the tripaper.

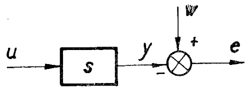

Fig. 1. Open-Loop Control System.

Complexity of these problems depends heavily upon the prescribed optimality criterion. There are two basic criteria which make the problem treatable for linear discrete systems, namely, the time optimal control problem and the least squares control problem. Loosely speaking, in the former problem we are to zero e and possibly u as fast as possible, while in the latter problem we are to minimize a quadratic functional involving e and possibly u.

In either problem we have to ensure stability for e and u. Otherwise the results would be of limited engineering relevance.

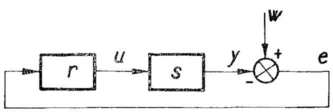

Fig. 2. Closed-Loop Control System.

The current trend in solving the above problems is to use either frequency-domain (z-transform) or time-domain (state space) approach. The former approach [1], [8], [9] transforms the essentially time-domain problem into the language of functions of a complex variable. This simplifies and visualizes the manipulations but requires a rather advanced mathematical tool (the z-transform theory, contour integration, the residue theorem, the theory of analytic functions, etc.) to lend mathematical respectability to those methods. Moreover, we are not able to give a rigorous definition of a system within this framework since we are confined to input-output properties. Further, this theory does not apply to finite automata.

On the other hand, the latter approach [3], [4], [7] introduces the idea of state, thus making an exact definition of a system possible. It works solely in the time domain and profits from the theory of differential equations in matrix form. Finite automata are accounted for. However, a control engineer may be disappointed. The state of a system is an abstract entity and frequently not accessible in a real system. Another objection involves computational aspects since the matrices often convey a good deal of superfluous structural information.

In contrast to both methods we can characterize the approach presented here as an algebraic one. It reflects the most recent trend in linear system theory. It combines the advantages of both previous methods, namely, it is conceptually simple, requires no advanced mathematical machinery (the optimal problems are solved even without any appeal to the calculus of variations), applies to discrete linear constant systems over an arbitrary field, and finally it yields effective and unified computational algorithms.

This approach introduces the concept of state just to define a system. For the control purposes we work only with the input–output responses viewed as abstract polynomials or formal power series. The synthesis procedure for all problems consists in solving a simple Diophantine equation in polynomials, whereby reducing mathematical complexity as much as possible and forming a neat and coherent whole.

POLYNOMIALS AND FORMAL POWER SERIES

We first introduce several modern algebraic notions [4], [5], [10].

A set $\mathfrak{G}$ in which two laws of composition are given, the first written additively and the second multiplicatively, is called a (commutative) *ring* if the following axioms hold.

A_0 (Consistency):	$a, b \in \mathfrak{G}$ implies $a + b \in \mathfrak{G}$.
A_1 (Associativity):	$a, b, c \in \mathfrak{G}$ implies $a + (b + c) = (a + b) + c$.
A_2 (Commutativity):	$a, b \in \mathfrak{G}$ implies $a + b = b + a$.
A_3 (Zero element):	$a \in \mathfrak{G}$, there exists $0 \in \mathfrak{G}$ such that $0 + a = a$.
A_4 (Additive inverse):	$a \in \mathfrak{G}$, there exists $-a \in \mathfrak{G}$ such that $-a + a = 0$.
M_0 (Consistency):	$a, b \in \mathfrak{G}$ implies $ab \in \mathfrak{G}$.
M_1 (Associativity):	$a, b, c \in \mathfrak{G}$ implies $a(bc) = (ab)\, c$.
M_2 (Commutativity):	$a, b \in \mathfrak{G}$ implies $ab = ba$.
M_3 (Identity element):	$a \in \mathfrak{G}$, there exists $1 \in \mathfrak{G}$ such that $1a = a$.
D (Distributivity):	$a, b, c \in \mathfrak{G}$ implies $a(b + c) = ab + ac$.

If an element $e \in \mathfrak{G}$ has a multiplicative inverse, we call e a *unit* of $\mathfrak{G}$. If every nonzero element of $\mathfrak{G}$ has a multiplicative inverse,

M_4 (Multiplicative inverse): $a \in \mathfrak{G}$, $a \neq 0$, there exists $a^{-1} \in \mathfrak{G}$ such that $a^{-1}a = 1$,

then $\mathfrak{G}$ is called a *field*.

For example, the set $\mathfrak{Z}$ of integers constitute a ring, while the rationals $\mathfrak{Q}$, reals $\mathfrak{R}$ and complex numbers $\mathfrak{C}$ all form fields. The set $\mathfrak{Z}_p$ of residue classes of integers modulo a prime is an example of a finite field.

It is seen that division is the only nontrivial operation in a ring. If $a, b \in \mathfrak{G}$, $b \neq 0$, we say that b *divides* a, and write $b \mid a$, if there exists a $c \in \mathfrak{G}$ such that $a = bc$.

For $a, b \in \mathfrak{G}$, a *greatest common divisor* of a and b is an element $d \in \mathfrak{G}$, denoted by (a, b), which is defined as follows:

(i) $$d \mid a\,, \quad d \mid b\,,$$

(ii) $$c \in \mathfrak{G}\,, \quad c \mid a\,, \quad c \mid b \ \text{ implies } \ c \mid d\,.$$

The greatest common divisor is uniquely determined up to units in $\mathfrak{G}$.

If $(a, b) = 1$ modulo units of $\mathfrak{G}$, the elements $a, b \in \mathfrak{G}$ are said to be *relatively prime.*

Given a field $\mathfrak{F}$, we shall consider sequences

$$a = \{\alpha_0, \alpha_1, \ldots, \alpha_n\}\,, \quad \alpha_k \in \mathfrak{F}\,, \; n < \infty\,.$$

If $\alpha_n \neq 0$, then n is the degree of a denoted as ∂a. We define $\partial a = -1$ for $a = 0$.

If a and b,

$$b = \{\beta_0, \beta_1, \ldots, \beta_m\}\,, \quad \beta_k \in \mathfrak{F}\,, \; m < \infty\,,$$

belong to the set of all such sequences, we define

(1) $$a + b = \{\gamma_0, \gamma_1, \ldots\}\,,$$

(2) $$ab \quad = \{\delta_0, \delta_1, \ldots\}$$

where

$$\gamma_k = \alpha_k + \beta_k\,, \quad \delta_k = \sum_{i+j=k} \alpha_i \beta_j\,.$$

Then the set becomes a ring.

Define

$$\zeta = \{0, 1, 0, \ldots, 0\}\,,$$

then

$$\zeta^k = \{0, \ldots, 0, 1, 0, \ldots, 0\}$$

with 1 at the k-th position, and we can write

$$a = \alpha_0 + \alpha_1\zeta + \ldots + \alpha_n\zeta^n\,.$$

That is why this ring is referred to as the ring of *polynomials over* $\mathfrak{F}$ *in the indeterminate* ζ and will be denoted by $\mathfrak{F}[\zeta]$.

The abstract algebraic construction above is to emphasize that we regard a polynomial as an algebraic object, not as a function of a complex variable. A polynomial is simply an alternate way of viewing finite sequences in $\mathfrak{F}$, the indeterminate ζ being a position-marker.

The units of $\mathfrak{F}[\zeta]$ are polynomials of zero degree, which are viewed as isomorphic with $\mathfrak{F}$.

Similarly, a *formal power series over $\mathfrak{F}$ in the indeterminate* ζ is an infinite sequence

$$a = \{\alpha_0, \alpha_1, \alpha_2, \ldots\}, \quad \alpha_k \in \mathfrak{F},$$

which we denote formally as

$$a = \alpha_0 + \alpha_1\zeta + \alpha_2\zeta^2 + \ldots.$$

Formal power series are a ring, denoted by $\mathfrak{F}[[\zeta]]$, if we define addition and multiplication as in (1) and (2). Observe that (2) is equivalent to convolution of sequences.

It is also clear that $\mathfrak{F}[\zeta]$ is a subring of $\mathfrak{F}[[\zeta]]$.

The greatest common divisor of two polynomials $a, b \in \mathfrak{F}[\zeta]$ is effectively determined through the euclidean algorithm. Write

(3)
$$\begin{aligned} a &= q_1 b + r_1, \quad \partial r_1 < \partial b, \\ b &= q_2 r_1 + r_2, \quad \partial r_2 < \partial r_1, \\ r_1 &= q_3 r_2 + r_3, \quad \partial r_3 < \partial r_2, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ r_{n-2} &= q_n r_{n-1}. \end{aligned}$$

We note that all q_i and r_i are uniquely determined and the algorithm terminates when $r_n = 0$.

Then $(a, b) = r_{n-1}$ to within a unit of $\mathfrak{F}[\zeta]$.

DIOPHANTINE EQUATIONS IN POLYNOMIALS

Consider the equation

(4)
$$ax + by = c$$

for unknowns $x, y \in \mathfrak{F}[\zeta]$ and given polynomials $a, b, c \in \mathfrak{F}[\zeta]$. By analogy with a similar equation in integers [2], [6], we shall call (4) a *Diophantine equation.*

Any pair x, y satisfying (4) will be called a solution. A particular solution of (4) can be found by guessing the degrees of x and y and equating the coefficients of like powers of ζ. However, this is ineffective and not very appealing. The solvability criterion as well as the general solution is given below.

Theorem 1. *Equation* (4) *has a solution if and only if* $(a, b) \mid c$.
If x_0, y_0 *is a particular solution of* (4), *then all solutions are of the form*

(5)
$$x = x_0 + \frac{b}{(a, b)} t,$$

$$y = y_0 - \frac{a}{(a, b)} t ,$$

where t is an arbitrary polynomial of $\mathfrak{F}[\zeta]$.

We can obtain

$$(6) \qquad x_0 = (-1)^n z_{n-1} \frac{c}{r_{n-1}} , \quad \frac{b}{(a, b)} = z_n ,$$

$$y_0 = (-1)^{n-1} w_{n-1} \frac{c}{r_{n-1}} , \quad \frac{a}{(a, b)} = w_n ,$$

where w_{n-1}, w_n and z_{n-1}, z_n are given via the recurrent equations

$$(7) \qquad w_0 = 1 , \quad w_1 = q_1 , \quad w_k = q_k w_{k-1} + w_{k-2} ,$$

$$z_0 = 0 , \quad z_1 = 1 , \quad z_k = q_k z_{k-1} + z_{k-2} ,$$

$$k = 2, 3, \ldots, n ;$$

the $q_1, q_2, \ldots, q_n$ and r_{n-1} come from the euclidean algorithm (3) for (a, b).

Proof. We rewrite (4) into the matrix form

$$[a, b] \begin{bmatrix} x \\ y \end{bmatrix} = c .$$

By the euclidean algorithm,

$$[a, b] \begin{bmatrix} 1, & 0 \\ -q_1, & 1 \end{bmatrix} = [r_1, b] ,$$

$$[r_1, b] \begin{bmatrix} 1, & -q_2 \\ 0, & 1 \end{bmatrix} = [r_1, r_2] ,$$

$$[r_1, r_2] \begin{bmatrix} 1, & 0 \\ -q_3, & 1 \end{bmatrix} = [r_3, r_2]$$

....

$$[r_{n-1}, r_{n-2}] \begin{bmatrix} 1, & -q_n \\ 0, & 1 \end{bmatrix} = [r_{n-1}, 0] , \quad n \text{ even} ,$$

$$[r_{n-2}, r_{n-1}] \begin{bmatrix} 1, & 0 \\ -q_n, & 1 \end{bmatrix} \begin{bmatrix} 0, & 1 \\ 1, & 0 \end{bmatrix} , \quad n \text{ odd} .$$

In view of (7),

$$\begin{bmatrix} 1, & 0 \\ -q_1, & 1 \end{bmatrix} = \begin{bmatrix} z_1, & -z_0 \\ -w_1, & w_0 \end{bmatrix},$$

$$\begin{bmatrix} z_{k-1}, & -z_{k-2} \\ -w_{k-1}, & w_{k-2} \end{bmatrix} \begin{bmatrix} 1, & -q_k \\ 0, & 1 \end{bmatrix} = \begin{bmatrix} z_{k-1}, & -z_k \\ -w_{k-1}, & w_k \end{bmatrix},$$

$$\begin{bmatrix} z_{k-1}, & -z_k \\ -w_{k-1}, & w_k \end{bmatrix} \begin{bmatrix} 1, & 0 \\ -q_{k+1}, & 1 \end{bmatrix} = \begin{bmatrix} z_{k+1}, & -z_k \\ -w_{k+1}, & w_k \end{bmatrix},$$

$$k = 2, 3, \ldots, n-1 .$$

Hence

$$[a, b]\, Q = [r_{n-1}, 0]$$

where

$$Q = \begin{bmatrix} (-1)^n & z_{n-1}, (-1)^{n-1} & z_n \\ (-1)^{n-1} & w_{n-1}, (-1)^n & w_n \end{bmatrix}.$$

Set

(8) $$\begin{bmatrix} x \\ y \end{bmatrix} = Q \begin{bmatrix} s \\ t \end{bmatrix}.$$

Then

$$[r_{n-1}, 0] \begin{bmatrix} s \\ t \end{bmatrix} = c ,$$

that is,

$$s = \frac{c}{r_{n-1}},$$

$$t \in \mathfrak{F}[\zeta] \quad \text{arbitrary} .$$

It follows that a solution of (4) exists if and only if $r_{n-1} = (a, b)$ divides c; then (8) results in (5) and (6). □

Remark 1. The euclidean algorithm also yields the finite continued fraction expansion for a/b, viz.

$$\frac{a}{b} = q_1 + \cfrac{1}{q_2 + \cfrac{1}{q_3 + \ldots + \cfrac{1}{q_n}}} =_{\text{df}} [q_1, q_2, \ldots, q_n] .$$

Then

$$\frac{w_k}{z_k} = [q_1, q_2, \ldots, q_k], \quad 1 \leqq k \leqq n,$$

is called the k-th convergent to $[q_1, q_2, \ldots, q_n]$, see [2] and [6]. In particular, $w_n = a/r_{n-1}$ and $z_n = b/r_{n-1}$.

As in [6], we can arrange the steps of solving (4) into the table below:

$$(9) \qquad \begin{array}{cccccc} a & b & r_1 & \ldots & r_{n-2} & r_{n-1} \\ & q_1 & q_2 & \ldots & q_{n-1} & q_n \\ 1 & q_1 & w_2 & \ldots & w_{n-1} & w_n \\ 0 & 1 & z_2 & \ldots & z_{n-1} & z_n \end{array}$$

In applications we often seek for a particular solution $\hat{x}$, $\hat{y}$ such that the degree of one polynomial, say $\hat{x}$, is minimal. For this purpose we rewrite (5) and (6) as

$$x = x_0 + z_n t$$

$$y = y_0 - w_n t$$

and let

$$x_0 = q_0 z_n + r_0, \quad \partial r_0 < \partial z_n.$$

Then

$$x = r_0 + z_n(q_0 + t),$$

$$y = y_0 - w_n t$$

and, obviously,

$$\hat{x} = r_0,$$

$$\hat{y} = y_0 + w_n q_0$$

is *uniquely* determined by setting $t = -q_0$.

Two examples are included to demonstrate how Theorem 1 works.

Example 1. Let $\mathfrak{F} = \mathfrak{Q}$ and solve the equation

$$\zeta^3 x + (1 - \zeta)\, y = 1 - \zeta^2 + \zeta^3.$$

Expressions (3) and (7) result in the table

$$\begin{array}{ccc} \zeta^3 & 1 - \zeta & 1 \\ & -1 - \zeta - \zeta^2 & 1 - \zeta \\ 1 & -1 - \zeta - \zeta^2 & \zeta^3 \\ 0 & 1 & 1 - \zeta\,. \end{array}$$

 Since $r_{n-1} = 1$, the equation has a solution and all solutions are given via (5), (6) as follows

$$x = 1 - \zeta^2 + \zeta^3 + (1 - \zeta)\,t\,,$$
$$y = 1 + \zeta + \zeta^5 - \zeta^3 t\,,$$

where $t \in \mathfrak{Q}[\zeta]$ arbitrary.

The solution induced by the condition $\partial \hat{x} = \min$ is obtained by computing $q_0 = -\zeta^2$, $r_0 = 1$:

$$\hat{x} = 1\,,$$
$$\hat{y} = 1 + \zeta\,.$$

Example 2. As a second example consider the equation

$$(1 - \zeta)^2\, x + (\zeta^2 - 0{\cdot}5\zeta^3)\, y = \zeta - \zeta^2$$

over the field $\mathfrak{R}$.

The table becomes

$1 - 2\zeta + \zeta^2$	$\zeta^2 - 0{\cdot}5\zeta^3$	$1 - 2\zeta + \zeta^2$	0.5ζ	1
	0	$-0{\cdot}5\zeta$	$-4 + 2\zeta$	$0{\cdot}5\zeta$
1	0	1	$-4 + 2\zeta$	$1 - 2\zeta + \zeta^2$
0	1	$-0{\cdot}5\zeta$	$1 + 2\zeta - \zeta^2$	$\zeta^2 - 0{\cdot}5\zeta^3$

Since $r_{n-1} = 1$, the equation has a solution and, generally,

$$x = \zeta + \zeta^2 - 3\zeta^3 + \zeta^4 + (\zeta^2 - 0{\cdot}5\zeta^3)\,t\,,$$
$$y = 4\zeta - 6\zeta^2 + 2\zeta^3 - (1 - 2\zeta + \zeta^2)\,t\,.$$

The t is again an arbitrary polynomial of $\mathfrak{R}[\zeta]$.

Imposing the condition $\partial \hat{x} = \min$, we obtain $q_0 = 2 - 2\zeta$, $r_0 = \zeta - \zeta^2$, and hence

$$\hat{x} = \zeta - \zeta^2\,,$$
$$\hat{y} = 2 - 2\zeta\,.$$

Notice that it may well happen that $(\hat{x}, \hat{y})$ is not a unit.

NOTATION CONVENTION

Let

$$m \in \mathfrak{F}[\zeta]\,,$$
$$m = \mu_0 + \mu_1\zeta + \ldots + \mu_n\zeta^n\,,$$

and let $\partial m = n \geqq 0$.

Then we define* the *reciprocal polynomial* of m, denoted by $m^\sim$ or $\tilde{m}$, as

$$\tilde{m} = \mu_0\zeta^n + \mu_1\zeta^{n-1} + \ldots + \mu_n .$$

It is obvious that $\partial\tilde{m} \leqq \partial m$, the equality sign holding if and only if $(\zeta, m) = 1$. Further

$$(10) \qquad \zeta^{\partial m}(\tilde{m})^\sim = \zeta^{\widetilde{\partial m}} m ,$$
$$(m_1 m_2)^\sim = \tilde{m}_1\tilde{m}_2 .$$

We also consider the factorization

$$m = m^- m^+ , \quad m^-, m^+ \in \mathfrak{F}[\zeta]$$

where m^+ is a stable polynomial of largest degree. These factors are unique to within units in $\mathfrak{F}[\zeta]$, $m = m^- e^{-1} e m^+$. Note that if $\mathfrak{F} = \mathfrak{Q}$, then generally m^-, $m^+ \in \mathfrak{R}[\zeta]$. However, since the rationals are everywhere dense in the reals, they both can be approximated by m^-, $m^+ \in \mathfrak{Q}[\zeta]$ with any desired accuracy.

Further let**

$$m^* = \tilde{m}^- m^+ .$$

It follows that

$$\zeta^{\partial m^*} \tilde{m}\, m = \zeta^{\partial m} \tilde{m}^* m^* .$$

The m^* is stable if and only if $\tilde{m}^-$ is stable. Note also that $\partial m^* \leqq \partial m$ where the equality sign holds if and only if $(\zeta, m) = 1$. If $\mathfrak{F} = \mathfrak{Q}$ then m^* does not generally exist in $\mathfrak{Q}[\zeta]$.

It is also useful to define

$$\langle m \rangle = \mu_0 .$$

In words, $\langle . \rangle$ extracts the absolute term of a polynomial.

Further consider a formal power series $p \in \mathfrak{F}[[\zeta]]$ which is a ratio of two polynomials l, $m \in \mathfrak{F}[\zeta]$,

$$p = \frac{l}{m} = \pi_0 + \pi_1\zeta + \pi_2\zeta^2 + \ldots .$$

Then the *reciprocal series* $p^\sim$ is defined by

$$p^\sim = \left(\frac{l}{m}\right)^\sim = \frac{\tilde{l}\zeta^{\partial m}}{\tilde{m}\zeta^{\partial l}}$$

* If $\mathfrak{F} = \mathfrak{C}$, then $\tilde{m} = \bar{\mu}_0\zeta^n + \bar{\mu}_1\zeta^{n-1} + \ldots + \bar{\mu}_n$, where $\bar{\mu}_k$ is the complex conjugate of μ_k.

** For typographical reasons, the symbols $\tilde{m}^-$, $\tilde{m}^*$, etc. are used in place of $(m^-)^\sim$, $(m^*)^\sim$, etc. throughout all parts of the tripaper.

 and it can be formally written as

$$p^{\sim} = \pi_0 + \pi_1\zeta^{-1} + \pi_2\zeta^{-2} + \dots .$$

Also

$$\langle p \rangle = \pi_0 \, .$$

SYSTEM DESCRIPTION

By a system we shall essentially mean a discrete, constant, linear system. However, a formal definition is necessary [4].

Let

$\mathscr{T}$ = time set = $\mathfrak{Z}$ = (ordered) set of integers,
$\mathscr{U}$ = input values = $\mathfrak{F}^m$ = vector space of m-tuples over a field $\mathfrak{F}$,
$\mathscr{Y}$ = output values = $\mathfrak{F}^l$,
$\mathscr{X}$ = state space = $\mathfrak{F}^n$.

Then a *finite dimensional, discrete, constant, linear, m-input, l-output system over a field* $\mathfrak{F}$ is a triple $\{\mathbf{A}, \mathbf{B}, \mathbf{C}\}$ of homomorphisms

$$\mathbf{A} : \mathscr{X} \to \mathscr{X} ,$$
$$\mathbf{B} : \mathscr{U} \to \mathscr{X} ,$$
$$\mathbf{C} : \mathscr{X} \to \mathscr{Y} ,$$

defining the equations

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k$$
$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k ,$$

where $k \in \mathscr{T}$, $\mathbf{x} \in \mathscr{X}$, $\mathbf{u} \in \mathscr{U}$, $\mathbf{y} \in \mathscr{Y}$.

The n is dimension of the system.

We shall not usually make a distinction between **A**, **B**, and **C** as homomorphisms or as matrices representing these homomorphisms with respect to a given basis.

Throughout the tripaper we shall adopt the following assumptions.

(11) The system is *canonical*, i.e.,

$$\operatorname{rank}\,[\mathbf{B}, \mathbf{AB}, \dots, \mathbf{A}^{n-1}\mathbf{B}] = n ,$$

$$\operatorname{rank}\begin{bmatrix}\mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^{n-1}\end{bmatrix} = n .$$

This actually means that the system is assumed to be completely reachable and completely observable [3], [4].

(12) $m = l = 1$, that is, we consider single-input single-output systems (single-variable systems) only.

Our definition covers a fairly large class of systems. In particular, if $\mathfrak{F} = \mathfrak{R}$, the reals, we have sampled-data systems or intrinsically discrete systems in the usual sense. If $\mathfrak{F} = \mathfrak{Z}_p$, the residue classes of integers modulo a prime, we have a finite automaton.

A polynomial

$$a = \alpha_0 + \alpha_1 \zeta + \ldots + \alpha_p \zeta^p \in \mathfrak{F}[\zeta]$$

is said to be the *annihilating* polynomial of a matrix $\mathbf{A}$ if

$$\alpha_0 \mathbf{A}^p + \alpha_1 \mathbf{A}^{p-1} + \ldots + \alpha_p \mathbf{I} = \mathbf{0}$$

and no polynomial of less degree has this property. Observe that the annihilating polynomial is unique modulo units in $\mathfrak{F}[\zeta]$.

We shall see that the annihilating polynomial makes it possible to obtain a polynomial description of a system.

Let

$$\sigma_0 = 0\,,$$

$$\sigma_k = \mathbf{C}\mathbf{A}^{k-1}\mathbf{B}\,, \quad k = 1, 2, \ldots$$

and denote

$$s = \sigma_0 + \sigma_1 \zeta + \sigma_2 \zeta^2 + \ldots.$$

Actually, $\{\sigma_0, \sigma_1, \ldots\}$ is the impulse response of the system $\{\mathbf{A}, \mathbf{B}, \mathbf{C}\}$. Further set

$$d = \min_k \{k : \sigma_k \neq 0\}\,.$$

We recognize that d is the *discrete-time delay* of the system and that $d > 0$ by definition.

Now introduce a polynomial $b \in \mathfrak{F}[\zeta]$ such that

$$s = \frac{\zeta^d b}{a}\,. \tag{13}$$

Observe that $(b, \zeta) = 1$ in (13).

If $\mathfrak{F} = \mathfrak{R}$ or $\mathfrak{F} = \mathfrak{C}$, it is customary to call $\zeta^d b/a$ the *transfer function* of the system. We shall use the same terminology, but remember that for us this is not a function of a complex variable.

It is well-known [4] that the transfer function (13) characterizes the system completely if and only if the system is canonical. This is equivalent to the condition

$$(a, \zeta^d b) = 1,$$

which will be assumed throughout. In view of (11), therefore, we can use the transfer function to rigorously describe a system.

To obtain the dimension n of the system, observe that

$$s = \mathbf{C}\zeta(\mathbf{I} - \zeta\mathbf{A})^{-1}\,\mathbf{B}$$

and

$$s^{\sim} = \mathbf{C}(\zeta\mathbf{I} - \mathbf{A})^{-1}\,\mathbf{B}.$$

The n is the degree of the characteristic polynomial of $\mathbf{A}$, which appears as the denominator in $s^{\sim}$. By definition

$$s^{\sim} = \left(\frac{\zeta^d b}{a}\right)^{\sim} = \frac{\tilde{b}\zeta^{\partial a}}{\tilde{a}\zeta^{\partial\zeta^d b}}$$

and let $l = \max\{\partial\zeta^d b - \partial a, 0\}$. Then $n = l + \partial a$ provided the system is canonical.

Condition (11) might seem too restrictive. However, this is not the case since the *control* problems for a noncanonical system become either trivial or meaningless [4].

A system $\{\mathbf{A}, \mathbf{B}, \mathbf{C}\}$ is defined to be stable if

$$\mathbf{A}^k \to \mathbf{0} \quad \text{for} \quad k \to \infty$$

or if a, the annihilating polynomial of $\mathbf{A}$, is stable.

Hence a compatible definition of a stable polynomial is as follows. A polynomial $a \in \mathfrak{F}[\zeta]$ is *stable* if the sequence obtained by long division of $1/a$ into ascending powers of ζ has the form

$$\delta_0 + \delta_1\zeta + \delta_2\zeta^2 + \ldots$$

and approaches zero, i.e.,

$$\delta_k \to 0 \quad \text{for} \quad k \to \infty.$$

Note that if $(a, \zeta) \neq 1$, the a is not stable as $1/a$ does not attain the form required.

If the ground field is $\mathfrak{Q}$, $\mathfrak{R}$ or $\mathfrak{C}$, we encounter both stable and unstable polynomials. The situation is different, however, if $\mathfrak{F} = \mathfrak{Z}_p$. A careful analysis shows that no polynomial in $\mathfrak{Z}_p[\zeta]$ is stable save the units of $\mathfrak{Z}_p[\zeta]$.

The reader will have noticed that our system description involves the polynomials of $\mathfrak{F}[\zeta]$ rather than those of $\mathfrak{F}[z]$, $\zeta = z^{-1}$. Although the latter approach has become traditional in the literature, we find it more convenient to work with the indeterminate ζ. In doing so we bypass many difficulties regarding causality of the optimal system being synthesized.

Inasmuch as the reference input may also be viewed as a response of a system, we identify w with a ratio of two relatively prime polynomials of $\mathfrak{F}[\zeta]$,

$$w = \frac{q}{p},$$

for which $(p, \zeta) = 1$. A similar identification can be done for u and e.

For further reference, let $(a, p) = w$ and

$$(14) \qquad \begin{aligned} a &= a_0 w , \\ p &= \quad w p_0 . \end{aligned}$$

CONCLUSIONS

This paper is the first part of a tripaper on the algebraic theory of discrete optimal control for single-variable systems. Our aim here was to establish the mathematical machinery for defining a system and solving various optimal control problems. Specifically, we have defined abstract polynomials and formal power series and have shown how to solve Diophantine equations in polynomials.

In the two remaining parts the open-loop and the closed-loop optimal control problems will be discussed.

(Received June 30, 1972.)

REFERENCES

[1] S. S. L. Chang: Synthesis of Optimum Systems Control. McGraw-Hill, New York 1961.

[2] G. H. Hardy, E. M. Wright: An Introduction to the Theory of Numbers. Clarendon Press, Oxford 1960.

[3] R. E. Kalman: On the general theory of control systems. Proc. 1st IFAC Congress, Moscow 1960.

[4] R. E. Kalman, P. L. Falb, M. A. Arbib: Topics in Mathematical Systems Theory. McGraw-Hill, New York 1969.

[5] S. Lang: Algebra. Addison-Wesley, 1965.

[6] Ш. Х. Михелович: Теория чисел. Высшая школа, Москва 1967.

[7] A. P. Sage: Optimum Systems Control. Prentice-Hall, Englewood Cliffs, N.J. 1968.

[8] V. Strejc et al.: Syntéza regulačních obvodů s číslicovým počítačem. NČSAV, Praha 1965.

[9] J. T. Tou: Digital and Sampled-data Control Systems. McGraw-Hill, New York 1959.

[10] O. Zariski, P. Samuel: Commutative Algebra, Vol. 1. Van Nostrand, Princeton, N.J. 1958.

Ing. Vladimír Kučera, CSc., Ústav teorie informace a automatizace ČSAV (Institute of Information Theory and Automation — Czechoslovak Academy of Sciences), Vyšehradská 49, 128 48 Praha 2, Czechoslovakia.