IFAC

INTERNATIONAL FEDERATION
OF AUTOMATIC CONTROL

WARSZAWA 1969

Identification
General Principles

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

TECHNICAL
SESSION
19

Organized by
Naczelna Organizacja Techniczna w Polsce

INTERNATIONAL FEDERATION OF AUTOMATIC CONTROL

Identification
General Principles

TECHNICAL SESSION No 19

FOURTH CONGRESS OF THE INTERNATIONAL
FEDERATION OF AUTOMATIC CONTROL
WARSZAWA 16 – 21 JUNE 1969

Organized by
Naczelna Organizacja Techniczna w Polsce

Contents

Wydawnictwa Czasopism Technicznych NOT - Polska

Zakład Poligraficzny WCT NOT. Zam. 101/69.

THE IDENTIFICATION OF PROCESSES BY MINIMIZATION OF THE DISTANCE BETWEEN SETS OF SIGNALS

S.WEGRZYN, G.DENIS, J.DELISLE
Université de Sherbrooke
Sherbrooke, Québec, Canada

INTRODUCTION

Historically, we can distinguish two periods in the evolution of automation:

- the first period covers the classical automation of unit processes (with one or many variables).

- while the second deals with the automation of a plant as a whole (many interacting unit processes).

A plant, or a complex process, is determined by

a) a set of input signals which shall be called vector $\underline{u} = \begin{vmatrix} u_1 \\ \vdots \\ u_e \end{vmatrix}$

b) a set of intermediate output signals, called vector $\underline{v} = \begin{vmatrix} v_1 \\ \vdots \\ v_s \end{vmatrix}$

c) and by a set of disturbance signals, called vector $\underline{z} = \begin{vmatrix} z_1 \\ \vdots \\ z_p \end{vmatrix}$

The global output $\underline{Y}$ of the plant is a function of $\underline{u}$, $\underline{v}$ and $\underline{z}$, while the purpose of the control is to maximize, through a convenient choice of $\underline{u}$, a function of that global output

$$\max_{\underline{u}} \| Q\{[\underline{Y}(\underline{u}, \underline{v}, \underline{z})]\} \|$$

The complex automation is concerned with the introduction of a control system, normally by means of a computer, which, depending upon the values of $\underline{v}$ and upon statistical information on the values of $\underline{z}$, determines the optimum value $\underline{u}_{opt}$. Fig. 1.

MATHEMATICAL MODEL

The plant global output $\underline{Y}$ generally depends on the variables $\underline{u}$, $\underline{v}$, $\underline{z}$

$$\underline{Y} = \Phi^{X} (\underline{u}, \underline{v}, \underline{z}) \qquad (1)$$

This function Φ^{X}, which shows how the process depends on the variables $\underline{u}$, $\underline{v}$, $\underline{z}$, is called the mathematical model of the process.

It will no longer be necessary to distinguish between signals $\underline{u}$ and $\underline{v}$; we will denote them by a single vector:

$$\underline{X} = \begin{vmatrix} \underline{u} \\ \underline{v} \end{vmatrix}$$

This will allow us to simplify formula (1):

$$\underline{Y} = \Phi(\underline{X}, \underline{Z}) \qquad (2)$$

The Ideal Mathematical Model

We assume that we can directly measure the value of the global output $\underline{Y}$

$$\underline{y}_1, \underline{y}_2, \cdots, \underline{y}_k, \underline{y}_m$$

and the corresponding values:

$$\underline{x}_1, \underline{x}_2, \cdots, \underline{x}_k, \cdots, \underline{x}_m$$

$$\underline{z}_1, \underline{z}_2, \cdots, \underline{z}_k, \cdots, \underline{z}_m$$

After this series of measurements, three sets are available

a) set $E(\underline{Y}) = (\underline{y}_1, \underline{y}_2, \ldots, \underline{y}_k, \ldots, \underline{y}_m)$

b) sets $E(\underline{X}) = (\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k, \ldots, \underline{x}_m)$

$E(\underline{Z}) = (\underline{z}_1, \underline{z}_2, \ldots, \underline{z}_k, \ldots, \underline{z}_m)$

Using equation (2) we can calculate the elements of a new set, the elements of which are the values of $\underline{Y}$. (This time, those values are not measured but computed).

$$E[\underline{\Phi}(\underline{X}_k, \underline{Z}_k)] = [\underline{\Phi}(\underline{X}_1, \underline{Z}_1), \ldots, \underline{\Phi}(\underline{X}_i, \underline{Z}_i), \ldots, \underline{\Phi}(\underline{X}_m, \underline{Z}_m)]$$

The mathematical model is <u>ideal</u> if, for any value of k, equation (2) is satisfied; that is to say

$$\underline{Y}_k = \underline{\Phi}(\underline{x}_k, \underline{z}_k) \quad ; \quad \forall k$$

<u>Approximate Mathematical Model</u>

Since any information we may have on the process is very seldom complete, and since there always exists a relative uncertainty represented by vector $\underline{Z}$, the model derived will never be but an approximate model:

$$\underline{Y} \simeq \underline{F}^*(\underline{\alpha}\ \underline{X}, \underline{\gamma}) \tag{3}$$

where $\underline{F}^*$ is the function proposed by the experts who know the process intimately. This function shall be as accurate as the elements of vector $\underline{\alpha}$.

The constant vector $\underline{\gamma}$ takes into account the uncertainty existing in the process itself (noise signals included). We can rewrite $F^*(\underline{\alpha}\ \underline{X}, \underline{\gamma})$ as follows:

$$\underline{F}^*(\underline{\alpha}\ \underline{X}, \underline{\gamma}) = \underline{F}(\underline{X}, \underline{A}) \tag{4}$$

where $\underline{A}$ is a matrix with constant coefficients. Thus

$$\underline{Y} \simeq \underline{F}(\underline{X}, \underline{A})$$

GENERAL IDENTIFICATION ALGORITHM

Let us suppose that one can directly measure the values of $\underline{Y}$ and the corresponding values of $\underline{X}$. After this series of measurements, two sets are at one's disposal

a) set $E(\underline{Y}) = (\underline{y}_1, \underline{y}_2, \cdots, \underline{y}_k, \cdots, \underline{y}_m)$

b) set $E[\underline{F}(\underline{X}, \underline{A})] = [\underline{F}(\underline{X}_1, \underline{A}), \cdots, \underline{F}(\underline{X}_k, \underline{A}), \cdots, \underline{F}(\underline{X}_m, \underline{A})]$

the elements of which are calculated using equation (5), with the precision of the coefficients of matrix $\underline{A}$.

The identification problem is then to find the optimal values of the coefficients of matrix $\underline{A}$. They will be optimal in relation to the purpose of the automation system which will be built using the approximate mathematical model.

In this work, it will be shown that this problem can be solved by minimizing, in a conveniently chosen metric space, the distance between the two sets

$$E(\underline{Y}) \text{ and } E[\underline{F}(\underline{X}, \underline{A})]$$

$$d\{E(\underline{y})/\, E[\underline{F}(\underline{X}, \underline{A})]\} = \frac{1}{m} \sum_{k=1}^{k=m} d[\underline{Y}_k / \underline{F}(\underline{X}_k, \underline{A})] \qquad (6)$$

It is assumed that, in all the cases considered in this paper, the limit of equation (6) exists when m approaches infinity:

$$d\{E(\underline{y}) \,/\, E[\underline{F}(\underline{X}, \underline{A})]\} = \lim_{m \to \infty} \frac{1}{m} \sum_{k=1}^{k=m} d[\underline{Y}_k \,/\, \underline{F}(\underline{X}_k, \underline{A})] \qquad (7)$$

The matrix $\underline{A}$ composed of the coefficients which minimalize the distance between set $E(\underline{Y})$ and $E\{\underline{F}(\underline{X}, \underline{A})\}$ will be labeled $\underline{A}_{opt}$.

One then has

$$d\{E[\underline{Y}]/\ E[\underline{F}(\underline{X},\ \underline{A}\ \text{opt})\} = \quad = \underset{\underline{A}}{\text{Min}}\ \{E[\underline{Y}]/\ E[\underline{F}(\underline{X},\ \underline{A})]\} \tag{8}$$

where the number of elements in each set may be finite or infinite.

Equation (8) is a general identification algorithm.

In this equation, the form of the function F is derived from tests carried out on the processes, and is of course subjected to the precision of the coefficients of matrix $\underline{A}$.

The numerical determination of the coefficients of matrix $\underline{A}$ requires a two-step procedure: experimental measurements and minimization of the distance between two sets of signals. Figure 2 is a geometrical representation of the general identification algorithmwhich is based on the principle of minimizing the distance between the set of measured signals $E(\underline{Y})$ and the set of calculated signals $E(\underline{Y})$.

Conditions of Model Stability and Accuracy

Equation 8 is general and remains valid whether the various sets E have a finite or infinite number of elements. If the sets are infinite and if limit (7) exists, it is sufficient to add the following additional precision condition:

$$d[E(\underline{y})/\ E[\underline{F}(\underline{X},\ \underline{A}\ \text{opt})]] < \xi \qquad m = \infty \tag{9}$$

where ξ is a positive number determined by precision constraints.

In practice, finite sets composed of m values can be used providing m is large enough so as to yield a significant sample.

In such a case, one has to check if the methematical model of the process, $F(X, A_{opt})$, (where function $\underline{F}$ is the result of physical tests, and where the elements of matrix A_{opt} have been found from a finite set of measurements using the identification technique) satisfies the following conditions

1. Precision condition: as defined above by equation (3), with m having a finite value.

2. Stability condition: equation (9) has to be verified for all significant samples of the process.

IDENTIFICATION, IN EUCLIDEAN SPACE, OF NONLINEAR MODELS WITH NO DELAY

Nonlinear models with no delay can be divided into two categories:

a) differential models
b) finite-difference models

For all those, and for the cases in which the dynamic system has e input signals and s output signals, function $\underline{F}$ can be expressed as follows

$$\underline{F} = \underline{A}\,\underline{X} \qquad \text{où} \qquad \underline{X} = \begin{vmatrix} \underline{e}_1 \\ \vdots \\ \underline{e}_k \\ \underline{e}_e \end{vmatrix} \tag{10}$$

a) For the differential model

$$\underline{e}_k = \begin{vmatrix} e_k \\ e_k^{(1)} \\ \vdots \\ e_k^{(p_k)} \end{vmatrix} \qquad e_k^{(p_k)} = \frac{d^{p_k}(e_k)}{dt^{p_k}}$$

b) For the finite-difference models

$$\underline{e}_k = \begin{vmatrix} (e_k)_n \\ (e_k)_{n+1} \\ \vdots \\ (e_k)_{n+p_k} \end{vmatrix}$$

$(e_k)_n$ is the step function corresponding to the continuous function $e_k(t)$, with a suitable value of Δt.

Where

$$(e_k)_n = \lfloor e_k(n) \quad \text{where} \quad n = \frac{t}{\Delta t}$$

$$(e_k)_{n+p} = \lfloor e_k(n+p)$$

Thus, vector $\underline{Y} \cong \underline{AX}$ becomes

a) for the differential model

$$\underline{Y} = \begin{vmatrix} \underline{s}_1 \\ \vdots \\ \underline{s}_k \\ \vdots \\ \underline{s}_s \end{vmatrix} \qquad \text{where } s_k = \begin{vmatrix} s_k \\ s_k^{(1)} \\ s_k^{(2)} \\ \vdots \\ s_k^{(1k)} \end{vmatrix} \qquad s_k^{(1k)} = \frac{d^{1k} s_k}{d+^{1k}}$$

b) for the finite-difference model

$$\underline{s}_k = \left| \begin{array}{l} (s_k)_n \\ (s_k)_{n+1} \\ (s_k)_{n+1k} \end{array} \right|$$

(s_k) is the step function corresponding to the continuous function $s_k(t)$, with a proper value of Δt.

Thus matrix $\underline{A}$ will be the following rectangular matrix

$$\underline{A} = |a_{ij}| \; i = 1, 2, \ldots, \sum_{k=1}^{s} l_k$$

$$j = 1, 2, \ldots, \sum_{k=1}^{e} p_k$$

A series of measurements are necessary to identify the coefficients a_{ij} of matrix $\underline{A}$

$$\begin{aligned} E(\underline{X}) &= (\underline{X}_1, \underline{X}_2, \ldots, \underline{X}_k, \ldots, \underline{X}_m) \\ E(\underline{Y}) &= (\underline{Y}_1, \underline{Y}_2, \ldots, \underline{Y}_k, \ldots, \underline{Y}_m) \end{aligned} \tag{11}$$

To represent this measurement series, let us define two measurement vectors

$$\begin{aligned} \underline{X}_m^T &= \left| \underline{X}_1^T, \underline{X}_2^T, \ldots, \underline{X}_m^T \right| \\ \underline{Y}_m^T &= \left| \underline{Y}_1^T, \underline{Y}_2^T, \ldots, \underline{Y}_m^T \right| \end{aligned} \tag{12}$$

The general identification algorithm(8), in the case of the delayless linear model, reads as follows:

$$d[E(\underline{Y})/\, E(\underline{A}_{opt}\, \underline{X})] = \operatorname*{Min}_{\underline{A}} d[E(\underline{Y})/\, E(\underline{A}\, \underline{X})] \tag{13}$$

If, for minimizing the distance d between the sets $E(\underline{Y})$ and $E(\underline{A}\ \underline{X})$, we choose to work in a Euclidean space in which distance d is defined by

$$d^2[\underline{Y}/\ \underline{A}\ \underline{X}] = [\underline{Y} - \underline{A}\ \underline{X}]^T\ [\underline{Y} - \underline{A}\ \underline{X}] \tag{14}$$

it can be shown that equation (13) has the following solution

$$A_{opt} = [(X_m^T\ X_m)^1\ X_m^T\ Y_m]^T \tag{15}$$

This last equation implies that

$$\det \left| X_m^T\ X_m \right| \neq 0 \tag{16}$$

which can be called the condition of identifyability of a series of measurements.

It requires that the length of a series of measurements (number of components being linearly independant of vector X_m) be greater than the number of components of vector $\underline{X}$.

IDENTIFICATION OF NONLINEAR MODELS WITH NO DELAY

For this nonlinear model, function F may be represented as

$$\underline{F} = \sum_{k=1}^{n} \underline{A}_k\ \underline{F}_k\ (\underline{X}) \tag{17}$$

where $\underline{f}_1(X)$, $\underline{f}_2(X)$, ... , $\underline{f}_n(X)$ is the chosen series of nonlinear functions. Equation (17) may be written in a more general form:

$$\underline{F} = \underline{A}\ \underline{X} \tag{18}$$

where $\underline{A} = |\ \underline{A}_1, \ \dots\ , \underline{A}_k, \ \dots\ , \underline{A}_n\ |$

and $\underline{X}^T = |\ \underline{f}_1(\underline{X}), \ \dots\ , \underline{f}_k(\underline{X}), \ \dots\ , \underline{f}_n(\underline{X})\ |$

As in the preceeding case, one has to make a series of measurements

$$\underline{X}_1, \ \underline{X}_2, \ \dots\ , \underline{X}_m$$

$$\underline{Y}_1, \ \underline{Y}_2, \ \dots\ , \underline{Y}_m$$

to identify matrices $\underline{A}_1, \ldots, \underline{A}_k, \ldots, \underline{A}_n$, and one also has to calculate the series

$$\underline{X}_m = | \underline{f}_{1m}(\underline{X}), \ldots, \underline{f}_{km}(\underline{X}), \ldots, \underline{f}_{nm}(\underline{X}) |$$

where $$\underline{f}^T_{km}(X) = | f_k(X_1), \ldots, f_k(X_k), \ldots, f_k(X_m) |$$

Let $\underline{Y}^T_m$ be defined as

$$\underline{Y}^T_m = | \underline{Y}_1, \underline{Y}_2, \ldots, \underline{Y}_m |$$

For the minimal distance in a Euclidean space, equation (15) yields

$$| \underline{A}_{k;opt} |_{k=1, \ldots, n} = (\underline{X}^T_m \underline{X}_m)^1 \underline{X}_m \; | \underline{Y}_k |_{k=1, \ldots, n}$$

whence

$$\underline{A}_{k,opt} = (\underline{X}^T_m \underline{X}_m)^1 \underline{X}_m \underline{Y}_k$$

IDENTIFICATION OF MODELS WITH DELAYS

Let us suppose the existence of a delay which causes a shift of the independant variable in $\underline{X}$. We get

$$\underline{X} = | \underline{e}_1, \ldots, \underline{e}_k, \ldots, \underline{e}_e |^T \qquad \text{and}$$

a) for the differential model

$$\underline{e}_k = | e_k(t-T_k), \ldots, e^{(p_k)}(t-T_k) |$$

b) for the finite-difference model

$$\underline{e}_k = | (e_k)_{n-T_k}, \ldots, (e_k)_{n+p_k} - T_k |$$

In the case of a model with delays $T_1, T_2, \ldots, T_e$, function F can be expressed as

$$\underline{F} = F[A, X(T)] \tag{19}$$

where $T = |T_1, T_2, \ldots, T_e|$

The problem here is to find the optimal values of the coefficients of matrix $\underline{A}$ and of matrix $\underline{T}$:

$$d\{E(Y)/ E[F(A_{opt}, X(T_{opt})]\} =$$

$$= \underset{A,\ T}{\text{Min}}\ d\{E(Y)/ E[A, X(T)]\} \tag{20}$$

A practical approach to finding the solution of equation (20) is as follows. First, one sets estimated values for $T_1, T_2, \ldots, T_p$ and then, for each of those, one finds, using equation (15), the corresponding values for $A_{k,opt}$ and for distance $d_{k,min}$. This process leads to a discrete relation between $d_{k,min}$ and T_k. Figure 3 illustrates this geometrical relation from which T_{opt} can be evaluated along with the values of the corresponding matrix $\underline{A}$.

For a small value of the delay, Paddy's approximation is sufficient

$$x(t-\tau) = \sum_{k=0}^{n} (-1)^k \frac{\tau^k}{k!} x^{(k)}(t)$$

and it brings the delay problem to that of identifying coefficients τ^k, (this problem has been solved in earlier sections).

CONCLUSION

This work has shown how the identification problem can be solved by minimizing distances between various sets of signals. Equation (8) can be

regarded as a general identification algorithm. If the space is Euclidean, the solution is analytical (equation (15)), and if not, numerical techniques should yield accurate results.

One of the most interesting feature of this method is that it can deal with delays and non-linearities which are too often encountered in the actual modeling of industrial processes.

BIBLIOGRAPHY

[1] Identification par minimalisation des distances
S. Wegrzyn, G. Denis, J. Delisle. - Université de Sherbrooke, 1967.

[2] Théorie des opérations linéaires
S. Banach. - Varsovie, 1932.

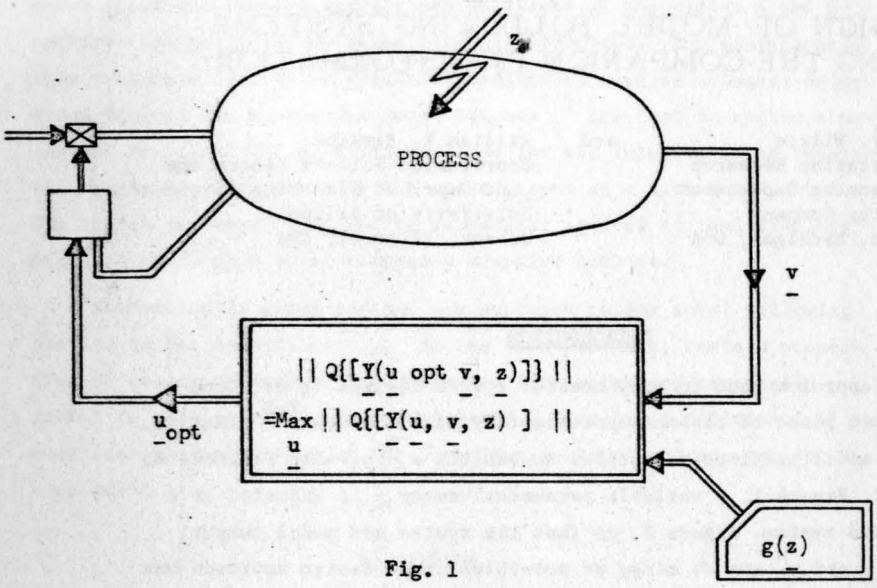

Fig. 1

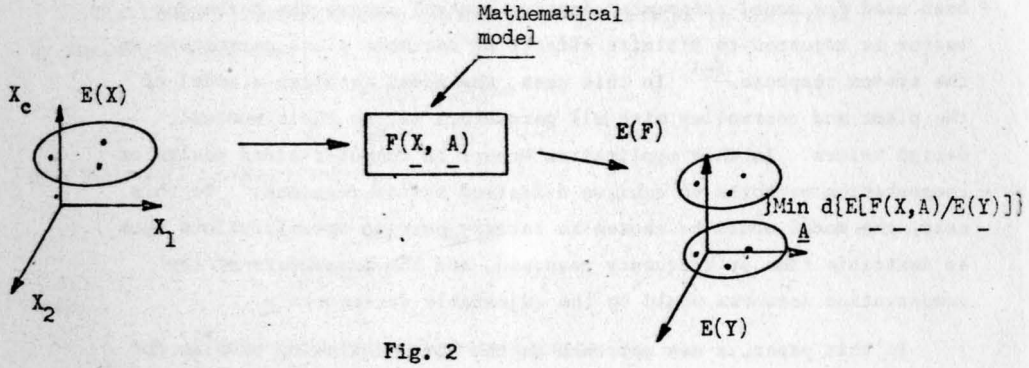

Fig. 2

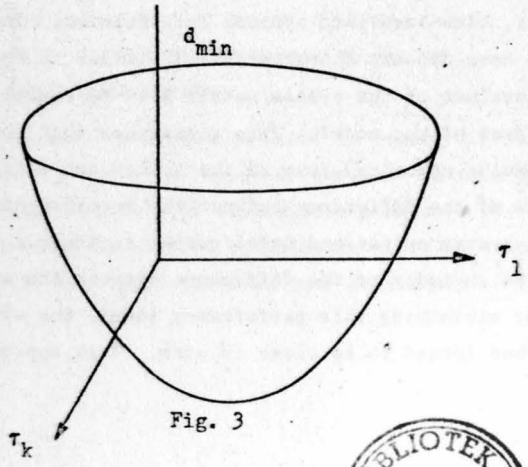

Fig. 3

DESIGN OF MODEL FOLLOWING SYSTEMS USING THE COMPANION TRANSFORMATION

Dennis F. Wilkie
Transportation Research
and Planning Department
Ford Motor Company
Dearborn, Michigan, USA

and

William R. Perkins
Coordinated Science Laboratory
and Dept. of Electrical Engineering
University of Illinois
Urbana, Illinois, USA

INTRODUCTION

An approach used in many control system designs is to force a controlled plant to respond approximately like a model.[1-4] That is, given a model previously selected to exhibit a desirable response to an input, Figure 1, a variable parameter vector $\underset{\sim}{v}$ is adjusted in a controlled system, Figure 2, so that the system and model output vectory $\underset{\sim}{y}$ and $\underset{\sim}{y}^m$ are as close as possible. This design approach has been used for model reference adaptive control, where the parameter vector is adjusted to minimize effects of variable plant parameters on the system response.[1-3] In this case, the model is often a model of the plant and controller with all parameters set to their nominal design values. Another application occurs in computer-aided design of compensation networks to achieve a desired system response.[4] In this case, the model would be chosen to satisfy certain specifications such as desirable time or frequency response, and the parameters of the compensation networks would be the adjustable parameters $\underset{\sim}{v}$.

In this paper, a new approach to the model following problem for single-input, linear, time-invariant systems is presented. The actual system is forced to have dynamic characteristics similar to the model by causing the eigenvalues of the system matrix A to be exactly the same as the eigenvalues of the model. This guarantees that the state vectors of the companion canonical form of the system and model are identical. The norm of the difference between the transformations relating the actual system output and model output to their companion states is selected as an index of the difference between the system and model responses. By minimizing this performance index, the system and model outputs are then forced to be close in norm. This approach is

shown to involve only algebraic manipulations of the system A and $\underset{\sim}{b}$ matrices and minimization of an algebraic function. The advantages of this technique over those requiring minimization of an integral square error between the system and model outputs[1-3] are that no system simulations are required, no averaging time for the integration is required, the result is independent of the system input as a time function, and the system and model need not be repeatedly excited in real time optimization as in some model reference adaptive schemes.

Motivation is given for the new approach to the model following problem in the second section. In the third section, certain properties of the transformation to companion form which are especially useful in this approach are demonstrated. Finally, a discussion of the numerical implementation of the technique is given and an example is considered.

MOTIVATION

The linear, time-invariant system to be controlled is described by the state equations

$$\dot{\underset{\sim}{x}} = A(\underset{\sim}{v})\underset{\sim}{x} + \underset{\sim}{b}(\underset{\sim}{v})u \qquad (1)$$
$$\underset{\sim}{y} = C\underset{\sim}{x},$$

and the model is described by

$$\dot{\underset{\sim}{x}}^m = A_m\underset{\sim}{x}^m + \underset{\sim}{b}_m u_o \qquad (2)$$
$$\underset{\sim}{y}^m = C_m\underset{\sim}{x}^m,$$

where

$\underset{\sim}{x}$, $\underset{\sim}{x}^m$ are n dimensional state vectors
$\underset{\sim}{y}$, $\underset{\sim}{y}^m$ are p dimensional output vectors
$\underset{\sim}{v}$ is an r dimensional parameter vector
u, u_o are scalar inputs.

The form of the controlled system is shown in Figure 2. It is desired to adjust $\underset{\sim}{v}$ to make $\underset{\sim}{y}$ and $\underset{\sim}{y}^m$ as close as possible in some sense. The method of achieving this will be to minimize a measure of the norm $\|\underset{\sim}{y}-\underset{\sim}{y}^m\|$. One way of doing this is to actually minimize a chosen norm by an iterative procedure. This has been done previously[1-3] for the norm

$$\|\underset{\sim}{y}-\underset{\sim}{y}^m\| = \tfrac{1}{2}\int_0^T < (\underset{\sim}{y}-\underset{\sim}{y}^m),\ Q(\underset{\sim}{y}-\underset{\sim}{y}^m) > dt.$$

However, such a minimization requires repeated simulations of the dynamic system and sensitivity equations which is time consuming. The approach here is to minimize a bound on the norm $\|\underset{\sim}{y}-\underset{\sim}{y}^m\|$, and it is shown to result in an optimization involving only algebraic equations (as opposed to differential equations). As shown in the considered examples, this approach can lead to a fast numerical minimization procedure, and the results are very close to those obtained by actually minimizing $\|\underset{\sim}{y}-\underset{\sim}{y}^m\|$.

Since experience has shown that a special choice of state variables may lead to an easier numerical minimization procedure, consider only two nonsingular, time-invariant, linear transformations T and T_m of the system and model state vectors, and denote the transformed states by $\underset{\sim}{z}$ and $\underset{\sim}{z}^m$, so that

$$\underset{\sim}{x} = T(\underset{\sim}{v})\underset{\sim}{z},\ \underset{\sim}{x}^m = T_m\underset{\sim}{z}^m. \tag{3}$$

Note that in general T is allowed to depend on $\underset{\sim}{v}$.

In terms of the transformed state variables, the problem of causing $\underset{\sim}{y}$ to follow $\underset{\sim}{y}^m$ by minimizing the norm of their difference becomes

$$\min_{\underset{\sim}{v}}\|C\underset{\sim}{x}-C_m\underset{\sim}{x}^m\| = \min_{\underset{\sim}{v}}\|CT(\underset{\sim}{v})\underset{\sim}{z}-C_mT_m\underset{\sim}{z}^m\|. \tag{4}$$

It would be desirable if the eigenvalues of the system could be forced to be the same as the model, since these characterize the unforced response of the system. The system and model may then be called "dynamically similar." Further, since the eigenvalues of a system are invariant under any nonsingular linear transformation, the system would be "dynamically similar" to the model independent of the choice of the state coordinates.

Pursuing this approach, it will become apparent that a particularly suitable transformation to use in (3) and (4) is the transformation to companion form, i.e., the A matrix in the state description of the transformed system is in companion form. It has been shown[5,6] that if a system as represented by (1) (or (2)) is controllable, then a

nonsingular transformation T (as in (3)) exists such that

$$\dot{\underset{\sim}{z}} = A_c \underset{\sim}{z} + \underset{\sim}{b}_c u \tag{5}$$

where

$$A_c = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 \cdots & 0 \\ \vdots & \vdots & 0 & & \vdots \\ & & \vdots & & 1 \\ -\alpha_1 & -\alpha_2 & & \cdots & -\alpha_n \end{bmatrix}, \quad \underset{\sim}{b}_c = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \tag{6}$$

and where $\alpha_1,\ldots,\alpha_n$ are the coefficients of the characteristic equation

$$s^n + \alpha_n s^{n-1} + \ldots + \alpha_2 s + \alpha_1 = 0. \tag{7}$$

Since $\alpha_1,\ldots,\alpha_n$ are the only parameters in state description of the companion system, it is evident that they completely determine $\underset{\sim}{z}$. Further, $\alpha_1,\ldots,\alpha_n$ are uniquely determined by the system eigenvalues, and vice-versa. Thus, if the system eigenvalues are forced to be the same as those of the model by requiring

$$\underset{\sim}{\alpha} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = \underset{\sim}{\alpha}^m = \begin{bmatrix} \alpha_1^m \\ \vdots \\ \alpha_n^m \end{bmatrix}, \tag{8}$$

it follows that $\underset{\sim}{z} = \underset{\sim}{z}^m$. Thus, the problem in (4) becomes a minimization with the equality constraint $\underset{\sim}{\alpha} = \underset{\sim}{\alpha}^m$, and then

$$\min_{\underset{\sim}{v}} \|\underset{\sim}{y}-\underset{\sim}{y}^m\| = \min_{\underset{\sim}{v}} \|(CT(\underset{\sim}{v}) - C_m T_m)\underset{\sim}{z}^m\|. \tag{9}$$

From this, it follows that

$$\min_{\underset{\sim}{v}} \|\underset{\sim}{y}-\underset{\sim}{y}^m\| \leq \min_{\underset{\sim}{v}} \{\|CT(\underset{\sim}{v}) - C_m T_m\|\} \, \|\underset{\sim}{z}^m\|, \tag{10}$$

where the operator norm indicated in (10) is some matrix norm. Suppose now that the upper bound on $\min_{\underset{\sim}{v}} \|\underset{\sim}{y}-\underset{\sim}{y}^m\|$ indicated in (10) is minimized. Then the minimization problem in (9) becomes the constrained <u>algebraic</u> minimization problem,

$$\min_{\underset{\sim}{v}} \|CT(\underset{\sim}{v}) - C_m T_m\|, \tag{10a}$$

with the equality constraint $\underset{\sim}{\alpha} = \underset{\sim}{\alpha}^{m}$. Thus, by choosing to minimize the upper bound indicated in (10) rather than $\|\underset{\sim}{y}-\underset{\sim}{y}^{m}\|$, the problem is an algebraic optimization. This minimization indicated in (10a) is logical, since if $\underset{\sim}{z} = \underset{\sim}{z}^{m}$, then in order for $\underset{\sim}{y}$ to be close to $\underset{\sim}{y}^{m}$ the difference of the transformations relating $\underset{\sim}{y}$ and $\underset{\sim}{y}^{m}$ to $\underset{\sim}{z}$ and $\underset{\sim}{z}^{m}$ should be small. This is exactly what is indicated in (10a). Further, the considered examples show that minimization of this bound can produce good results in making $\underset{\sim}{y}$ close to $\underset{\sim}{y}^{m}$. A suitable norm in the space of considered matrices has been found to be

$$\tfrac{1}{2} \sum_{i} \sum_{j} (CT(\underset{\sim}{v}) - C_{m}T^{m})^{2}_{ij} \qquad (11)$$

although other choices for the norm may yield good results in particular cases.

The question arises as to whether any desired $\underset{\sim}{\alpha}^{m}$ can be achieved by the adjusted system. It is shown that in the next section that if a suitable controller structure is assumed (namely, state feedback is included), the requirement (8) can be satisfied for any $\underset{\sim}{\alpha}^{m}$. Further, due to invariance of the companion transformation under state feedback, it is shown that the constrained minimization (10) can be reduced to an unconstrained minimization.

INVARIANCE OF THE COMPANION TRANSFORMATION

Morgan[7] used an important property of the companion form transformation to establish that any eigenvalues can be achieved with state feedback. This property is:

Property I:

Consider the system described by

$$\dot{\underset{\sim}{\hat{x}}} = (A(\underset{\sim}{v}) + \underset{\sim}{b}(\underset{\sim}{v})\underset{\sim}{k}')\underset{\sim}{\hat{x}} + \underset{\sim}{b}(\underset{\sim}{v})u_{o} = \hat{A}(\underset{\sim}{v},\underset{\sim}{k})\underset{\sim}{\hat{x}} + \underset{\sim}{b}(\underset{\sim}{v})u_{o}. \qquad (12)$$

Then the transformation $\hat{T}(\underset{\sim}{v},\underset{\sim}{k})$ of the system (12) to companion form is the same as $T(\underset{\sim}{v})$ for the system (1).

Proof:

Using the transformation $T(\underset{\sim}{v})$ in (3), let

$$\hat{\underset{\sim}{x}} = T(\underset{\sim}{v})\hat{\underset{\sim}{z}}. \tag{13}$$

Then (12) becomes

$$\dot{\hat{\underset{\sim}{z}}} = [T^{-1}(\underset{\sim}{v})A(\underset{\sim}{v})T(\underset{\sim}{v}) + T^{-1}(\underset{\sim}{v})\underset{\sim}{b}(\underset{\sim}{v})\underset{\sim}{k}'T(\underset{\sim}{v})]\hat{\underset{\sim}{z}} + T^{-1}(\underset{\sim}{v})\underset{\sim}{b}(\underset{\sim}{v})u_o \tag{14}$$

but using (6) this becomes

$$\dot{\hat{\underset{\sim}{z}}} = \left[A_c + \begin{bmatrix}0\\ \vdots\\ 0\\ 1\end{bmatrix} k'T(\underset{\sim}{v})\right]\hat{\underset{\sim}{z}} + \begin{bmatrix}0\\ \vdots\\ 0\\ 1\end{bmatrix} u_o \equiv \tilde{A}\,\hat{\underset{\sim}{z}} + \begin{bmatrix}0\\ \vdots\\ 0\\ 1\end{bmatrix} u_o. \tag{15}$$

Define

$$\hat{\underset{\sim}{k}}' = \underset{\sim}{k}'T(\underset{\sim}{v}) = (\hat{k}_1, \hat{k}_2, \ldots, \hat{k}_n). \tag{16}$$

Then, it follows from (14) that

$$\tilde{A} = \begin{bmatrix} 0 & 1 & 0 & & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \vdots & & & & \\ -\alpha_1+\hat{k}_1 & -\alpha_2+\hat{k}_2 & \cdots & & -\alpha_n+\hat{k}_n \end{bmatrix} \tag{17}$$

and thus $T(\underset{\sim}{v})$ transforms system (12) to companion form, <u>independent</u> of $\underset{\sim}{k}$, i.e., $T(\underset{\sim}{v}) = \hat{T}(\underset{\sim}{v},\underset{\sim}{k})$.

Now, as noticed by Morgan,[7] it is evident from (17) that any desired poles of the system (with corresponding $\alpha_1^m, \ldots, \alpha_n^m$) can be obtained for any $\underset{\sim}{v}$ by the following choice for $\underset{\sim}{k}$,

$$\underset{\sim}{k} = (T'(\underset{\sim}{v}))^{-1}(\underset{\sim}{\alpha} - \underset{\sim}{\alpha}^m). \tag{18}$$

Property I along with (18) can now be utilized to simplify the algebraic minimization problem stated in (10). Assume that the system to be optimized (with respect to a given model) is in the form (12) as shown in Figure 3. Also, assume for convenience of notation that $C=C^m$. For a system in this form, the constrained minimization (10) becomes

$$\min_{\underset{\sim}{v},\underset{\sim}{k}} \|\hat{T}(\underset{\sim}{v},\underset{\sim}{k}) - T^m\| \tag{19}$$

subject to the equality constraint

$$\underset{\sim}{\alpha}(\underset{\sim}{v},\underset{\sim}{k}) = \begin{bmatrix} \alpha_1(\underset{\sim}{v},\underset{\sim}{k}) \\ \vdots \\ \alpha_n(\underset{\sim}{v},\underset{\sim}{k}) \end{bmatrix} = \underset{\sim}{\alpha}^m. \tag{20}$$

Now utilizing Property I, it is evident that the minimization in (19) is independent of $\underset{\sim}{k}$. Thus, the problem is simply a non-constrained minimization with respect to $\underset{\sim}{v}$ of $\|T(\underset{\sim}{v}) - T^m\|$, and the condition (20) is satisfied after the optimal $\underset{\sim}{v}^*$ is found by solving (16) for the feedback vector $\underset{\sim}{k}$. Thus, by choosing the controller structure of Figure 3, the designer can greatly facilitate the computational solution.

NUMERICAL IMPLEMENTATION OF THE METHOD

The algebraic optimization of (10) and (11) involves the minimization

$$\min_{\underset{\sim}{v}} J(\underset{\sim}{v}) = \min_{\underset{\sim}{v}} \|CT(\underset{\sim}{v}) - C_m T_m\| = \min_{\underset{\sim}{v}} \frac{1}{2} \sum_i \sum_j (Ct_{ij}(\underset{\sim}{v}) - C_m t^m_{ij})^2. \tag{21}$$

The required minimization can be carried out by a variety of optimization techniques. A method found to be particularly successful is Davidon's method,[8,9] in which the iteration scheme at the i+1 st step is given by

$$\underset{\sim}{v}^{i+1} = \underset{\sim}{v}^i - \alpha^i H^i (\nabla_{\underset{\sim}{v}} J(\underset{\sim}{v}))_i \tag{22}$$

where α^i is a positive scalar constraint determined at each iteration step and H^i is a positive definite symmetric matrix (rxr if $\underset{\sim}{v}$ is an r vector) which is updated at each step of the minimization. The procedures for determining H^i and α^i at each step of the minimization are given in Appendix A. It should be noted that only simple matrix products are required for updating and no more information is needed in Davidon's method than is needed in the steepest descent method. For (21), the gradient components are given by

$$(\text{grad}_{\underset{\sim}{v}} J(\underset{\sim}{v}))_k = (\nabla_{\underset{\sim}{v}} J(\underset{\sim}{v}))_k = \frac{\partial J}{\partial v_k} = \sum_i \sum_j (Ct_{ij}(\underset{\sim}{v}) - C_m t^m_{ij}) \frac{\partial t_{ij}}{\partial v_k}. \tag{23}$$

Thus, some means of determining $\frac{\partial t_{ij}}{\partial v_k}$, i=1,...,n, j=1,...,n, k=1,...,r without finding $T(\underset{\sim}{v})$ as a general function of $\underset{\sim}{v}$ and then differentiating

is needed. The authors have shown[10] how this can be done, and the necessary computer algorithms are given in Appendix B. Using these algorithms, and assuming the functional dependences of the system matrices $A(\underset{\sim}{v})$ and $\underset{\sim}{b}(\underset{\sim}{v})$ are known, one can pointwise generate $T(\underset{\sim}{v})$ and $\frac{\partial T(\underset{\sim}{v})}{\partial v_k}$, $k=1,\ldots,r$ with relative ease on a digital computer. These results were implemented in example given in the next section.

EXAMPLE

Consider the system of Figure 4. Without the state feedback parameters (shown dotted), this system is a realization of an arbitrary fourth order transfer function if suitable values for $v_1,\ldots,v_8$ are chosen. Suppose now that this system is to be controlled in a model reference adaptive scheme, where $v_1,\ldots,v_6$ are to be adjusted to compensate for changes in v_7 and v_8. The model used in this case is the system with the nominal parameter values $v_1^o,\ldots,v_8^o$ all unity. Consider first the case where the output vector is exactly the state vector, $\underset{\sim}{y}=\underset{\sim}{x}$. For 5% changes in v_7 and v_8 from the nominal, $v_7 = 1.05$ and $v_8 = 0.95$, the computation of $v_1,\ldots,v_6$ and $k_1,\ldots,k_4$ according to the method of the paper required 6 iterations which took 31 seconds on a CDC 1604 computer. The computed values for $\underset{\sim}{v}$ and $\underset{\sim}{k}$ were

$$\underset{\sim}{v} = \begin{bmatrix} 1.024 \\ 1.0 \\ 0.992 \\ 0.971 \\ 1.07 \\ 1.02 \end{bmatrix}, \quad \underset{\sim}{k} = \begin{bmatrix} 0.01 \\ 0.081 \\ -0.08 \\ -0.048 \end{bmatrix}.$$

Note that k_2 and v_2 together constitute only one parameter adjustment. The step response for the state x_4 along with the corresponding model state is shown in Figure 5. The other states of the system and model compared equally well and are not shown. For comparison, the same problem was considered from the standpoint of adjusting $\underset{\sim}{v}$ to minimize the integral-square-error criterion

$$\hat{J}(\underset{\sim}{v}) = \frac{1}{2}\int_0^T \langle(\underset{\sim}{x}-\underset{\sim}{x}^m),\ (\underset{\sim}{x}-\underset{\sim}{x}^m)\rangle\, dt.$$

In this case, the solution for the optimal $\underset{\sim}{v}^*$ required 5 iterations and 10 minutes 50 seconds of computer time. The excessive solution time

for this approach was due to the repeated solution of the system and model state equations. The state trajectories were found to be essentially the same as those given by the technique of the paper (with respect to closeness to the model states).

Consider now the case where $y = x_4$, a scalar output. Then $\underset{\sim}{C}' = (0,0,0,1)$ and only the differences $t_{ij} - t^m_{ij}$ for $j=4$ are considered in the performance index to be minimized. The computation of $v_1,\ldots,v_6$ and $k_1,\ldots,k_4$ in this case required 1 iteration (8 seconds). The computed values for $\underset{\sim}{v}$ and $\underset{\sim}{k}$ are

$$\underset{\sim}{v} = \begin{bmatrix} 1.05 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \end{bmatrix}, \quad \underset{\sim}{k} = \begin{bmatrix} 0.053 \\ 0.048 \\ -0.050 \\ -0.048 \end{bmatrix}.$$

It is interesting to note that only one parameter, v_1, is changed in this case. The step response of state x_4 for this case is also shown in Figure 5, and it is evident that x_4 in the system and model are closer than in the previous case. This is perhaps to be expected, since all of the effect of adjusting $\underset{\sim}{k}$ is concentrated toward making only four terms in $T(\underset{\sim}{v})$ close to those of T_m, rather than on making all 16 terms of the transformation the same. It is interesting to note however, that in this problem all of the states $x_1,\ldots,x_4$ for the case $y = x_4$ were closer to the corresponding model responses than in the previous case when $\underset{\sim}{y} = \underset{\sim}{x}$. From these results, it is apparent that one can obtain good results in forcing a system to follow a model using the method of the paper. Also, due to the short computation time, the utility of the method for an adaptive scheme is apparent.

CONCLUSIONS

A new approach to solving the model following problem for single input, linear, time-invariant control system design was given. The method consists of forcing the system to be "dynamically equivalent" to the model by causing the system eigenvalues to be the same as the model. The index of the difference between the system and model responses chosen to be minimized is the norm of the difference between their companion transformations. This is then an algebraic minimization

with an equality side constraint. By utilizing the invariance of the companion transformation under state feedback, it was shown that the constrained algebraic minimization problem could be reduced to an unconstrained minimization, if a controller including state feedback was chosen. Finally, an example demonstrated how the approach of the paper can lead to a fast solution of the model following problem.

ACKNOWLEDGMENTS

This work was supported by the Joint Services Electronics Program (U. S. Army, U. S. Navy, and U. S. Air Force) under Contract No. DAAB-07-67-C-0199; and in part by the Air Force Office of Scientific Research under Grant No. 68-1579.

Mr. Wilkie was also supported by an NDEA Title IV Fellowship.

The authors would like to thank Professor P. Kokotović and Professor J. B. Cruz, Jr., for helpful discussions.

REFERENCES

1. H. P. Whitaker, "Design Capabilities of Model Reference Adaptive Systems," Proc. NEC, 1962, Chicago, pp. 241-249.

2. P. C. Parks, "Liapunov Redesign of Model Reference Adaptive Control Systems," IEEE Trans. on Automatic Control, Vol. AC-11, No. 3, July, 1966, pp. 362-367.

3. P. V. Kokotović, et. al., "Sensitivity Method in the Experimental Design of Adaptive Control Systems," Third IFAC Congress, London, June, 1966, pp. 45B.1-45B.12.

4. T. C. Coffey, "The Application of Modern Computing Technology to Control System Analysis and Design Problems," Defense Documentation Center Document AD660-133, June, 1967, pp. 63-107.

5. W. M. Wonham and C. D. Johnson, "Optimal Bang-Bang Control with Quadratic Performance Index," Proc. 1963 4th JACC, pp. 101-112.

6. D. S. Rane, "A Simplified Transformation to (Phase-Variable) Canonical Form," IEEE Trans. on Automatic Control, Vol. AC-11, No. 3, July, 1966, p. 608.

7. B. S. Morgan, Jr., "Sensitivity Analysis and Synthesis of Multi-Variable Systems," IEEE Trans. on Automatic Control, Vol. AC-11, No. 3, July, 1966, pp. 506-512.

8. W. C. Davidon, "Variable Metric Method for Minimization," AEC Res. and Div. Report ANL-5990 (Rev.), November, 1959.

9. R. Fletcher and M. J. D. Powell, "A Rapidly Convergent Descent Method of Minimization," The Computer Journal, Vol. 6, No. 2, 1963, pp. 163-168.

10. D. F. Wilkie and W. R. Perkins, "Essential Parameters in Sensitivity Analysis," Proc. of the Second IFAC Symposium on System Sensitivity and Adaptivity, Dubrovnik, Yugoslavia, August, 1968.

APPENDIX A

The steps to be implemented in a minimization procedure using Davidon's method[8,9] are as follows:

(1) Choose $H^0 = I$ (rxr identity matrix).

(2) Having $\underset{\sim}{y}^i$ and H^i and $(\nabla_{\underset{\sim}{y}} J)_i$, find α^i such that

$$J(\underset{\sim}{y}^i - \alpha^i H^i (\nabla_{\underset{\sim}{y}} J)_i) = \min_{\lambda} \{J(\underset{\sim}{y}^i - \lambda H^i (\nabla_{\underset{\sim}{y}} J)_i)\}.$$

Then form $\underset{\sim}{y}^{i+1}$ according to (22).

(3) Find $(\nabla_{\underset{\sim}{y}} J)_{i+1}$ and define

$$\underset{\sim}{\gamma}^i \triangleq (\nabla_{\underset{\sim}{y}} J)_{i+1} - (\nabla_{\underset{\sim}{y}} J)_i .$$

(4) Form the matrix H^{i+1} by

$$H^{i+1} = H^i + A^i + B^i$$

where

$$A^i = \frac{(\alpha^i H^i \nabla_{\underset{\sim}{y}} J/_i)(\alpha^i H^i \nabla_{\underset{\sim}{y}} J/_i)^T}{(-\alpha^i H^i \nabla_{\underset{\sim}{y}} J/_i)^T \underset{\sim}{\gamma}^i}$$

and

$$B^i = - \frac{(H^i \underset{\sim}{\gamma}^i)(H^i \underset{\sim}{\gamma}^i)^T}{(\underset{\sim}{\gamma}^i)^T H^i \underset{\sim}{\gamma}^i} .$$

Return to step (2) and repeat the process, continuing until some predetermined stopping criteria (indicating the minimum is obtained) are satisfied.

APPENDIX B

The algorithm for computing the transformation to companion form for a given system is as follows:[5,6]

$$T = [\underset{\sim}{t}_1 \cdots \underset{\sim}{t}_n]$$

where

$$\begin{aligned} \underset{\sim}{t}_n &= \underset{\sim}{b} \\ \underset{\sim}{t}_{n-1} &= A\underset{\sim}{t}_n + \alpha_n \underset{\sim}{t}_n \\ \underset{\sim}{t}_{n-2} &= A\underset{\sim}{t}_{n-1} + \alpha_{n-1} \underset{\sim}{t}_n \\ &\vdots \\ \underset{\sim}{t}_1 &= A\underset{\sim}{t}_2 + \alpha_2 \underset{\sim}{t}_n . \end{aligned} \tag{B1}$$

The coefficients of the characteristic equation which are required to find the transformation can be calculated by using Leverrier's algorithm:

$$\alpha_{n+1} = 1 \qquad S_{n+1} = I \qquad \text{(nxn identity matrix)}$$

$$\alpha_{n-j+1} = -(1/j)\,\mathrm{tr}(AS_{n-j+2}), \quad S_{n-j+1} = \alpha_{(n-j+1)} I + AS_{n-j+2}. \tag{B2}$$

A check on the numerical calculations is that $S_1=0$ should be obtained.

Using these results, one can obtain the algorithm for computing $\frac{\partial T(\underset{\sim}{v})}{\partial v_i}$, i=1,...,r, which is as follows:

$$\frac{\partial T}{\partial v_i} = \left[\frac{\partial \underset{\sim}{t}_1}{\partial v_i} \cdots \frac{\partial \underset{\sim}{t}_n}{\partial v_i}\right] \tag{B3}$$

$$\frac{\partial \underset{\sim}{t}_n}{\partial v_i} = \frac{\partial \underset{\sim}{b}}{\partial v_i}$$

$$\frac{\partial \underset{\sim}{t}_{n-1}}{\partial v_i} = \frac{\partial A}{\partial v_i}\underset{\sim}{t}_n + A\frac{\partial \underset{\sim}{t}_n}{\partial v_i} + \frac{\partial \alpha_n}{\partial v_i}\underset{\sim}{b} + \alpha_n \frac{\partial \underset{\sim}{b}}{\partial v_i}$$

$$\frac{\partial \underset{\sim}{t}_{n-2}}{\partial v_i} = \frac{\partial A}{\partial v_i}\underset{\sim}{t}_{n-1} + A\frac{\partial \underset{\sim}{t}_{n-1}}{\partial v_i} + \frac{\partial \alpha_{n-1}}{\partial v_i}\underset{\sim}{b} + \alpha_{n-1}\frac{\partial \underset{\sim}{b}}{\partial v_i}$$

$$\vdots$$

$$\frac{\partial \underset{\sim}{t}_1}{\partial v_i} = \frac{\partial A}{\partial v_i}\underset{\sim}{t}_2 + A\frac{\partial \underset{\sim}{t}_2}{\partial v_i} + \frac{\partial \alpha_2}{\partial v_i}\underset{\sim}{b} + \alpha_2\frac{\partial \underset{\sim}{b}}{\partial v_i},$$

and the derivatives $\frac{\partial \alpha_k}{\partial v_i}$, k=1,...,n, i=1,...,r, can be pointwise calculated by the derivative of the Leverrier algorithm:

$$\frac{\partial \alpha_{n+1}}{\partial v_i} = 0 \qquad \frac{\partial S_{n+1}}{\partial v_i} = 0 \tag{B4}$$

$$\frac{\partial \alpha_{n-j+1}}{\partial v_i} = (-\frac{1}{j})\,\mathrm{tr}\left(\frac{\partial A}{\partial v_i}S_{n-j+2} + A\frac{\partial S_{n-j+2}}{\partial v_i}\right)$$

$$\frac{\partial S_{n-j+1}}{\partial v_i} = \frac{\partial \alpha_{n-j+1}}{\partial v_i}I + \frac{\partial A}{\partial v_i}S_{n-j+2} + A\frac{\partial S_{n-j+2}}{\partial v_i}.$$

Thus, by using (B1) through (B4), $T(\underset{\sim}{v})$, $\frac{\partial T}{\partial v_i}$, and $\frac{\partial \alpha_k}{\partial v_i}$, k=1,...,n, i=1,...,r at any point $\underset{\sim}{v}$ can be calculated assuming only the functional dependence of A and $\underset{\sim}{b}$ on $\underset{\sim}{v}$ are known.

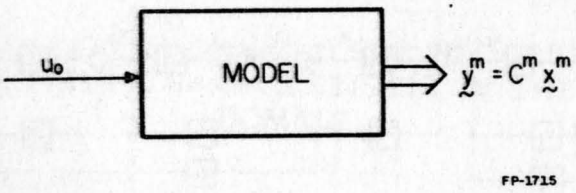

Fig. 1. Model.

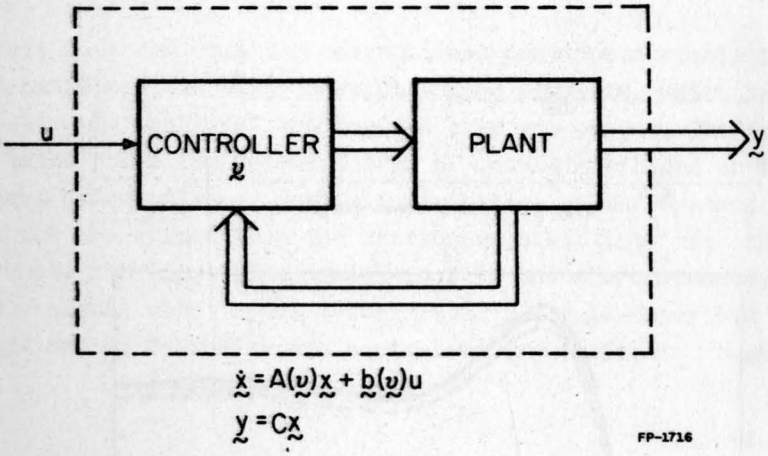

Fig. 2. System consisting of controller and plant.

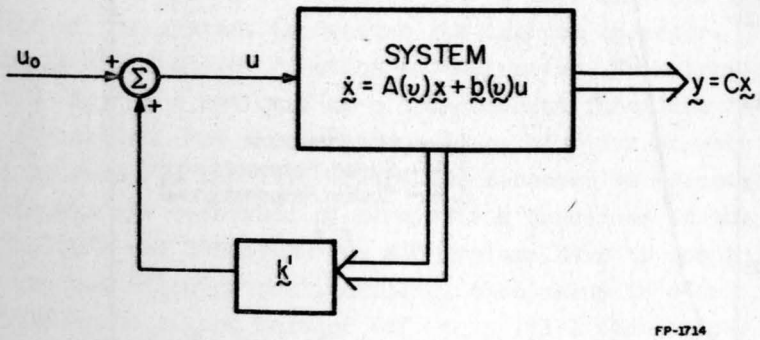

Fig. 3. System with state feedback added.

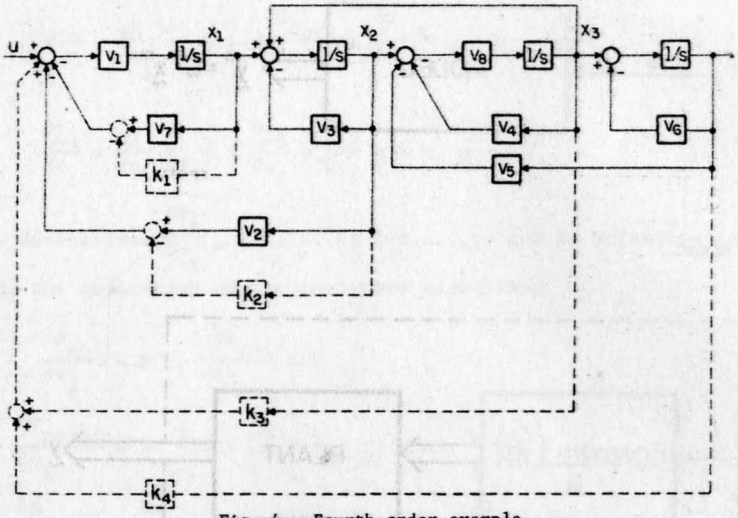

Fig. 4. Fourth order example.

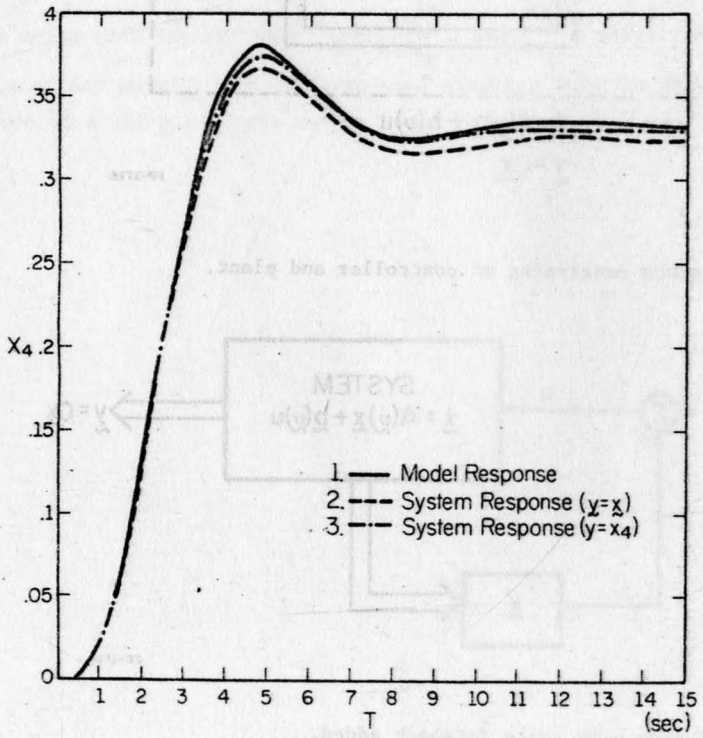

Fig. 5. Responses of state x_4 for the model and for the system in the cases $\underset{\sim}{y} = \underset{\sim}{x}$ and $y = x_4$.

DETERMINATION OF MODELS FOR IDENTIFICATION WITH QUALITY CHARACTERISTIC IN THE TIME DOMAIN

Gunter Schwarze
Humboldt-Universität zu Berlin
Berlin
German Democratic Republic

0. Introduction

In this paper we consider conventional continuous stable linearly dynamic systems with constant lumped elements, which have one and only one input and one and only one output. The following methods for the determination of models are based on the numerical calculation of such signals acting on the system, which describe the change from one stationary state into the other. We denote the input signal with $x_e(t)$ and the corresponding output signal with $x_a(t)$. The system $\underline{S}$ may be described unique of the following two equivalent expressions:

$$S = ((x_e(t), x_a(t)) \quad , \tag{1}$$

$$F_S(s) = L(x_a(t)) \,/\, L(x_e(t)) \quad \text{with } (x_e(t), x_a(t)) \in S \quad . \tag{2}$$

For eq. (1) we need the whole set of all permissible pairs of signals of the system and for eq. (2) we need only one pair of signals of the system. L denotes the Laplace operator. $F_S(s)$ is the transfer function of the system. This transfer function may be a rational or a transcendent function. We take such a notation that the functions describing the signals are equal to zero for negative arguments. Moreover we assume that all signals are described of permissible functions in the sense of FÖLLINGER and SCHNEIDER[3,5]. All systems have to comply with the condition of realizability, that means to every permissible input signal belongs (cf. eq. (1)) a permissible output signal. All the following models are systems $\underline{M}$ complying with the above conditions for systems. In the same way we take the terms and designations M, $F_M(s)$, $m_e(t)$, $m_a(t)$. A system $\underline{M}$ is called a model of the system $\underline{S}$, if special signals $m_e(t), m_a(t)$ and $x_e(t), x_a(t)$

$((m_e(t),m_a(t)) \in M$; $(x_e(t),x_a(t)) \in S)$ accomplish a quality characteristic. In the following we denote all signal functions without size.

We consider only global methods[5]. They are distinguished by single and repeated integrals as basis of the determination of models.

We assume one pair of signal functions $(x_e(t),x_a(t)) \in S$ may be given with the condition that both signal functions are empirical functions (graphics, point plottings, or tabulated functions). For the digital computation special numerical processes for integration must be investigated.

These methods are special results of a research team of the "Sektion Mathematik" (Mathematical Methods for Information Processing) under the direction of the author.

1. Characterization of rational models by integral equations and its solutions

1.1. Introductory considerations

A model with rational transfer function will be called a rational model. We assume the model being asymptotical stable. The transfer function of the model can now be written as

$$F_M(s) = Z_M(s) \,/\, N_M(s) \tag{3}$$

$$Z_M(s) = \sum_{l=0}^{n} b_l s^l \tag{4}$$

$$N_M(s) = \sum_{k=0}^{n} a_k s^k \tag{5}$$

with n as the order of the system, $a_k > 0$ $(k=0,\ldots,n)$, at least one $b_l \neq 0$ and $N_M(s)$ is a Hurwitz polynomial. Between two permissible input resp. output signals of a pair of signals with $t=0$ as unique exception point we have the integral equation

$$\sum_{k=0}^{n} a_k I^{n-k}(m_a(t)) = \sum_{l=0}^{n} b_l I^{n-l}(m_e(t)) \quad (0 < t < +\infty) \tag{6}$$

with

$$I(m(t)) = \int_0^t m(u)du \ , \quad I^r(m(t)) = \frac{1}{(r-1)!} \int_0^t (t-u)^{r-1} m(u)du \tag{7}$$

The polynom corresponds to the homogeneous integral equation

$$\sum_{k=0}^{n} a_k J^{n-k}(m(t)) = 0 \qquad (0 < t < +\infty) \tag{8}$$

with

$$J(m(t)) = \int_t^{+\infty} m(u)du, \quad J^r(m(t)) = \frac{1}{(r-1)!} \int_t^{+\infty} (t-u)^{r-1} m(u)du \ . \tag{9}$$

From both equations follows that the given functions are linear dependent. All following considerations will refer to linear normed function spaces. There we take different norms and different integral equations (eq. (6, 8)). It follows that the first method gives the coefficients (eq. (4, 5)) in one step and the other methods in two steps (at first coefficient of eq. (5), then of eq. (4)).

1.2. Evaluation of the inhomogeneous integral equation

We consider the linear space Z of all permissible functions with the single point of exception $t = 0$. For these functions there exist all derivatives in $(0,+\infty)$ and the limits from the left side for $t = 0$ exist for the functions and all their derivatives. Now we take one pair $(m_e(t), m_a(t)) \in M$ satisfying eq. (6). We denote

$$I^k(m_e(t)) = z_{k+1} \ , \quad I^k(m_a(t)) = z_{k+n+2} \tag{10}$$

$$b_k = c_{n+1-k} \ , \quad a_k = -c_{2n+2-k} \ , \tag{11}$$

Substituting eq. (10, 11) into eq. (6) yields with $a_0 = 1$:

$$z_{2n+2} = \sum_{i=1}^{2n+1} c_i z_i \ . \tag{12}$$

We define in the space Z a scalar product $(z,\hat{z})$, the norm $//z// = (z,z)^{1/2}$, and the metric $D(z,\hat{z}) = //z-\hat{z}//$. If $z_i \in Z$ $(i=1,\ldots,2n+1)$ are linearly independent, these are a base of a subspace Z_{2n+1}. May be $z \in Z_{2n+1}$. If

$D(z,\tilde{z}) = \min$ for any one $z \in Z$, it follows

$$(z-\tilde{z},z_i) = 0 \qquad (i=1,\ldots,2n+1) . \qquad (13)$$

$D(z,\tilde{z}) = 0$ if and only if z is linearly independent from $z_1,\ldots,z_{2n+1}$. Substituting $z = z_{2n+2}$ into eq. (12) and considering eq. (13) yields

$$\sum_{i=1}^{2n+1} c_i(z_i z_k) = (z_{2n+2},z_k) \qquad (k=1,\ldots,2n+1) . \qquad (14)$$

Eq. (14) is a system of linear equations for c_i. Hitherto we used no special form for the scalar product. This shows the universality of this method. In part 3 of this paper we used for the special numerical method a special form of the scalar product. Specially we mention that this method generally uses the metric not relative to the signal functions but relative to the eq. (6).

1.3. Evaluation of the homogeneous integral equation
- norm of $L^2(0,+\infty)$ -

We assume $m_a(t)$ is a linear combination of a full system of linearly independent functions of eq. (8). We get the full system from the characteristic equation belonging to eq. (8), which corresponds to $N_M(s)$. We assume that the model is asymptotic stable (cf. 1.1.). It follows $m_a(+\infty)=\lim_{t\to+\infty} m_a(t)=0$. The step response satisfies this condition for $b_0=0$ (cf.eq.(4)). If $m_a(+\infty) \neq 0$ we use $m_a(t) - m_a(+\infty)$. Studies being analogous to these are possible for input signal functions, e. g. if the signal function is the step response or a impulse response of an other rational system. Now we denote the signal function with $m(t)$ and assume $m(+\infty) = 0$.

s_i may be the roots of the polynomial $N_M(s)$ with the multiplicity v_i. May be r the number of the different roots, then we get $v_1+v_2+\ldots+v_r = n$. The functions

$$e^{s_1 t},\ldots,t^{v_1-1}e^{s_1 t},\ldots,e^{s_r t},\ldots,t^{v_r-1}e^{s_r t} \qquad (15)$$

are a full system of linearly independent functions of eq.(8). We denote these with $g_1(t),\ldots,g_n(t)$.

A function $m(t) \in L^2(0,+\infty)$ can be expressed in the form

$$m(t) = \sum_{i=1}^{n} c_i \, g_i(t) \quad , \tag{16}$$

if and only if $m(t)$ satisfies eq. (8) almost everywhere and the norm is given by

$$(//m(t)//)^2 = \int_0^{+\infty} /m(t)/^2 \, dt \quad . \tag{17}$$

At first we assume $c_i \neq 0$ $(i=1,\ldots,n)$ in eq. (16). For this model the left side of eq. (8) has the characteristic that

$$D(0, \sum_{i=1}^{\hat{n}} \hat{a}_i \, J^{\hat{n}-i}(m(t)) \tag{18}$$

as a function of the coefficients accepts for $\hat{n}<n$ a minimal value unlike zero and for $\hat{n} = n$ and $\hat{a}_i = a_i$ the minimal value zero. We determine the integral equation as follows. Since $m(t) \neq 0$ is $//m(t)// \neq 0$. We denote $m_0(t) = m(t)$ and

$$m_1(t) = a_{01} \, m_0(t) + J^1(m(t)) \quad . \tag{19}$$

We determine a_{01} from the problem $//m_1(t)//^2 \stackrel{!}{=}$ minimum. If this minimum is unlike zero, then $n > 1$. We take step for step the formula

$$m_i(t) = a_{01} \, m_0(t) + \ldots + a_{i-1,i} \, m_{i-1}(t) + J^i(m(t)) \tag{20}$$

and determine a_{ji} from the problem $//m_i(t)//^2 \stackrel{!}{=}$ minimum and finish, when this minimum has got the value zero. Then is $i = n$ and therefore n is determined. From eq. (20) for $i = n$ and the equations corresponding to those $i<n$ one gets the integral equation (8) immediately by successive putting in. It follows from the integral equation the polynomial eq. (5) and thus also the full system eq. (15). In the second step of the whole method we determine the polynomial eq. (4) immediately from

$$D^2(m(t) , \sum_{i=1}^{n} c_i \, g_i(t)) \stackrel{!}{=} \text{minimum} \quad . \tag{21}$$

In this way we determine c_i. It is well known how to solve

it. We mention that the second step is also possible for $L^2(0, T)$ $(T > 0)$. With this method one gets a formula for $m(t)$ and therefore for $m_a(t)$. If we also know the formula for $m_e(t)$ it follows $F_M(s)$ from eq. (2). This is for instance possible, if $m_e(t)$ is a special test signal.

Hitherto we assumed that $c_i \neq 0$ $(i=1,\ldots,n)$ in eq. (16). If $L(m_e(t))$ is a rational function of s, we can also use this method, in case this assumption is not valid. If $L(m_e(t))$ is not a rational function we can use this method just nearly (cf. chapter 2).

1.4. Evaluation of the homogeneous integral equation
- norm of Tschebyscheff -

This method is like the method in 1.3. We use the linear space of the continuous functions $m(t)$ for $0 < t < +\infty$ with the condition $m(+\infty) = 0$ and the norm

$$//m(t)// = \max_{0 \leq t < +\infty} (p(t)/m(t)/) \quad . \qquad (22)$$

In eq. (22) there is $p(t)$ a continuous function with $p(t) > 0$ for $0 < t < +\infty$. We mention that in this case eq. (15) is a Tschebyscheff function system.

2. Approximation problems and quality characteristics

The theory described in chapter 1 yields in practice the following difficulties:

a) The transfer function of the system $\underline{S}$ is a transcendental function and not a rational function, e. g. the equations of the system are special partial differential equations.

b) The transfer function of the system $\underline{S}$ is a rational function, but the signal functions are given with limited accuracity (graphics, point plottings, or tabulated functions).

c) When we compute the functions $I^k(m(t))$ and $J^k(m(t))$ for a model we must use numerical methods for the operation integration and consequently we get truncation errors. The same is valid for the scalar product and the norm.

There are logically two typs of errors:

1) The system $\underline{S}$ is not identical with the model $\underline{M}$.

2) Using numerical methods for the computation of the model we get truncation erros.

Consequently we can in each case formulate the minimum conditions, but the given conditions are not exactly realizable. There is for example no finite number n with $//m_n(t)//=0$ exactly valid. Convergence investigations are necessary for securing errors sufficient little. We got in the case of 1.3. the following result: If $x(t) \in L^2(0,+\infty)$ and

$$/x(t)/ < e^{-At} \qquad \text{mit} \qquad A>1 \quad , \tag{23}$$

is valid almost everywhere, then the procedure described in 1.3. converges with

$$x_n(t) = \sum_{i=0}^{n-1} a_{in}\, x_i(t) + J^n(x(t)) \tag{24}$$

and the estimation

$$//x_n(t)//^2 < A^{-(2n+1)} \tag{25}$$

is valid. Hitherto for the computation we choose a predetermined upper bound for the order n of the model and a predetermined lower bound for the norm. The computation will be finished, if one of both bounds has been reached. For the procedure described in 1.2. we use the norm for the determined inhomogeneous integral equation as quality characteristic. The procedures described in 1.3. and 1.4. use for quality characteristics the norm of the determined homogeneous integral equation and the norm of the approximated signal function. The respective bounds of the norms we must strive for, must be given as quality numbers. With the help of these numbers we decide for the acceptance or unacceptance of the determined model for the system. We have used a predetermined upper bound for the order of the model, because we do not want to determine such models, which we can not accept for too high order. Minimizing the truncation errors we developed further numerical processes

under the considered conditions. We declared that the numerical processes for the integration (cf. 1.2.) with a fixed order of the interpolation polynomial for models with high order are not enough exactly. We generated a special numerical process for repeated integrals with variable upper bound. In this case, the function $m(t)$ will be substituted for a interpolation polynomial of the order 2 relative to three interpolation points following immediately one against the other. Then the integration can be exactly performed like eq. (7). This method is evident possible for each order.

3. Describing of the programmed procedures and results of computation

The following described procedures are approximate algorithms for the procedures in chapter 1. We identificate the signal functions of the system $\underline{S}$ with those of the models over the basic set of points (fixed points), in consideration of the accuracy of the measuring values (corresponding to the signal functions of the system). We note down the equations in the denotation of the models, as we determine indeed the equations of the model. Besides the numerical processes for the integration described in chapter 2 we used conventional numerical processes for resolving simultaneous linear equations and for determination of the roots of polynomials. Giving a first estimation of the usefulness of the methods we have computed examples. We used as systems analog computer programs and analyzed the functions registered by a plotter (graphics). Besides we used tabulated functions. Thereby had been computed the pairs of signals for predetermined transfer functions. Partially we tabulated with high accuracy in order to watch the implications of the truncation errors.

3.1. Method of the inhomogeneous integral equation

We used the procedure described in 1.2. and the numerical process of integration described in chapter 2, and the following scalar product (eq. (26)). Thereby must be $2p + 1$

$$(z,\hat{z}) = \sum_{k=1}^{2p} z(k\underline{dt})\,\hat{z}(k\underline{dt}) \qquad (26)$$

the number of elements of the basic set of points. One gives up the following parameters to the program: order n , incrementation parameter $\underline{dt}$, number p , the values of the input and output signals of the system over the basic set of points. The result will be the coefficients of $Z_M(s)$ (eq. (4)) and $N_M(s)$ (eq. (5)) with $a_0 = 1$.

We analyzed the step response for two different systems of order 2 with p = 20 and $\underline{dt}$ = 0,1 . The numerical results are:

coefficients	system	model	system	model
a_0	1,0	1,00000	1,0	1,00000
a_1	1,5	1,50002	1,5	1,50002
a_2	0,5	0,50002	0,5	0,50002
b_0	1,0	1,00000	1,0	0,99999
b_1	0,0	0,00002	-1,5	-1,49999
b_2	0,0	0,00000	0,5	0,50002

3.2. Method of the inhomogeneous integral equation
- error square norm -

The procedure described in 1.2. (scalar product of $L^2(0,T)$) had been numerical computed with the following scalar product corresponding to Simpson's numerical procedure for integration:

$$(m,\hat{m}) = \frac{dt}{3}\,(m(0)\hat{m}(0) + m(2p)\hat{m}(2p) + 4\sum_{k=1}^{p} m(2k-1)\hat{m}(2k-1) + 2\sum_{k=1}^{p-1} m(2k)\hat{m}(2k)\ . \quad (27)$$

For this numerical procedure corresponding to the equations (19, 20) we predetermine a lower bound ε for $//m_n//^2=(m_n,m_n)$ and an upper bound $\bar{n}$ for n . The computer stops the computation, if one of the both parameters equals the corresponding bound. Then will be computed the coefficients of the integral equation (8) and under use of the characteristic poly-

nomial the full system (eq. (15)). The coefficients of the representation of $m(t)$ with the aid by the full system (eq. (16)) follow from the solution of the corresponding simultaneous equations. One gives up the following parameters to the program: upper bound $\bar{n}$, lower bound ε, incrementation parameter, the number of elements of basic set of points $2p + 1$. The result will be the order n, the characteristic values of the full system, and the coefficients of eq. (16). With a special program one must compute the coefficients of the transfer function.

Numerical example:
We analyzed an output signal of a system with 51 equidistant elements of the basic set of points, for $n = 3,4,5$ the values (m_n, m_n) had been printed and also the solutions corresponding to the models over the basic set of points. The following table shows a part of the numerical results. The values of the system function are exact, the values of the model functions are rounded.

system function	model functions		
	n=3	n=4	n=5
-0,036	-0,007	-0,081	-0,030
+0,065	+0,053	+0,074	+0,066
+0,121	+0,093	+0,107	+0,118
+0,077	+0,084	+0,077	+0,081
+0,026	+0,038	+0,033	+0,027
-0,003	-0,009	-0,003	-0,008
-0,017	-0,030	-0,020	-0,019
-0,017	-0,018	-0,018	-0,014
-0,007	-0,008	-0,007	-0,003
+0,007	+0,027	+0,006	+0,006
+0,013	+0,027	+0,015	+0,010
+0,011	+0,011	+0,017	+0,011
+0,007	-0,006	+0,014	+0,008
+0,004	-0,014	+0,008	+0,005
+0,002	-0,009	+0,003	+0,003
+0,001	+0,000	-0,002	+0,000
+0,000	+0,009	-0,004	+0,000
+0,000	+0,010	-0,004	-0,000

$(m_3, m_3) = 23\ 10^{-8}$

$(m_4, m_4) = 18\ 10^{-9}$

$(m_5, m_5) = 56\ 10^{-13}$

3.3. Method of the homogeneous integral equation
– Tschebyscheff norm –

The procedure described in 1.4. had been programmed in the case of the discrete Tschebyscheff approximation with the norm

$$//m// = \max_{(t_i)} \left(p(t_i) / m(t_i)/\right) \quad (i=1,\ldots,q) \; . \qquad (28)$$

The number of elements of the basic set of points q must satisfy the condition $q > n$ in eq. (28). Corresponding to the general theory[2] one determined algorithm for computing the point of the best approximation. The coordinates of this point and the corresponding value of the norm are the output of the program. This method is in principle analogous to that in 3.2. Therefore we show the efficiency of the method on the basis of the second step in the following example.

Numerical example:
The transfer function of the system must be

$$F_S(s) = \frac{1 + 8s}{(1+s)(1+2s)(1+4s)} \; . \qquad (29)$$

The step response of the system is a non monotonic function. With a program for an analog computer we registered the step response $h(t)$ with the aid by a plotter. We analyzed $m(t) = h(t) - 1$ and put $p(t) = 1$. As the full system by Tschebyscheff had been used

$$g_1(t) = e^{-t} \; , \quad g_2(t) = e^{-t/2} \; , \quad g_3(t) = e^{-t/4} \; . \qquad (30)$$

The number of elements of the basic set of points was 60 and we put the incrementation parameter $\underline{dt} = 0{,}27$. We computed with the program a maximal deviation for the solution of 0,01. The corresponding transfer function is

$$F_M(s) = \frac{1 + 7{,}7s - 0{,}33s^2 - 0{,}08s^3}{(1+s)(1+2s)(1+4s)} \qquad (31)$$

4. Further problems

The following problems will be treated: influence of the errors on the accuracy of the determination of models, systems and models with transcendental transfer functions, approximations of transcendental transfer functions to rational transfer functions. Further remarks will follow at the Congress.

References

1 Achieser, N. I.: Vorlesungen über Approximationstheorie. Berlin: Akademie-Verlag 1953.

2 Collatz, I.: Funktionalanalysis und numerische Mathematik. Berlin-Göttingen-Heidelberg: Springer Verlag 1964.

3 Föllinger, O., und O. Schneider: Lineare Übertragungssysteme. Berlin: Mitteilungen aus dem Institut für Automation der AEG 1961.

4 McDanough, R. N.: Matched exponents for the representation of signals. Dissertation. Baltimore: The John Hopkins University 1963.

5 Schwarze, G.: Algorithmische Ermittlung der Übertragungsfunktion linearer Modelle mit konstanten konzentrierten Parametern für analoge Systeme mit einem Eingang und einem Ausgang durch Analyse der zu charakteristischen Testsignalen gehörigen Ausgangssignale im Zeitbereich. Habilitationsschrift: Magdeburg: Technische Hochschule Otto von Guericke 1967.

6 Zurmühl, R.: Praktische Mathematik für Ingenieure und Physiker. Berlin-Göttingen-Heidelberg: Springer Verlag 1963.

ORDER AND FACTORIZATION OF THE IMPULSE RESPONSE MATRIX

by
C. BRUNI - A. ISIDORI - A. RUBERTI
Istituto di Elettrotecnica dell'Università, ROMA, Italy
Fondazione "Ugo Bordoni", ROMA, Italy

1. Introduction

In this work the problem is considered of evaluating the order of a linear continuous or discrete-time system, of finite order, strictly proper, completely controllable and completely observable, described by the impulse response matrix and the problem of factorization of such a matrix.

These problems are tackled using a method of construction of sets of solutions of the class of differential or difference equations associated with the impulse response matrix, starting from the latter. If these sets of solutions contain a fundamental set, the evaluation of the order can be carried out by computing the number of independent solutions. The factorization can be obtained by choosing fundamental set of solutions [1,2,3].

Problems and methods are examined for time invariant and time varying systems as well as for continuous and discrete time systems. In particular, for the time invariant case a theorem is given from which it is possible to deduce the algorithms for calculating the order based respectively on Markov's parameters and on the moments of the impulse response matrix [5,2]. For the time varying case, under some limiting hypothesis, a theorem is given which leads to an algorithm for the calculation of the order which generalizes that of Markov's parameters [1,2]. Under the same hypothesis a third theorem is stated which allows the factorization of the impulse response matrix, in the continuous [3] as well as in the discrete time case.

2. Construction of the solutions

2.1. Case of continuous systems.

A linear continuous system, of finite order, strictly proper, completely controllable and completely observable, can be described, as is known, by the impulse response matrix $\underline{W}(t,\tau)$. This matrix can be factorized in the form:

$$\underline{W}(t,\tau) = \underline{\Psi}(t,t_o)\underline{H}(t_o,\tau) \qquad t \geq \tau \qquad (1)$$

with $\underline{\Psi}(t,t_o)$ and $\underline{H}(t_o,\tau)$ respectively q.n and n.p matrices. In particular there exists always [6] a class of factorizations, said the class of reduced-form factorizations $\{\underline{\Psi}_o(t,t_o);\underline{H}_o(t_o,\tau)\}$ in which n takes on the minimum value n_o; for this class the column of $\underline{\Psi}_o(t,t_o)$ and respectively the rows of $\underline{H}_o(t_o,\tau)$ are sets of linearly independent functions.

The description of the systems considered can be made by means of ordinary linear differential equations; among these of particular importance are the input-state-output and the input-output equations. The first have the form:

$$\begin{aligned} D\underline{x}(t) &= \underline{A}(t)\underline{x}(t) + \underline{B}(t)\underline{u}(t) \\ \underline{y}(t) &= \underline{C}(t)\underline{x}(t) \end{aligned} \tag{2}$$

$\underline{u}(t)$ being the input vector, $\underline{x}(t)$ the state vector, $\underline{y}(t)$ the output vector and D being the differentiation operator with respect to time while the second have the form:

$$\underline{L}(D,t)\underline{y}(t) = \underline{M}(D,t)\underline{u}(t) \tag{3}$$

where $\underline{L}(D,t)$ and $\underline{M}(D,t)$ are respectively q.q and q.p matrices whose elements are polynomials in D.(°)

The number n of the linearly independent solutions of the homogeneous equation associated with (2) or with (3) is given by the dimension of the state vector for (2) and by the degree of the determinant of $\underline{L}(D, t)$ for (3); it constitutes the order of such equations.

With the operator $\underline{W}(t,\tau)$ can be associated equations of different orders. If these are of type (2) there exists always a class of minimum order and this order coincides with the dimensions n_o of the class of reduced form factorization of $\underline{W}(t,\tau)$ [6]. If the equations are of type (3), there exists a class of order n_o, if and only if the system described by $\underline{W}(t,\tau)$ is differentially observable.(°°) [1]

For the purpose of calculating the order of a system starting from $\underline{W}(t,\tau)$ the authors prefer to keep in mind the natural interpretation of the order as the number of linearly independent solutions of the class of homogeneous equations of minimum order associated with $\underline{W}(t,\tau)$, rather than the interpretation as the dimension of the reduced-form factorization of $\underline{W}(t,\tau)$. Infact it will be shown that it is possible, and in various ways,

---

(°) In a strictly proper system, if (3) is changed into a form in which $\underline{L}(D,t)$ is upper triangular with the degree of the off diagonal elements in each column lower than the degree of the corresponding diagonal element, the degree of each element belonging to the diagonal of $\underline{L}(D,t)$, is higher than the degrees of all the elements of the corresponding row of $\underline{M}(D,t)$.

(°°) The same consideration can be repeated for the equations associated with the adjoint system. In particular, equations of type (3) of order n_o exist for the adjoint system if and only if the given system is differentially controllable.

to construct sets of linearly independent solutions starting from $\underline{W}(t,\tau)$, and therefore deducing the order by means of their computation. Such a principle is also useful for the purposes of factorization of $\underline{W}(t,\tau)$, which can be achieved acting upon the independent solutions.

Thus, the construction of solutions can be achieved by means of the results stated in the following Assertions.

Assertion 1 - Given the impulse response matrix $\underline{W}(t,\tau)$, each function

$$\underline{y}(t) = \int_{-\infty}^{t_o} \underline{W}(t,\tau)\underline{u}(\tau)d\tau \tag{4}$$

where $\underline{u}(t)$ belongs to the space of input functions and is such that integral (4) is defined, is a solution of a homogeneous differential equation of minimum order for the output variable of the given system.

In fact, replacing in (4) a reduced-form factorization of $\underline{W}(t,\tau)$, one has

$$\underline{y}(t) = \underline{\Psi}_o(t,t_o)\int_{-\infty}^{t_o} \underline{H}_o(t_o,\tau)\underline{u}(\tau)d\tau = \underline{\Psi}_o(t,t_o)\underline{z}(t_o) \tag{5}$$

where $\underline{z}(t_o)$ is a constant vector. Since the columns of $\underline{\Psi}_o(t,t_o)$ constitute a fundamental set of solutions for the class of homogeneous equations of minimum order associated with the given system [6], Assertion 1 is justified.

The vector $\underline{z}(t_o)$ represents the state of the system at time t_o as is shown immediately from the comparison between (5) and the expression of the general solution of the homogeneous equation associated with (2)

$$\underline{y}(t) = \underline{C}(t)\underline{\Phi}(t,t_o)\underline{x}(t_o) \tag{6}$$

This interpretation makes it possible to specify that the linear operation defined by the integral which appears in (5) maps the space of the input functions, on the interval $(-\infty,t_o)$, into the state space, at time t_o.

Consider now a set of $N > n_o$ solutions $\{\underline{y}_i(t)\}$ constructed from a set of N inputs $\{\underline{u}_i(t)\}$ according to (4). Introducing the matrices

$$\underline{Y}_N(t) = \left[\underline{y}_1(t) \ \ldots \ \underline{y}_N(t)\right] \tag{7}$$

$$\underline{U}_N(t) = \left[\underline{u}_1(t) \ \ldots \ \underline{u}_N(t)\right] \tag{8}$$

it can be written:

$$\underline{Y}_N(t) = \int_{-\infty}^{t_o} \underline{W}(t,\tau)\underline{U}_N(\tau)d\tau \tag{9}$$

Taking into consideration the latter and keeping in mind what has been observed in connection with Assertion 1, it may be stated:

Assertion 2 - A necessary and sufficient condition for the matrix $\underline{Y}_N(t)$ to contain n_o linearly independent solutions is

that the input matrix $\underline{U}_N(t)$ be such that the rank of

$$\int_{-\infty}^{t_o} \underline{H}_o(t_o,\tau)\underline{U}(t)d\tau = \left[\underline{x}_1(t_o) \ \ldots \ \underline{x}_N(t_o)\right] \tag{10}$$

is n_o that is that the set of states $\{\underline{x}_i(t_o)\}$ into which the input matrix is mapped contain a set of n_o linearly independent states

Having established the procedure for constructing solutions starting from $\underline{W}(t,\tau)$, it is now necessary to find input matrices capable of giving forth sets of solutions which contain fundamental sets. This is done in the following sections, separately for time invariant and time varying cases. Once this problem has been resolved, the properties of linear independence of the solutions constructed by means of (9) must be examined. For this purpose it is useful to consider the properties of the solutions of homogeneous linear differential equations, and particularly those of their Wronskian.

In this connection consider the homogeneous differential equation of order n, associated with (3)

$$\underline{L}(D,t)\underline{y}(t) = \underline{0} \tag{11}$$

and a corresponding triangular form

$$\begin{bmatrix} \sum_{0}^{\nu_1} a_{11}^{(h)}(t)D^h & \sum_{0}^{\nu_2-1} a_{12}^{(h)}(t)D^h & \ldots & \sum_{0}^{\nu_q-1} a_{1q}^{(h)}(t)D^h \\ 0 & \sum_{0}^{\nu_2} a_{22}^{(h)}(t)D^h & \ldots & \sum_{0}^{\nu_q-1} a_{2q}^{(h)}(t)D^h \\ \cdot & \cdot & \ldots & \cdot \\ 0 & 0 & & \sum_{0}^{\nu_q} a_{qq}^{(h)}(t)D^h \end{bmatrix} \underline{y}(t) = \underline{0} \tag{12}$$

In reference to these notations may be stated the following theorem which generalizes the well known result of the scalar case.

Theorem 1 - Given the class of homogeneous linear differential equations of type (1), which may be transformed into the triangular form (12) with

$$a_{ii}^{(\nu_i)}(t) \neq 0 \qquad \forall i, \forall t \tag{13}$$

and which admit the same fundamental matrix of solutions $\underline{Y}(t)$, the Wronskian matrix

$$\left[\underline{Y}^T(t) \quad D\underline{Y}^T(t) \quad \ldots \quad D^{\nu-1}\underline{Y}^T(t)\right] \tag{14}$$

has rank n for every t, provided ν is not less than a fixed minimum value ν_o which depends upon the considered class of equations and satisfies the constraints

$$\frac{n}{q} \leqslant \nu_o \leqslant n \tag{15}$$

2.2. Case of discrete-time systems

The results obtained in the preceding section for continuous systems remain substantially valid also for discrete-time systems. From a formal point of view it is necessary to introduce the discrete variables k, j in place of the continuous ones t, τ and the advance operator E in place of the differentiation operator D. Thus the various results can be stated in terms of discrete-time, and only in some cases variations and specifications will be necessary.

Regarding (1), one obtains for the discrete-time case:

$$\underline{W}(k,j) = \underline{\Psi}(k,k_o)\underline{H}(k_o,j) \qquad k \geq k_o > j \tag{1'}$$

Regarding Assertion 1 in place of (4) one obtains:

$$\underline{y}(k) = \sum_{-\infty}^{k_o-1} \underline{W}(k,j)\underline{u}(j) \tag{4'}$$

Regarding Assertion 2 in place of (9) and (10) one obtains:

$$\underline{Y}(k) = \sum_{-\infty}^{k_o-1} \underline{W}(k,j)\underline{U}_N(j) \tag{9'}$$

$$\sum_{-\infty}^{k_o-1} \underline{H}_o(k_o,j)\underline{U}(j) = \left[\underline{x}_1(k_o) \ \ldots \ \underline{x}_N(k_o)\right] \tag{10'}$$

Regarding Theorem 1 the class of difference equations must be considered:

$$\underline{L}(E,k)\underline{y}(k) = \underline{0} \tag{11'}$$

transformable into the triangular form analogous to (12) with:

$$a_{ii}^{(\nu_i)}(k) \neq 0 \qquad a_{ii}^{(o)}(k) \neq 0 \qquad \forall i, \ \forall k \tag{13'}$$

Beside the Casoratian matrix

$$\left[\underline{y}^T(k) \quad E\underline{y}^T(k) \quad \ldots \quad E^{\nu-1}\underline{y}^T(k)\right] \tag{14'}$$

corrisponding to (14) it is useful, as is herein shown, to introduce the following matrix:

$$\left[\underline{y}^T(k) \quad \Delta\underline{y}^T(k) \quad \ldots \quad \Delta^{\nu-1}\underline{y}^T(k)\right] \tag{14"}$$

in which $\Delta = E - 1$. This matrix has the same properties as (14') in regard to the rank being a linear non-singular transformation of the latter.

3. Order of time invariant systems

3.1. Construction of a fundamental set of solutions.

Consider the functions

$$\gamma_i(t) = \begin{cases} D^i \delta(t) & i \geq 0 \\ -\dfrac{t^{-i-1}}{(-i-1)!} & i < 0 \end{cases} \tag{16}$$

where $\delta(t)$ is the unit impulse, and consider the input matrix

$$\underline{U}_N(t) = \left[\gamma_m(t)\underline{I}_p \quad \gamma_{m+1}(t)\underline{I}_p \quad \dots \quad \gamma_{m+\mu-1}(t)\underline{I}_p\right] \tag{17}$$

in which $\underline{I}_p$ is the unit p.p matrix. The following theorem may be stated in relation to this matrix:

Theorem 2 - For a time invariant continuous system, described by the impulse response matrix $\underline{W}(t)$, the input matrix (17), by means of operator (9) with $t_o=0^+$, gives forth an output matrix which contains a fundamental set of solutions provided μ is not less than a fixed minimum value μ_o which depends on $\underline{W}(t)$, and satisfies constraints

$$\frac{n_o}{p} \leqslant \mu_o \leqslant r_o \tag{18}$$

where r_o is the degree of the minimal polynomial associated with the system.

m can be given:

a) the value 0 for any $\underline{W}(t)$;
b) any positive integer value if no element of $\underline{W}(t)$ takes on a polinomial form;
c) any positive or negative integer value if $\underline{W}(t)$ is strictly stable.

Proof: placing (17) into (10) one has:

$$\int_{-\infty}^{0^+} \underline{H}_o(0,\tau)\underline{U}_N(\tau)\,d\tau = \int_{-\infty}^{0^+} e^{-\underline{A}\tau}\,\underline{B}\underline{U}_N(\tau)\,d\tau = \tag{19}$$
$$= \underline{A}^m\left[\underline{B} \quad \underline{AB} \quad \dots \quad \underline{A}^{\mu-1}\underline{B}\right]$$

With reference to Assertion 2 it is necessary to examine the rank of the latter matrix. In this connection it can be seen that matrix $\left[\underline{B} \quad \underline{AB} \quad \dots \quad \underline{A}^{\mu-1}\underline{B}\right]$ is the Wronskian matrix of $\underline{H}_o(0,\tau)$ calculated at the origin According to Theorem 1 it has rank n_o because $\underline{H}_o(0,\tau)$ is a fundamental set of solutions of the class of homogeneous differential equations associated with the adjoint system [1]. However in this case the well known controllability condition may be used. The inequality (18) replaces (15) because every power of $\underline{A}$ can be expressed as a combination of the first r_o powers starting from the zero power.

The theorem is therefore proved for m=0. For the other values of m it is necessary and sufficient that the integrals appearing in (19) exist and that $\underline{A}^m$ is nonsingular. Thus the conditions specified in b) and c) are derived. Q.E.D.

An analogous theorem may be stated in the case of discrete-time systems by introducing the following functions:

$$\gamma_i(k) = \begin{cases} \Delta^i\delta(k) & i \geqslant 0 \\ -1 & i = -1 \\ -\dfrac{(k-1)(k-2)\dots(k+i-1)}{(-i-1)!} & i < -1 \end{cases} \tag{16'}$$

where $\delta(k)$ is the unit discrete-time impulse. The statement of the theorem is obtained by interpreting in discrete-time terms the various relations considered in the continuous case, replacing $t_o=0^+$ with $k_o=1$.

3.2. Application to continuous systems.

3.2.1. Method of Markov's parameters.

A first algorithm for calculating the order which can be deduced from the general method is based on Markov's parameters of $\underline{W}(t)$ and has been already introduced by B.L. Ho and R.E. Kalman[4]. Defining these parameters by

$$\underline{S}_i = \left[D^i \underline{W}(t)\right]_{t=0^+} \tag{20}$$

and constructing the matrix

$$\underline{\mathcal{H}}(\mu,\nu) = \begin{bmatrix} \underline{S}_0 & \underline{S}_1 & \cdots & \underline{S}_{\mu-1} \\ \underline{S}_1 & \underline{S}_2 & \cdots & \underline{S}_{\mu} \\ \cdot & \cdot & \cdots & \cdot \\ \underline{S}_{\mu-1} & \underline{S}_{\nu} & \cdots & \underline{S}_{\nu+\mu-2} \end{bmatrix} \tag{21}$$

the following Corollary can be stated:

Corollary 1 - The matrix of Markov's parameters $\underline{\mathcal{H}}(\mu,\nu)$ of a given $\underline{W}(t)$ has rank n_o provided μ and ν take on values respectively not less than quantities μ_o and ν_o . The latter depend on $\underline{W}(t)$ and satisfy constraints

$$\frac{n_o}{q} \leq \nu_o \leq r_o \qquad\qquad \frac{n_o}{p} \leq \mu_o \leq r_o \tag{22}$$

Proof: the input matrix (17) with m=0 gives forth, by means of operator (9) with $t_o=0^+$, the output matrix

$$\underline{Y}_N(t) = \left[\underline{W}(t) \quad D\underline{W}(t) \quad \ldots \quad D^{\mu-1}\underline{W}(t)\right] \tag{23}$$

which for Theorem 2 contains a fundamental set of solutions provided $\mu \geq \mu_o$. According to Theorem 1 the Wronskian matrix of $\underline{Y}_N(t)$ has therefore rank n_o for every t provided $\nu \geq \nu_o$ .The upper bound for ν_o , which appears in (22), replaces the one given in Theorem 1. This possibility is based on the time invariance of the system, as shown in the proof of Theorem 2 in relation to μ_o . To complete the proof it is sufficient to observe that $\underline{\mathcal{H}}(\mu,\nu)$ coincides with the Wronskian matrix, calculated for $t_o=0^+$. Q.E.D.

The property stated in Corollary 1 allows the determination of n_o through the calculation of the rank of the matrix of Markov's parameters $\underline{\mathcal{H}}(\mu,\nu)$, carried out for values of μ and ν certainly greater than μ_o and ν_o . Since these are not known a priori it is necessary practically to choose values of μ and ν sufficiently high. Otherwise it is possible to carry out the calculation with gradually increasing values of μ and ν until the rank tends to remain definitely constant.

If instead of $\underline{W}(t)$ its Laplace-transform is known, the

number r_o can be immediately calculated as the degree of the minimum common denominator of the elements of $\underline{W}(s)$. Therefore it is possible according to (22) to determine n_o as the rank of $\underline{\mathcal{H}}(r_o, r_o)$.

3.2.2. Method of the moments.

A second algorithm, already discussed by the authors[5], which can be deduced from the general method is based on the moments of the impulse response matrix. Defining for a strictly stable $\underline{W}(t)$, the moments by

$$\underline{M}_i = \int_o^\infty \underline{W}(t) t^i dt \tag{24}$$

and constructing the matrix

$$\underline{\mathcal{C}}(\mu,\nu) = \begin{bmatrix} \underline{C}_o & \underline{C}_1 & \cdots & \underline{C}_{\mu-1} \\ \underline{C}_1 & \underline{C}_2 & \cdots & \underline{C}_\mu \\ \cdot & \cdot & \cdots & \cdot \\ \underline{C}_{\nu-1} & \underline{C}_\nu & \cdots & \underline{C}_{\mu+\nu-2} \end{bmatrix} \tag{25}$$

where

$$\underline{C}_i = \frac{(-1)^i}{i!} \underline{M}_i \tag{26}$$

it is possible to state the following Corollary:

Corollary 2 - The matrix of the moments $\underline{\mathcal{H}}(\mu,\nu)$ of a given strictly stable $\underline{W}(t)$ has rank n_o provided μ and ν take on values respectively not less than quantities μ_o and ν_o. The latter depende on $\underline{W}(t)$ and satisfy constraints (22).

Proof: the proof is analogous to that of Corollary 1, choosing $m = -(\mu+\nu-1)$ and $t_o = 0$. It is possible in this way to prove for the matrix

$$- \begin{bmatrix} \underline{C}_{\nu+\mu-2} & \underline{C}_{\nu+\mu-3} & \cdots & \underline{C}_{\nu-1} \\ \underline{C}_{\nu+\mu-3} & \underline{C}_{\nu+\mu-4} & \cdots & \underline{C}_{\nu-2} \\ \cdot & \cdot & \cdots & \cdot \\ \underline{C}_{\mu-1} & \underline{C}_{\mu-2} & \cdots & \underline{C}_o \end{bmatrix} \tag{27}$$

the property stated above. Consequently also (25), obtained from (27) by exchanging the order of the rows and the columns and changing the sign of all the elements, has the same properties. Q.E.D.

Also in this case the properties of the matrix of the moments allow the determination of n_o by a similar procedure as the one shown in the preceding section.

3.2.3. Remarks on the preceding methods.

The first remark to be made on the preceding methods is that they depend on particular choices of input matrix (17) defined in Theorem 2. It is clear that with different choices of m other matrices can be obtained, suitable for the calculation of the order. Thus, for example, matrices can be constructed from any sequence of $\mu+\nu-1$ consecutive Markov's parameters or moments pro

vided the conditions stated in Theorem 2 are satisfied. Otherwise mixed matrices can be constructed from both Markov's parameters and moments.

A second consideration important for the application of these methods, concerns the determination of the quantities for constructing the matrices $\mathcal{S}(\mu,\nu)$ and $\mathcal{C}(\mu,\nu)$. In this regard, if $\underline{W}(t)$ is available Markov's parameters and moments can be determined using the definition formulas. If the Laplace transform $\underline{W}(s)$ is available they can be determined taking into account the following expansions:

$$\underline{W}(s) = \sum_{0}^{\infty}{}_{i} \underline{S}_i \, s^{-i-1} \tag{28}$$

$$\underline{W}(s) = \sum_{0}^{\infty}{}_{i} \underline{C}_i \, s^{i} \tag{29}$$

To verify (28) it is sufficient to antitransform term by term (the series being uniformly convergent for every s the modulus of which is greater than the convergence radius) and to observe that the McLaurin expansion of $\underline{W}(t)$ is obtained. Regarding (29) it is sufficient to keep in mind that

$$\left[\frac{d^i \underline{W}(s)}{ds^i}\right]_{s=0} = (-1)^i \int_0^{\infty} \underline{W}(t) t^i dt \tag{30}$$

if $\underline{W}(t)$ is strictly stable.

A third consideration concerns the use of the above mentioned methods for calculating the order starting from noise-affected data. In this case the method of the moments is undoubtedly preferable for the filtering action of the integral operation instead of that of Markov's parameters. This abvious consideration has been confirmed by various calculations based on data of gradually decreasing precision.

3.3. Application to discrete-time systems.

In the discrete-time case an analogous formulation can be given for the two methods developed above. The matrices under consideration are those corresponding to (21) and (25) constructed this time starting from quantities.

$$\underline{S}_i = \left[E^i \underline{W}(k)\right]_{k=1} \tag{20'}$$

and respectively (26), with

$$\underline{M}_i = \sum_{0}^{\infty}{}_{k} \underline{W}(k) k\,(k+1)\ldots(k+i-1) \tag{24'}$$

In regard to the proof of the analogue of Corollary 1, $m=0$ and $k_0=1$ are chosen and the Casoratian matrix of the solution is constructed. Form (14") instead of (14') has been used in order to have operators Δ either with respect to index μ or index ν. In this way it is possible to prove the properties stated in Corollary 1 for a matrix having as its elements the

quantities $\left[\Delta^i \underline{W}(k)\right]_{k=1}$ instead of $\underline{S}_i$. Keeping in mind that $\Delta = E-1$, it is easy to see that the two matrices can be transformed into each other by nonsingular transformations, thus conserving the properties of the rank.

In regard to the analogue of Corollary 2, proceeding in the same way it is possible to prove the properties of the matrix constructed with quantities (24').

The procedure for using the two corollaries and the remarks made for the continuous case can be directly transferred Regarding the expansions (28) and (29) with reference to the z-transform, one obtains:

$$\underline{W}(z) = \sum_{o}^{\infty}{}_i \; \underline{S}_i \; z^{-i-1} \tag{28'}$$

$$\underline{W}(z) = \sum_{o}^{\infty}{}_i \; \underline{C}_i \; (z-1)^i \tag{29'}$$

The verification is obvious if one keeps in mind the definition of the z-transform and respectively, the analogous form of (30):

$$\left[\frac{d^i \underline{W}(z)}{dz^i}\right]_{z=1} = (-1)^i \sum_{o}^{\infty}{}_k \; \underline{W}(k)k(k+1)\ldots(k+i-1) \tag{30'}$$

4. Order of time varying systems.

In the case of continuous time varying systems, the following theorem can be used for the construction of an output matrix which contains a fundamental set of solutions.

Theorem 3 - Given a continuous differentially controllable system, if the class of homogeneous differential equations of minimum order associated with its adjoint system satisfies the conditions (13), the input matrix (17), with m=0, gives forth, through operator (9) with $t=0^+$, an output matrix which contains a fundamental set of solutions provided μ is not less than a fixed minimum value μ_o, which depends on $\underline{W}(t,\tau)$ and satisfies constraints

$$\frac{n}{p}_o \leqslant \mu_o \leqslant n_o \tag{31}$$

Proof: placing (17) into (10), with m=0 and $t_o=0^+$, one obtains the matrix:

$$\int_{-\infty}^{0^+} \underline{H}_o(0,\tau)\underline{U}_N(\tau)d\tau = \left[\underline{H}_o(0,\tau) \; \ldots \; D^{\mu-1}\underline{H}_o(0,\tau)\right]_{\tau=0} \tag{32}$$

$\underline{H}_o(0,\tau)$, as previously mentioned, is a fundamental set of solutions of the class of homogeneous differential equations associated with the adjoint system. Therefore, the r.h.s. matrix of (32) is the Wronskian matrix of $\underline{H}_o(0,\tau)$ calculated at the origin and has rank n_o in the hypothesis of Theorem 1. Constraints (31) are deduced from (15) keeping in mind, that, the dimensions of $\underline{H}_o(0,\tau)$ are $n_o.p$. Due to Assertion 2 the theorem is therefore proven. Q.E.D.

An algorithm for the calculation of the order can be deduced from the preceding theorem and is based on parameters which extend those of Markov in the time varying case. Definining the parameters

$$\underline{S}_{hi} = \left[D_t^h \, D_\tau^i \, \underline{W}(t,\tau) \right]_{t=\tau=0} \tag{33}$$

and constructing the matrix

$$\underline{\mathcal{S}}(\mu,\nu) = \begin{bmatrix} \underline{S}_{00} & \underline{S}_{01} & \cdots & \underline{S}_{0,\mu-1} \\ \underline{S}_{10} & \underline{S}_{11} & \cdots & \underline{S}_{1,\mu-1} \\ \cdot & \cdot & \cdots & \cdot \\ \underline{S}_{\nu-1,0} & \underline{S}_{\nu-1,1} & \cdots & \underline{S}_{\nu-1,\mu-1} \end{bmatrix} \tag{34}$$

the following corollary may be stated:

Corollary 3 - Given a differentially controllable and differentially observable system described by $\underline{W}(t,\tau)$, if the classes of homogeneous differential equations of minimum order associated with such a system and its adjoint satisfy the conditions (13), the matrix $\underline{\mathcal{S}}(\mu,\nu)$ has rank n_o provided μ and ν take on values respectively not less than quantities μ_o and ν_o. The latter depend on $\underline{W}(t,\tau)$ and satisfy constraints

$$\frac{n_o}{p} \leq \mu_o \leq n_o \qquad\qquad \frac{n_o}{q} \leq \nu_o \leq n_o \tag{35}$$

Proof: the proof is analogous to that of Corollary 1. The hypothesis made are necessary and sufficient in order to apply Theorems 1 and 2. In this way it is proven that the properties stated above hold for a matrix which is different from (34) due to the signs of the submatrices constituting the even columns, and consequently for (34) itself. Q.E.D.

A theorem analogous to Theorem 3 can be stated for discrete-time systems. Regarding its validity, it is necessary to assume that the class of homogeneous difference equations of minimum order associated with the adjoint system, satisfy conditions (13'). Regarding the input matrix, it is necessary to consider:

$$\underline{U}_N(k) = \left[\delta(k)\underline{I}_p \quad E\delta(k)\underline{I}_p \quad \ldots \quad E^{\mu-1}\delta(k)\underline{I}_p \right]$$

This gives forth, through operator (9') with $k_o=1$ output matrix $\underline{Y}_N(k)$ which contains a fundamental set of solutions.

Placing $\underline{U}_N(k)$ into (10'), with $k_o=1$, one obtains the matrix

$$\sum_{-\infty j}^{o} \underline{H}_o(1,j)\underline{U}_N(j) = \left[\underline{H}_o(1,j) \; E^{-1}\underline{H}_o(1,j) \; \ldots \; E^{-(\mu-1)}\underline{H}_o(1,j) \right]_{j=0} =$$

$$= \left[E^{\mu-1}\underline{H}_o(1,j) \; \ldots \; E\underline{H}_o(1,j) \; \underline{H}_o(1,j) \right]_{j=-\mu+1}$$

$\underline{H}_o(1,j)$ is a fundamental set of solutions of the class of homogeneous difference equations associated with the adjoint system Therefore, the r.h.s. matrix is the Casoratian matrix of $\underline{H}_o(1,j)$ calculated for $j=-\mu+1$ and has, as in the previous hypothesis, rank n_o. The proof can be completed keeping in mind the analogue of Assertion 2 in the discrete-time case.

An algorithm for the calculation of the order can be deduced from the preceding theorem and is based on the parameters corresponding to those defined in (33). In the discrete time case, these may be written as:

$$\underline{S}_{hi} = \left[E_k^h E_j^{-i} \underline{W}(k,j) \right]_{\substack{k=1 \\ j=0}} \tag{33'}$$

It is easily verified that an analogue of Corollary 3 holds for matrix (34) constructed with quantities (33').

5. Factorization of the impulse response matrix.

The method discussed in section 2 for the construction of a fundamental set of solutions allows also the resolution of the problem of the factorization in reduced form of $\underline{W}(t,\tau)$ according to

$$\underline{W}(t,\tau) = \underline{\Psi}_o(t,t_o)\underline{H}_o(t_o,\tau) \tag{36}$$

It allows in fact the calculation of the matrices $\underline{\Psi}_o(t,t_o)$ and $\underline{H}_o(t_o,\tau)$. This problem is of notable interest because the knowledge of $\underline{\Psi}_o(t,t_o)$ and $\underline{H}_o(t_o,\tau)$ allows the calculation of the state and of the output, starting from any state, while the knowledge of $\underline{W}(t,\tau)$ allows only the calculation of the zero state response. In addition the factorization in the form (36) is the usual starting point for the realization of $\underline{W}(t,\tau)$ according to the matrices $\underline{A}(t)$, $\underline{B}(t)$, $\underline{C}(t)$.

In the case of continuous systems, in the hypothesis of Corollary 3, consider the matrix $\underline{\mathcal{Y}}(\mu,\nu)$, defined in (34) and assign μ and ν two values so that the rank of $\underline{\mathcal{Y}}(\mu,\nu)$ is equal to the order n_o of the system. Let $\underline{\mathcal{Y}}_t$ be a nonsingular $n_o.n_o$ matrix, constructed with rows $i_1, i_2, \ldots i_{n_o}$ and columns $j_1, j_2 \ldots j_{n_o}$ of $\underline{\mathcal{Y}}(\mu,\nu)$.

Consider furthermore matrices

$$\left[\underline{W}(t,\tau) \quad D_\tau \underline{W}(t,\tau) \quad \ldots \quad D_\tau^{\mu-1} \underline{W}(t,\tau) \right]_{\tau=0} \tag{37}$$

$$\left[\underline{W}^T(t,\tau) \quad D_t \underline{W}^T(t,\tau) \quad \ldots \quad D_t^{\nu-1} \underline{W}^T(t,\tau) \right]_{t=0}^T \tag{38}$$

and let $\underline{\Psi}_t(t,t_1)$ and $\underline{H}_t(t_2,\tau)$ be the matrices constructed respectively with columns $j_1, j_2, \ldots j_{n_o}$ of (37) and rows $i_1, i_2, \ldots i_{n_o}$ of (38).

It is now possible to state the following theorem:

Theorem 4 - Given a differentially controllable and differentially observable system described by $\underline{W}(t,\tau)$, if the classes of homogeneous differential equations of minimum order associated with

such a system and its adjoint satisfy the conditions (13), the impulse response matrix admits the two reduced-form factorizations:

$$\underline{\Psi}_r(t,t_1) \; ; \; \underline{\mathcal{G}}_r^{-1} \, \underline{H}_r(t_2,\tau) \tag{39}$$

$$\underline{\Psi}_r(t,t_1)\,\underline{\mathcal{G}}_r^{-1} \; ; \; \underline{H}_r(t_2,\tau) \tag{40}$$

An analogous theorem can be proven for the discrete time case starting from the matrix $\mathcal{G}(\mu,\nu)$ constructed with parameters (33'). The matrices that must be considered instead of (37) and (38) are:

$$\left[\underline{W}(k,j) \quad E_j^{-1}\,\underline{W}(k,j) \quad \ldots \quad E_j^{-(\mu-1)}\,\underline{W}(k,j)\right]_{j=0} \tag{37'}$$

$$\left[\underline{W}^T(k,j) \quad E_k\,\underline{W}^T(k,j) \quad \ldots \quad E_k^{\nu-1}\,\underline{W}^T(k,j)\right]_{k=1}^{T} \tag{38'}$$

References

1. C. BRUNI-A.ISIDORI-A. RUBERTI: The order of linear continuous systems described by the impulse response matrix - Report 46/bis, Istituto di Elettrotecnica dell'Università di Roma, May 1968.

2. C. BRUNI-A.ISIDORI-A. RUBERTI: The order of linear discrete-time systems described by the impulse response matrix - To appear.

3. C. BRUNI-A.ISIDORI-A. RUBERTI: A method of factorization of the impulse-response matrix - Report 50, Istituto di Elettrotecnica dell'Università di Roma, May 1968.

4. B.L. HO-R.E. KALMAN: Effective construction of linear state variable models from input/output functions - Regelungstechnik, Vol. 14, n. 12, p. 545-548, 1966.

5. C. BRUNI-A.ISIDORI-A. RUBERTI: A method of realization based on the moments of the impulse response matrix - Report 40, Istituto di Elettrotecnica dell'Università di Roma, Dec. 1967.

6. L. WEISS-R.E. KALMAN: Contributions to linear system theory - Int. J. Engng. Sci., Vol. 3, p. 141-171, 1965.

C. BRUNI - A. ISIDORI - A. RUBERTI

Order and Factorization of the Impulse-Response Matrix

Page	Row		ERRATA	CORRIGE
	from the top	from the bottom		
4	2		$\int_{-\infty}^{t_o} \underline{H}_o(t_o,\tau)\underline{U}(t)\,d\tau$	$\int_{-\infty}^{t_o} \underline{H}_o(t_o,\tau)\underline{U}_N(\tau)\,d\tau$
5	15		$\underline{Y}(k)$	$\underline{Y}_N(k)$
5	16		$\sum_{-\infty}^{k_o-1} j\ \underline{H}_o(k_o,j)\underline{U}(j)$	$\sum_{-\infty}^{k_o-1} j\ \underline{H}_o(k_o,j)\underline{U}_N(j)$
6		1	$\frac{(k-1)\,(k-2)\,..\,(k+i-1)}{(-i-1)\,!}$	$\frac{(k-1)\,(k-2)\,..\,(k+i+1)}{(-i-1)\,!}$
8	16		$\underline{\mathcal{L}}(\mu,\nu)$	$\underline{\mathcal{C}}(\mu,\nu)$
10		16	$t=0^+$	$t_o=0^+$