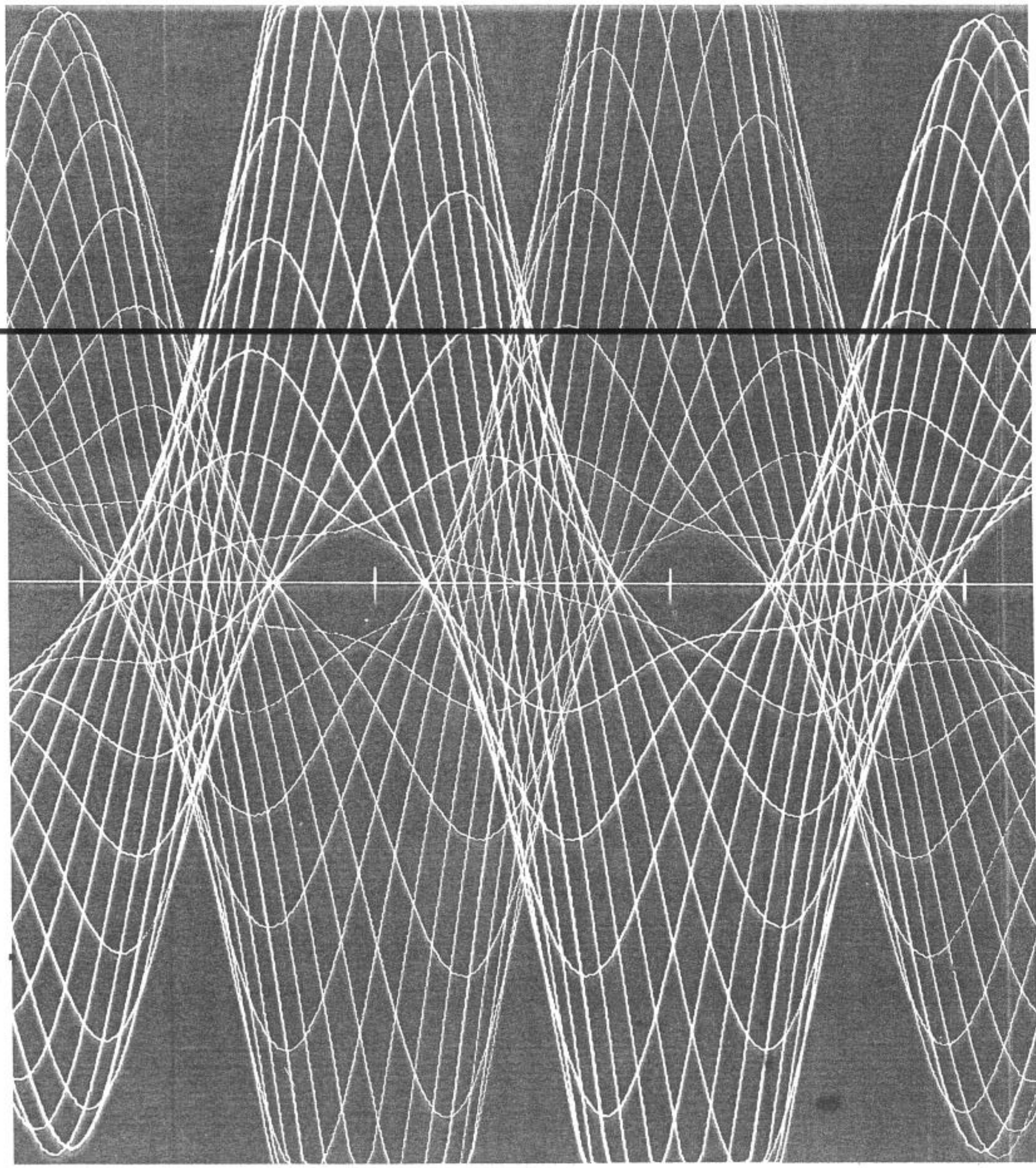

RJ 375

CANONICAL FORMS FOR LINEAR SYSTEMS - I

W. G. Tuel, Jr.

March 14, 1966

IBM RESEARCH

CANONICAL FORMS FOR LINEAR SYSTEMS - I

W. G. Tuel, Jr.

IBM San Jose Research Laboratory
San Jose, California

ABSTRACT: A canonical representation for multivariable linear control systems is derived from the concepts of controllability. An analogous result is obtained for the optimal estimation problem, using the concept of observability. A constructive method of obtaining the appropriate canonical form from any linear time-invariant state-variable system description is presented. For the single input case, the method yields the well-known companion normal form of the system matrix.

Research Paper
RJ-375
March 14, 1966

I. INTRODUCTION

In many, if not most, industrial processes, it is nearly impossible to construct models based on physical characteristics which describe the behavior of the plant adequately. Hence, one must fit an approximate model, consisting of an assumed structure and its associated parameters, to the observed characteristics of the system. Naturally, it is desirable to associate the fewest possible parameters with the model structure so that the identification task is minimized. For instance, if one assumes the plant to be fourth order with two inputs, it is well to know that only nine parameters completely describe the model, not twenty-four, as might first be supposed. In addition, one would hope that the description of the system in a minimal form would shed light on the internal mathematical structure of the process.

An analogous situation arises in the selection of a suitable performance index for control purposes. In a great majority of cases, the mathematical criterion is chosen to reflect certain subjective judgements on system performance. In these instances, it is clearly useful to specify the performance index with the fewest possible number of parameters, to reduce the number of alternatives confronting the designer.

Thus, the problem is to find minimal (canonical) forms for system description and performance specification. This paper is devoted to a study of these canonical forms for a special, but important, case--the linear regulator problem for a constant coefficient discrete system with quadratic performance index. Specifically, the class of systems considered may be represented in the form

$$\underline{x}(k+1) = A\underline{x}(k) + B\underline{u}(k) + C\underline{v}(k)$$
$$\underline{z}(k) = H\underline{x}(k) + G\underline{w}(k), \qquad (1.1)$$

with the performance evaluated by

$$J = \lim_{N\to\infty} E\left\{\frac{1}{N}\sum_{i=1}^{N}\left[\underline{x}'(i)Q\underline{x}(i) + \underline{u}'(i-1)R\underline{u}(i-1)\right]\right\} \qquad (1.2)$$

where $\underline{x}(k)$ is an n-dimensional state vector; $\underline{u}(k)$ is an r-dimensional control vector; $\underline{z}(k)$ is an m-dimensional observation vector; $\underline{w}(k)$ and $\underline{v}(k)$ are m- and n-dimensional white Gaussian noise processes with

$$E[\underline{v}(k)\underline{v}'(j)] = I\delta_{kj}; \quad E[\underline{w}(k)\underline{w}'(j)] = I\delta_{kj}; \quad E[\underline{v}(k)\underline{w}'(j)] = 0.$$

E denotes statistical expectation over the distribution of $\underline{v}$ and $\underline{w}$.

Equation (1.1) may describe the incremental behavior of an industrial process near a desired operating point. As is well known,[1] the optimum control policy $\underline{u}(k)$ is given by

$$\underline{u}(k) = -K\hat{\underline{x}}(k), \qquad (1.3)$$

where K is a constant feedback gain matrix, and $\hat{\underline{x}}(k)$ is the minimum variance estimate of $\underline{x}(k)$ obtained from the observations $\underline{z}(k)$, $\underline{z}(k-1)$,

... $\underline{z}(0)$ by recursive estimation, using what is commonly known as the Kalman filter. It has previously been shown that the control and estimation problems may be split and solved separately.[1] In fact, they are mathematical duals, so that any results derived for one problem may be applied to the other. This property will be exploited in the following presentation.

This report is divided into two parts which are essentially self-contained. This paper, Part I, contains the development of particularly useful canonical forms for the system (1.1). These forms display the internal mathematical structure of the process as interconnected sub-systems. In Part II entitled "Canonical Forms for Linear Systems-II," these canonical forms are used to simplify the control and estimation equations. It will be shown that there are classes of systems having the same Kalman filter or feedback control law; a characterization of these equivalence classes will be derived. Finally, these results will be used to obtain reduced expressions for solving the control and estimation problems.

II. CANONICAL REPRESENTATIONS OF SYSTEMS

In order to study the control and estimation problems in detail, it is convenient to be able to describe a system in a canonical form which displays the internal mathematical structure of the process. (This is not to be confused with the internal structure of the physical process, which is generally complex and largely unknown.) The standard forms developed in this part will play a central role in the analyses in Part II, much as the canonical Jordan form is a primary tool in matrix theory.

A system may be expressed in two different ways, depending upon whether the control or estimation problem is being considered. These possibilities are embodied in the two theorems stated below--the first for the control problem; the second, the dual theorem, for the estimation problem.

Theorem I. Given the system (1.1), assume that the related process

$$\underline{x}(k+1) = A\underline{x}(k) + B\underline{u}(k) \tag{2.1}$$

has the following properties:

a. B is of full rank $(=r)$.

b. The system is completely controllable.

Then there exists an alternative set of states, $\underline{\xi}(k)$, related to $\underline{x}(k)$ by the non-singular transformation

$$\underline{x}(k) = T_1\underline{\xi}(k)$$

for which

$$\underline{\xi}(k+1) = \hat{A}\underline{\xi}(k) + \hat{B}\underline{u}(k) \tag{2.2}$$

has the following properties.

The matrices $\hat{A}$, $\hat{B}$ may be partitioned into

$$\hat{A} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \qquad \hat{B} = \begin{bmatrix} 0 \\ \hat{B}_2 \end{bmatrix} \tag{2.3}$$

where

1. $\hat{B}_2$ is a non-singular, $r \times r$ upper triangular matrix with $(\hat{B}_2)_{ii} = 1, \; i = 1, 2, \ldots, r$.
2. The matrices $\hat{A}_{21}$ and $\hat{A}_{22}$ are $r \times (n-r)$ and $r \times r$, respectively, containing all the parameters of the canonic form.
3. There exists a set of r positive integers $\{p_i\}$ defined by the structure of the system, such that

$$\sum_{i=1}^{r} p_i = n.$$

The method of obtaining these p_i's is given in the proof.

4.

$$\hat{A}_{11} = \begin{bmatrix} Z_1 & 0 & \cdots & 0 \\ 0 & Z_2 & & \\ \vdots & 0 & \ddots & \\ 0 & & 0 & Z_r \end{bmatrix} \tag{2.4}$$

$$\hat{A}_{12} = \begin{bmatrix} \underline{e}_1 & \underline{0} & \cdots & \underline{0} \\ \underline{0} & \underline{e}_2 & & \\ \vdots & & \ddots & \\ \underline{0} & & & \underline{e}_r \end{bmatrix} \tag{2.5}$$

where $\hat{A}_{11}$ is $(n-r) \times (n-r)$; $\hat{A}_{12}$ is $(n-r) \times r$; Z_i is a $(p_i-1) \times (p_i-1)$ matrix of the form

$$Z_i = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & & \ddots & \ddots & \\ \vdots & & & \ddots & 1 \\ 0 & & & 0 & 0 \end{bmatrix}$$

and $\underline{e}_i$ is a (p_i-1) column vector of the form

$$\underline{e}_i = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

If some $p_i = 1$, the corresponding terms are deleted from the arrays.

For example, the matrices

$$\hat{A} = \left[\begin{array}{ccccc|cc} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \hline -.3 & .1 & .8 & .7 & 0. & .1 & .01 \\ .51 & 0. & 1.2 & .68 & -.5 & .4 & -2. \end{array}\right]$$

$$\hat{B} = \left[\begin{array}{cc} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ \hline 1 & .2 \\ 0 & 1 \end{array}\right] \qquad (p_1 = 4, \; p_2 = 3)$$

are in canonical form. Hereafter, the form (2.2) will be referred to as the <u>control canonic form</u>.

Theorem II: Given the system (1.1), assume that the related process

$$\underline{x}(k+1) = A\underline{x}(k) \qquad (2.6)$$

$$\underline{z}(k) = H\underline{x}(k)$$

has the following properties:

a. H is of full rank (= m).

b. The system is completely observable.

Then there exists an alternative set of states, $\underline{\eta}(k)$, related to $\underline{x}(k)$ by the non-singular transformation

$$\underline{\eta}(k) = T_2\underline{x}(k)$$

for which

$$\underline{\eta}(k+1) = \widetilde{A}\underline{\eta}(k); \quad \underline{z}(k) = \widetilde{H}\underline{\eta}(k). \qquad (2.7)$$

The matrices $\widetilde{A}$, $\widetilde{H}$ may be partitioned into

$$\widetilde{A} = \begin{bmatrix} \widetilde{A}_{11} & \widetilde{A}_{12} \\ \widetilde{A}_{21} & \widetilde{A}_{22} \end{bmatrix}, \qquad \widetilde{H} = \begin{bmatrix} \widetilde{H}_1 & 0 \end{bmatrix}. \qquad (2.8)$$

where

1. $\widetilde{H}_1$ is a non-singular $m \times m$ lower traingular matrix with $(\widetilde{H}_1)_{ii} = 1, \quad i = 1, 2, \ldots, m.$

2. The matrices $\widetilde{A}_{11}$ and $\widetilde{A}_{21}$ are $m \times m$ and $(n-m) \times m$, respectively. They contain all the parameters of the canonic form.

3. There exists a set of m positive integers $\{q_i\}$, defined by the structure of the system, such that

$$\sum_{i=1}^{m} q_i = n.$$

The method of obtaining these q_i's is given in the proof.

4. $$\widetilde{A}_{22} = \begin{bmatrix} Z_1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & Z_2 & & & & \cdot \\ \cdot & & \cdot & & & \cdot \\ \cdot & & & \cdot & & \cdot \\ \cdot & & & & \cdot & 0 \\ 0 & \cdot & \cdot & \cdot & & Z_m \end{bmatrix} \qquad (2.9)$$

$$\widetilde{A}_{12} = \begin{bmatrix} g'_1 & \underline{0}' & \cdot & \cdot & \cdot & \underline{0}' \\ \underline{0}' & g'_2 & & & & \cdot \\ \cdot & & \cdot & & & \cdot \\ \cdot & & & \cdot & & \cdot \\ \cdot & & & & \cdot & \underline{0}' \\ \underline{0}' & \cdot & \cdot & \cdot & & g'_m \end{bmatrix} \qquad (2.10)$$

where $\widetilde{A}_{22}$ is $(n-m) \times (n-m)$; $\widetilde{A}_{12}$ is $m \times (n-m)$; Z_i is of dimension $(q_i-1) \times (q_i-1)$, and of the form of Theorem I; and g'_i is a (q_i-1) row vector of the form

$$g'_i = [1, 0, \ldots, 0].$$

If some $q_i = 1$, the corresponding terms are deleted from the arrays. This form will be known as the <u>observation canonic form.</u>

The proofs of these theorems are straightforward and constructive, although lengthy. To illustrate the results obtainable, the following examples are presented before proceeding with the proof.

Example 1: Consider the system (1.1) with

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & .1 & 1 & 0 \\ .5 & 0 & 0 & 1 \\ 0 & 0 & .2 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad H = [1\ 0\ 0\ 0].$$

For this system $p_1 = p_2 = 2$, and the control canonic form becomes

$$\hat{A}_{21} = \begin{bmatrix} .2 & 0 \\ 1.0 & 0 \end{bmatrix}, \qquad \hat{A}_{22} = \begin{bmatrix} 0 & .5 \\ 0 & 0.1 \end{bmatrix}$$

$$Z_1 = [0], \quad Z_2 = [0], \quad \underline{e}_1 = [1], \quad \underline{e}_2 = [1].$$

Thus,

$$\hat{A}_{11} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \qquad \hat{A}_{12} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \text{and } \hat{B} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Hence,

$$\underline{\xi}(k+1) = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ .2 & 0 & 0 & .5 \\ 1.0 & 0 & 0 & 0.1 \end{bmatrix} \underline{\xi}(k) + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \underline{u}(k).$$

The transformation T_1 is given by

$$T_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -.5 & 1 & 0 \end{bmatrix}.$$

For the observation canonic form, $q_1 = 4$,

$$\tilde{A}_{11} = [0.1]; \quad \tilde{A}_{21} = \begin{bmatrix} 0.2 \\ 0.48 \\ 0 \end{bmatrix};$$

$$Z_1 = \tilde{A}_{22} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}; \quad \text{and } \tilde{A}_{12} = [1\ \ 0\ \ 0] = g_1'.$$

$$\tilde{H} = [1\ \ 0\ \ 0\ \ 0].$$

Thus,

$$\underline{\eta}(k+1) = \begin{bmatrix} 0.1 & 1 & 0 & 0 \\ 0.2 & 0 & 1 & 0 \\ 0.48 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \underline{\eta}(k)$$

$$z(k) = [1\ \ 0\ \ 0\ \ 0]\,\underline{\eta}(k).$$

Here,

$$T_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -.1 & 1 & 0 & 0 \\ -.2 & 0 & 1 & 0 \\ .02 & -.2 & 0 & 1 \end{bmatrix}.$$

This example is worked out in detail in Appendix I, following the procedure outlined in the proofs of the theorems.

Example 2: Let

$$A = \begin{bmatrix} .7 & 1 & 0 & 0 & 0 & 0 \\ 0 & .1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1. & 0 & 1 & 0 \\ 0 & 0 & 0 & .5 & 0 & 0 \\ .2 & 0 & 0 & 0 & -.1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 2. & 0 \end{bmatrix},$$

and consider the problem of obtaining the control canonical form. (This example will also be worked out in Appendix I.) The sub-partition indices become $p_1 = 1$, $p_2 = 3$, $p_3 = 2$, and the transformation T_1 is found to be

$$T_1 = \begin{bmatrix} -.2 & 1. & 1. & 0 & 0 & 0 \\ .14 & -.9 & -.7 & 0 & 1 & 1 \\ 2. & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ -2. & 2. & 0 & 0 & 0 & 0 \\ -.16 & -2. & -.2 & 0 & 2. & 0 \end{bmatrix} .$$

The resultant canonic form is

$$A = \left[\begin{array}{ccc|ccc} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & .5 & 0 & 0 \\ 0 & .08 & 0 & 0 & 1. & .1 \\ .014 & -.31 & -.07 & 0 & 0 & .7 \end{array}\right] .$$

Notice that the terms Z_1, $\underline{e}_1$ are deleted from $\hat{A}_{11}$ and $\hat{A}_{12}$, since $p_1 = 1$.

The proofs of Theorems I and II will now be given. The principal tools are the controllability and observability conditions.

III. PROOF OF THEOREM I

Assumption (b) (controllability) is satisfied if, and only if,[2]

$$\text{rank } \{B, AB, \ldots, A^{n-1} B\} = n.$$

Define separately the columns of B by the notation $B = \{\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_r\}$. These vectors are linearly independent, since B is of full rank by assumption (a).

Consider the rectangular array whose elements are the vectors $A^j\underline{b}_i$, i.e., the array

$$\begin{array}{ccccccccc} \underline{b}_1 & A\underline{b}_1 & \cdots & & A^{p_1-1}\underline{b}_1 & A^{p_1}\underline{b}_1 & \cdots & A^n\underline{b}_1 \\ \underline{b}_2 & A\underline{b}_2 & \cdots & & \cdot & \cdot & & \cdot \\ \cdot & \cdot & A^j\underline{b}_{k-1} & \cdots & \cdot & \cdot & & \cdot \\ \cdot & \cdot & A^j\underline{b}_k & A^{j+1}\underline{b}_k & \cdots & A^{p_k}\underline{b}_k & \cdots & \\ \underline{b}_r & A\underline{b}_r & \cdots & & \cdots & & & A^n\underline{b}_r \end{array} \qquad (3.1)$$

It will be shown that it is always possible to separate the array into a left-hand and right-hand side in such a way that the elements (vectors) on the left-hand side form a basis for the n-dimensional space E_n. The dashed line in equation (3.1) is an illustration of such a separating line.

Considering the vectors in the order $\underline{b}_1$; $\underline{b}_2$; ...; $\underline{b}_r$; $A\underline{b}_1$; ...; $A\underline{b}_r$; $A^2\underline{b}_1$; ...; $A^n\underline{b}_r$, define a basis with the first n that are linearly independent. It is always possible to find n independent vectors, since the set $\{A^j\underline{b}_i\}$ spans the n-dimensional space E_n by the controllability assumption. It is now necessary to show that the vectors thus obtained are bunched on the left side of the array.

Since by hypothesis B is of full rank, the vectors $\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_r$ are members of the basis. Now in each row of (3.1), denoted here by k, (k=1, 2, ..., r), there exists an index p_k, where $1 \le p_k \le n-r+1$, such that the vectors

$$\underline{b}_k;\ A\underline{b}_k;\ \ldots;\ A^{p_k-1}\underline{b}_k$$

are all members of the basis, but that $A^{p_k}\underline{b}_k$ is not. (Since there are r members of the basis in the first column, there can be at most n - r + 1 basis vectors in any row.) Since $A^{p_k}\underline{b}_k$ is not a basis vector, by construction it must be a linear combination of the previously considered elements of the array; i.e.,

$$A^{p_k}\underline{b}_k = \sum_{i=1}^{r}\sum_{j=0}^{p_k-1} \alpha_{ij}(k)A^j\underline{b}_i + \sum_{i=1}^{k-1} \alpha_{i,p_k}(k)A^{p_k}\underline{b}_i \ , \quad (3.2)$$

for some set of $\alpha_{ij}(k)$, not necessarily unique.

Consider the next element in the row, the vector

$$A^{p_k+1}\underline{b}_k \ .$$

From (3.2),

$$A^{p_k+1}\underline{b}_k = \sum_{i=1}^{r}\sum_{j=0}^{p_k-1} \alpha_{ij}(k)A^{j+1}\underline{b}_i + \sum_{i=1}^{k-1} \alpha_{i,p_k}(k)A^{p_k+1}\underline{b}_i \ . \quad (3.3)$$

Thus, $A^{p_k+1}\underline{b}_k$ is also a linear combination of previously considered elements and cannot be in the basis. By induction, the rest of the row is similarly dependent. Thus, the array of equation (3.1) may be partitioned so that the basis vectors lie on the left-hand side of the separating line, as shown. As is readily seen, the element $A^{p_k}\underline{b}_k$ is the first entry in the k^{th} row to the right of the separation line.

The independent basis vectors may be written

$$\{A^j\underline{b}_i\}\ , \quad i = 1, 2, \ldots, r, \quad j = 0, 1, \ldots, p_i-1$$

Clearly,

$$\sum_{i=1}^{r} p_i = n \ ,$$

since there are precisely n independent vectors. The $\{p_i\}$ can be interpreted as the dimensions of the subsystems of the process.

Equation (3.2) may be written more compactly as

$$A^{p_k}\underline{b}_k = \sum_{j=0}^{p_k}\sum_{i=1}^{r} \alpha_{ij}(k)A^j\underline{b}_i, \quad k = 1, 2, \ldots, r \ , \qquad (3.4)$$

where

$$\alpha_{ij}(k) = 0, \qquad j > p_k$$

$$\alpha_{ip_k}(k) = 0, \quad i \ge k, \qquad (3.5)$$

since the basis vectors are bunched on the left-hand side of the array.

For any $m = 0, 1, 2, \ldots, p_k$, equation (3.4) may be written

$$A^{p_k-m}\left[A^m\underline{b}_k - \sum_{j=p_k-m}^{p_k} A^{j-p_k+m}\left(\sum_{i=1}^{r}\alpha_{ij}(k)\underline{b}_i\right)\right] =$$

$$= \sum_{j=0}^{p_k-m-1} A^j\left(\sum_{i=1}^{r} \alpha_{ij}(k)\underline{b}_i\right), \quad k = 1, 2, \ldots, r . \tag{3.6}$$

The bracket on the left defines a vector

$$\underline{v}_{m,k} = A^m \underline{b}_k - \sum_{j=p_k-m}^{p_k} A^{j-p_k+m}\left(\sum_{i=1}^{r} \alpha_{ij}(k)\underline{b}_i\right). \tag{3.7}$$

From equation (3.7) it is easily shown that

$$\underline{v}_{m,k} = A\underline{v}_{m-1,k} - \sum_{i=1}^{r} \alpha_{i,p_k-m}(k)\underline{b}_i \tag{3.8}$$

for $m = 1, 2, \ldots, p_k$, while

$$\underline{v}_{0,k} = \underline{b}_k - \sum_{i=1}^{r} \alpha_{i,p_k}(k)\underline{b}_i . \tag{3.9}$$

By comparing equations (3.7) and (3.4), it is clear that $\underline{v}_{p_k,k} = 0$, so that from equation (3.8)

$$A\underline{v}_{p_k-1,k} = \sum_{i=1}^{r} \alpha_{i,0}(k)\underline{b}_i . \tag{3.10}$$

What will be shown is that the matrix of vectors $\{\underline{v}_{ij}\}$, when arranged in a certain order, is the required transformation T_1 given in the statement of the theorem.

Using (3.5), equation (3.9) defines the matrix equation

$$[\underline{v}_{01}, \underline{v}_{02}, \ldots, \underline{v}_{0r}] = [\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_r] \begin{bmatrix} 1 & -\alpha_{1,p_2}(2) & \cdot & \cdot & \cdot & -\alpha_{1,p_r}(r) \\ 0 & 1 & -\alpha_{2,p_3}(3) & & & \cdot \\ 0 & 0 & \cdot & & & \cdot \\ \cdot & & & \cdot & & \cdot \\ \cdot & & & & 1 & -\alpha_{r-1,p_r}(r) \\ 0 & \cdot & \cdot & \cdot & 0 & 1 \end{bmatrix} \tag{3.11}$$

which may be written

$$V_0 = B\Phi \tag{3.12}$$

for which the solution $B = V_0\Phi^{-1}$ is easily obtainable, since Φ is triangular. In terms of the elements β_{ij} of Φ^{-1}, the columns of B may be written

$$\underline{b}_i = \sum_{k=1}^{r} \beta_{ik}\underline{v}_{0k}, \quad i = 1, 2, \ldots, r \tag{3.13}$$

where $\beta_{ij} = 0$ for $j > i$, since Φ is upper triangular.

Substituting (3.13) into (3.8), and rearranging,

$$A\underline{v}_{m-1,k} = \underline{v}_{m,k} + \sum_{j=1}^{r}\left(\sum_{i=1}^{r}\alpha_{i,p_k-m}(k)\beta_{ij}\right)\underline{v}_{0,j}$$

for $m = 1, 2, \ldots, p_k-1$; $k = 1, 2, \ldots, r$, while

$$A\underline{v}_{p_k-1,k} = \sum_{j=1}^{r}\left(\sum_{i=1}^{r} \alpha_{i,0}(k)\beta_{ij}\right)\underline{v}_{0,j}$$

from (3.11).

Define the inner sum

$$\sum_{i=1}^{r} \alpha_{i,p_k-m}(k)\, \beta_{ij} \;\overset{\Delta}{=}\; \gamma_{p_k-m,j}(k) \quad . \tag{3.14}$$

Then

$$A\underline{v}_{m-1,k} = \sum_{j=1}^{r} \gamma_{p_k-m,j}(k)\underline{v}_{0,j} + \underline{v}_{m,k} \tag{3.15}$$

$$m=1,\ 2,\ \ldots,\ p_k-1; \quad k=1,\ 2,\ \ldots,\ r$$

and

$$A\underline{v}_{p_k-1,k} = \sum_{j=1}^{r} \gamma_{0,j}(k)\underline{v}_{0,j} \quad . \tag{3.16}$$

Now define an $n \times n$ matrix T_1 with the columns

$$T_1 = \Big[\underline{v}_{p_1-1,1};\ \underline{v}_{p_1-2,1};\ \ldots;\ \underline{v}_{1,1};\ \underline{v}_{p_2-1,2};\ \ldots; \\ \underline{v}_{p_r-1,r};\ \ldots;\ \underline{v}_{1,r};\ \underline{v}_{0,1};\ \underline{v}_{0,2};\ \ldots;\ \underline{v}_{0,r}\Big]. \tag{3.17a}$$

This matrix will be written in the abbreviated notation

$$T_1 = [V_1;\ V_2;\ \ldots;\ V_r;\ \underline{v}_{0,1};\ \underline{v}_{0,2};\ \ldots;\ \underline{v}_{0,r}] \tag{3.17b}$$

by collecting the groups of vectors into rectangular matrices. If $p_i = 1$ for some i, the matrix V_i does not appear in the matrix T_1.

From (3.17a), AT_1 may be written

$$\Big[A\underline{v}_{p_1-1,1};\ \ldots;\ A\underline{v}_{1,r};\ A\underline{v}_{0,1};\ \ldots;\ A\underline{v}_{0,r}\Big] .$$

Using (3.15) and (3.16), the ℓ^{th} column of AT_1 may be represented as a linear combination of the elements $\underline{v}_{ij}$ in the form

$$\Big[\underline{v}_{p_1-1,1};\ \underline{v}_{p_1-2,1};\ \ldots;\ \underline{v}_{1,r};\ \underline{v}_{0,1};\ \ldots;\ \underline{v}_{0,r}\Big] (\hat{A})_\ell \ .$$

where $(\hat{A})_\ell$ is the ℓ^{th} column of an $n \times n$ matrix $\hat{A}$. Thus,

$$AT_1 = \Big[A\underline{v}_{p_1-1,1};\ \ldots;\ A\underline{v}_{1,r};\ A\underline{v}_{0,1};\ \ldots;\ A\underline{v}_{0,r}\Big] \\ = \Big[\underline{v}_{p_1-1,1};\ \ldots;\ \underline{v}_{1,r};\ \underline{v}_{0,1};\ \ldots;\ \underline{v}_{0,r}\Big]\hat{A} = T_1\hat{A} , \tag{3.18}$$

where the exact form of $\hat{A}$ is described below.

T_1 is a similarity transformation from A to $\hat{A}$,

$$\hat{A} = T_1^{-1}AT_1 . \tag{3.19}$$

It will be shown in Appendix II that T_1 is non-singular by proving that the $\{\underline{v}_{ij}\}$ are linearly independent by construction.

To determine what $\hat{A}$ is, write (3.18) as

$$\Big[AV_1;\ AV_2;\ \ldots;\ AV_r;\ A\underline{v}_{0,1};\ \ldots;\ A\underline{v}_{0,r}\Big] \\ = \Big[V_1;\ V_2;\ \ldots;\ V_r;\ \underline{v}_{0,1};\ \ldots;\ \underline{v}_{0,r}\Big]\left[\begin{array}{c|c} \hat{A}_{11} & \hat{A}_{12} \\ \hline \hat{A}_{21} & \hat{A}_{22} \end{array}\right], \tag{3.20}$$

where the partitioning induced on $\hat{A}$ defines $\hat{A}_{11}$ as an $(n-r) \times (n-r)$ matrix. Let $\hat{A}_{11}$ and $\hat{A}_{21}$ be further partitioned into

$$\hat{A}_{11} = \begin{bmatrix} Ⓗ_{11} & Ⓗ_{12} & \cdots & Ⓗ_{1r} \\ \vdots & & \ddots & \\ Ⓗ_{r1} & \cdot & \cdot & Ⓗ_{rr} \end{bmatrix}$$

and

$$A_{21} = \begin{bmatrix} \underline{\varphi}'_{11} & \cdots & \underline{\varphi}'_{1r} \\ \underline{\varphi}'_{21} & \cdots & \\ \vdots & \ddots & \\ \underline{\varphi}'_{r1} & \cdots & \underline{\varphi}'_{rr} \end{bmatrix}$$

by the partitioning induced by the matrices V_i and vectors $\underline{v}_{0i}$.

Then (3.20) becomes

$$\left[AV_1; \; \ldots; \; AV_r\right] = \left[V_1; \; \ldots; \; V_r\right]\hat{A}_{11} + \left[\underline{v}_{0,1}; \; \ldots; \; \underline{v}_{0,r}\right]\hat{A}_{21}$$

which can be further expanded

$$AV_i = \sum_{j=1}^{r} V_j \Theta_{ji} + \sum_{j=1}^{r} \underline{v}_{0,j}\underline{\varphi}_{ji}', \quad i = 1, 2, \ldots, r. \quad (3.21)$$

From (3.15), (3.16), and the definition of V_i as a collection of vectors, it is seen that

$$\Theta_{ji} = 0, \quad i \neq j,$$

and

$$\underline{\varphi}_{ji}' = \left[\gamma_{0,j}(i), \; \gamma_{1,j}(i), \; \ldots, \; \gamma_{p_i-2,j}(i)\right].$$

If $p_i = 1$ for some i, these terms are deleted from $\hat{A}$. Further,

$$\Theta_{ii} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & & & \ddots & \\ & & & & 1 \\ 0 & \cdots & & & 0 \end{bmatrix} \stackrel{\Delta}{=} Z_i, \quad (3.22)$$

where Z_i is of dimension $(p_i-1) \times (p_i-1)$. If $p_i = 1$, this term is again deleted.

Similarly, define

$$\hat{A}_{12} = \begin{bmatrix} \underline{\xi}_{11} & \cdots & \underline{\xi}_{1r} \\ \underline{\xi}_{21} & & \\ \vdots & \ddots & \\ \underline{\xi}_{r1} & \cdots & \underline{\xi}_{rr} \end{bmatrix}$$

and

$$\hat{A}_{22} = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \cdots & \lambda_{1r} \\ \lambda_{21} & \ddots & & \\ \vdots & & \ddots & \\ \lambda_{r1} & \cdots & & \lambda_{rr} \end{bmatrix}$$

and write

$$\left[A\underline{v}_{0,1}; \; \ldots; \; A\underline{v}_{0,r}\right] = \left[V_1; \; \ldots; \; V_r\right]\hat{A}_{12} + \left[\underline{v}_{0,1}; \; \ldots; \; \underline{v}_{0,r}\right]\hat{A}_{22},$$

or equivalently,

$$A\underline{v}_{0,i} = \sum_{j=1}^{r} V_j \underline{\xi}_{ji} + \sum_{j=1}^{r} \underline{v}_{0,j}\lambda_{ji}. \quad (3.23)$$

Here, it is seen that

$$\underline{\xi}_{ji} = 0, \quad j \neq i,$$

$$\lambda_{ji} = \gamma_{p_i-1,j}(i),$$

and

$$\underline{\xi}_{ii} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \stackrel{\Delta}{=} \underline{e}_i , \tag{3.24}$$

where $\underline{e}_i$ is of length p_i-1 unless $p_i=1$, in which case the term is deleted. This completes the characterization of $\hat{A}$. To summarize,

$$\hat{A} = \left[\begin{array}{cccc|cccc} Z_1 & & & & \underline{e}_1 & & & \\ & Z_2 & & & & \underline{e}_2 & & \\ & & \ddots & & & & \ddots & \\ & & & Z_r & & & & \underline{e}_r \\ \hline \underline{\phi}'_{11} & \underline{\phi}'_{12} & \cdots & \underline{\phi}'_{1r} & \lambda_{11} & \lambda_{12} & \cdots & \lambda_{1r} \\ \vdots & & \ddots & & \vdots & & \ddots & \\ \underline{\phi}'_{r1} & \cdot & \cdot & \underline{\phi}'_{rr} & \lambda_{r1} & \cdot & \cdot & \lambda_{rr} \end{array}\right] \tag{3.25}$$

as asserted in the statement of the theorem.

To obtain the canonical form for B, write

$$\underline{x}(k+1) = A\underline{x}(k) + B\underline{u}(k), \tag{2.1}$$

and define

$$x(k) = T_1\underline{\xi}(k) .$$

Then

$$T_1\underline{\xi}(k+1) = AT_1\underline{\xi}(k) + B\underline{u}(k) ,$$

or

$$\underline{\xi}(k+1) = T_1^{-1}AT_1\underline{\xi}(k) + T_1^{-1}B\underline{u}(k) .$$

Define

$$\hat{B} = T_1^{-1}B . \tag{3.26}$$

Then, from equation (3.19),

$$\underline{\xi}(k+1) = \hat{A}\underline{\xi}(k) + \hat{B}\underline{u}(k) . \tag{3.27}$$

Equation (3.26) implies that

$$T_1\hat{B} = B ,$$

or

$$T_1\hat{\underline{b}}_i = \underline{b}_i = \sum_{k=1}^{r} \beta_{ik}\underline{v}_{0k}, \quad \beta_{ik} = 0 \text{ for } k > i , \tag{3.28}$$

from (3.13). By comparing (3.28) and (3.17), it is seen that

$$\hat{B} = \left[\begin{array}{ccccc} & & 0 & & \\ \hline \beta_{11} & \beta_{21} & \cdots & \cdots & \beta_{r1} \\ 0 & \beta_{22} & & & \cdot \\ \vdots & 0 & \ddots & & \cdot \\ 0 & \cdot & \cdot & 0 & \beta_{rr} \end{array}\right] = \left[\begin{array}{c} 0 \\ \hline \hat{B}_2 \end{array}\right], \tag{3.29}$$

where $\hat{B}_2$ is upper triangular. This completes the proof.

It is interesting to examine the block diagram associated with the general canonical form given by the theorem. This is shown in Figure 1. The sketch shows that the representation is in terms of interconnected subsystems of p_i unit delays driven only at one end by the control inputs and the states of all the subsystems. Although this is a reasonable form, it is not a priori obvious why this should be

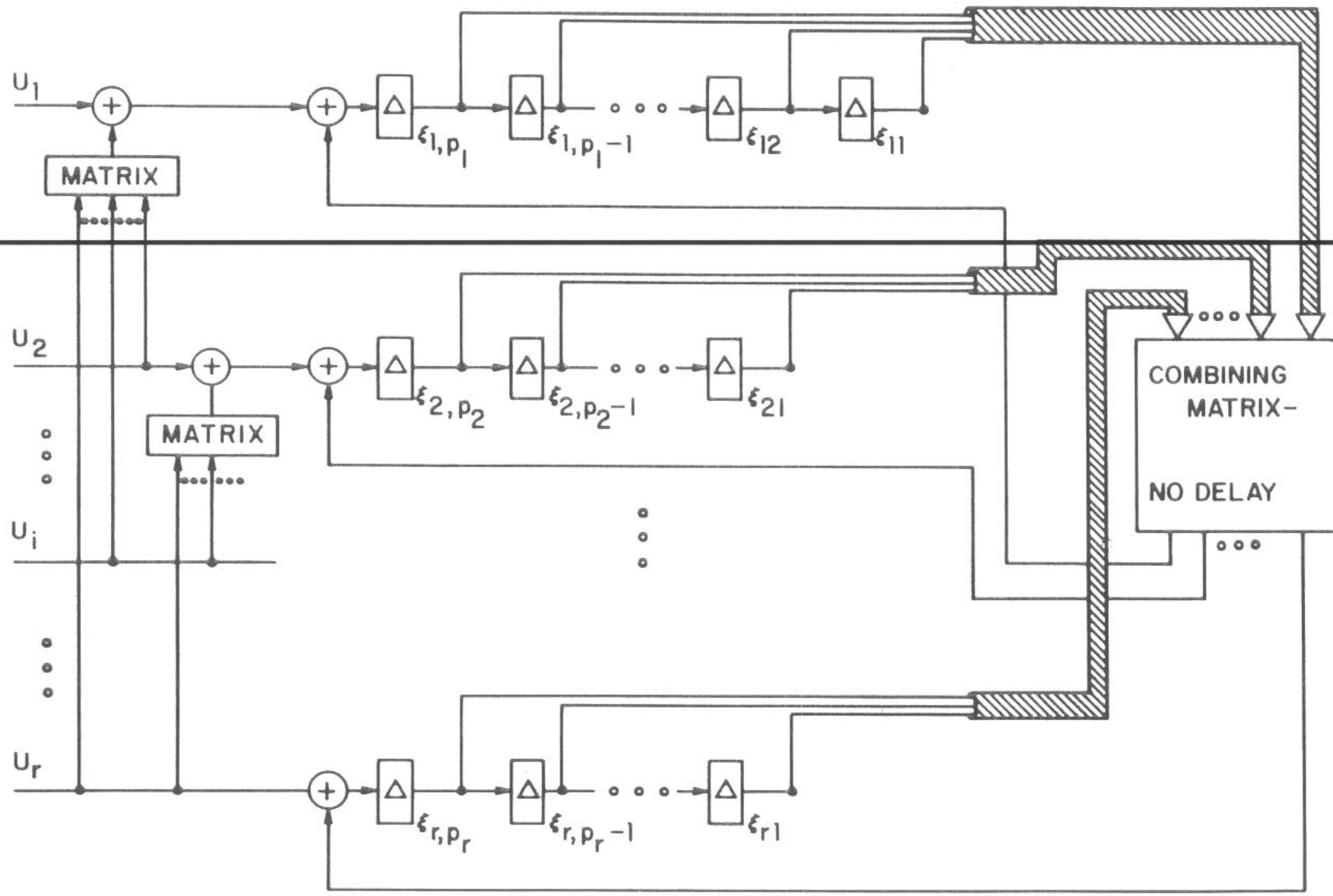

Figure 1 Block Diagram of System in Control Canonical Form.

preferred over some other form, such as Jordan blocks. The rationale is given in Part II of this report.

IV. PROOF OF THEOREM II.

Since Theorem II is the dual of Theorem I, the ideas behind the proof are identical. Hence, only the minor differences in notation will be indicated. For convenience, the equations in this section will be numbered to correspond to the analogous equations in section III.

The observability condition is equivalent to the requirement that

$$\text{rank}\begin{bmatrix} H \\ HA \\ \vdots \\ HA^{n-1} \end{bmatrix} = n, \quad \text{where } H = \begin{bmatrix} \underline{h}'_1 \\ \underline{h}'_2 \\ \vdots \\ \underline{h}'_m \end{bmatrix}.$$

Define the array

$$\begin{matrix} \underline{h}'_1 & \underline{h}'_1 A & \cdot & \cdot & \cdot & \underline{h}'_1 A^n \\ \underline{h}'_2 & \underline{h}'_2 A & \cdot & \cdot & \cdot & \cdot \\ \vdots & \vdots & & & & \cdot \\ \underline{h}'_m & & \cdot & \cdot & \cdot & \underline{h}'_m A^n \end{matrix} \tag{4.1}$$

Obtain the n independent vectors and the indices q_i in the same way as in Theorem I. Then,

$$\underline{h}'_k A^{q_k} = \sum_{j=0}^{q_k} \sum_{i=1}^{m} \delta_{ij}(k)\underline{h}'_i A^j , \tag{4.4}$$

where

$$\delta_{ij}(k) = 0, \quad j > q_k$$
$$\delta_{iq_k}(k) = 0, \quad i \geq k . \tag{4.5}$$

Define

$$\underline{t}'_{\ell,k} = \underline{h}'_k A^{\ell} - \sum_{j=q_k-\ell}^{q_k} \left(\sum_{i=1}^{m} \delta_{ij}(k)\underline{h}'_i \right) A^{j-q_k+\ell} . \tag{4.7}$$

Then

$$\underline{t}'_{\ell,k} = t'_{\ell-1,k}A - \sum_{i=1}^{m} \delta_{i,q_k-\ell}(k)\underline{h}'_i \tag{4.8}$$

for $\ell = 1, 2, \ldots, q_k$; $k = 1, 2, \ldots, m$.

$$\underline{t}'_{0,k} = \underline{h}'_k - \sum_{i=1}^{m} \delta_{i,q_k}(k)\underline{h}'_i , \tag{4.9}$$

and

$$\underline{t}'_{q_k-1,k}A = \sum_{i=1}^{m} \delta_{i0}(k)\underline{h}'_i . \tag{4.10}$$

Equation (4.9) defines a transformation

$$\begin{bmatrix} \underline{t}'_{0,1} \\ \underline{t}'_{0,2} \\ \cdot \\ \cdot \\ \cdot \\ \underline{t}'_{0,m} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdot & \cdot & \cdot & 0 \\ -\delta_{1q_2}(2) & 1 & 0 & \ldots & & 0 \\ -\delta_{1q_3}(3) & -\delta_{2q_3}(3) & & \ddots & & \\ \vdots & & & & & 0 \\ -\delta_{1q_m}(m) & \cdot & \cdot & \cdot & & 1 \end{bmatrix} \begin{bmatrix} \underline{h}'_1 \\ \underline{h}'_2 \\ \cdot \\ \cdot \\ \cdot \\ \underline{h}'_m \end{bmatrix} \tag{4.11}$$

from which

$$\underline{h}'_i = \sum_{k=1}^{m} \mu_{ik}\underline{t}'_{0,k}, \quad i = 1, 2, \ldots, m \tag{4.13}$$

where $\mu_{ik} = 0$, $i < k$. Substituting back into (4.8) and (4.9), define

$$\nu_{q_k-\ell,j}^{(k)} = \sum_{i=1}^{m} \delta_{i,q_k-\ell}^{(k)}\mu_{ij}, \quad \ell = 1, 2, \ldots, q_k \quad j = 1, 2, \ldots, m \quad k = 1, 2, \ldots, m. \tag{4.14}$$

Then the $n \times n$ matrix

$$T_2 = \begin{bmatrix} \underline{t}'_{0,1} \\ \underline{t}'_{0,2} \\ \vdots \\ \underline{t}'_{0,m} \\ \underline{t}'_{1,1} \\ \underline{t}'_{2,1} \\ \vdots \\ \underline{t}'_{q_1-1,1} \\ \underline{t}'_{1,2} \\ \vdots \\ \underline{t}'_{q_m-1,m} \end{bmatrix} \tag{4.17}$$

is the required transformation, defining

$$\widetilde{A} = T_2 A T_2^{-1}. \tag{4.19}$$

The sub-partitions become

$$\widetilde{A}_{22} = \begin{bmatrix} Z_1 & & & \\ & Z_2 & & \\ & & \ddots & \\ & & & Z_m \end{bmatrix}, \tag{4.25a}$$

$$\widetilde{A}_{12} = \begin{bmatrix} \underline{g}'_1 & & & \\ & \underline{g}'_2 & & \\ & & \ddots & \\ & & & \underline{g}'_m \end{bmatrix}, \tag{4.25b}$$

where $\underline{g}'_i = [1, 0, 0, \ldots, 0]$ is of dimension q_i-1, unless $q_i = 1$, in which case the entry is deleted. The rest of the elements of $\widetilde{A}$ are

$$\widetilde{A}_{11} = \begin{bmatrix} \nu_{q_1-1,1}^{(1)} & \cdots & \nu_{q_1-1,m}^{(1)} \\ \nu_{q_2-1,1}^{(2)} & \cdots & \cdot \\ \vdots & & \vdots \\ \nu_{q_m-1,1}^{(m)} & & \nu_{q_m-1,m}^{(m)} \end{bmatrix} \tag{4.25c}$$

and

$$\widetilde{A}_{21} = \begin{bmatrix} \nu_{q_1-2,1}^{(1)} & \cdots & \nu_{q_1-2,m}^{(1)} \\ \nu_{q_1-3,1}^{(1)} & \cdots & \nu_{q_1-3,m}^{(1)} \\ \vdots & & \vdots \\ \nu_{0,1}^{(1)} & \cdots & \nu_{0,m}^{(1)} \\ \nu_{q_2-2,1}^{(2)} & \cdots & \nu_{q_2-2,m}^{(2)} \\ \vdots & & \vdots \\ \nu_{0,1}^{(2)} & \cdots & \nu_{0,m}^{(2)} \\ \vdots & & \vdots \\ \nu_{0,1}^{(m)} & \cdots & \nu_{0,m}^{(m)} \end{bmatrix}. \tag{4.25d}$$

If some $q_i = 1$, the entry

$$\nu_{q_i-1,1}^{(i)} = \nu_{0,1}^{(i)}$$

belongs in $\tilde{A}_{11}$ and is deleted from $\tilde{A}_{21}$.

The matrix $\tilde{H}$, for which

$$\tilde{H}T_2 = H, \qquad (4.28)$$

is $[\tilde{H}_1 \mid \tilde{H}_2] = \tilde{H}$. The sub-matrices are $\tilde{H}_2 = 0$,

$$\tilde{H}_1 = \begin{bmatrix} 1 & 0 & . & . & . & 0 \\ \mu_{21} & 1 & & & & \\ \mu_{31} & \mu_{32} & \cdot & & & \\ & & & . & & \\ \vdots & \vdots & & & \cdot & 0 \\ \mu_{m1} & \mu_{m2} & . & . & . & 1 \end{bmatrix} . \qquad (4.29)$$

This completes the proof of Theorem II.

V. SUMMARY

Canonical forms for multivariable systems have been derived. These show that any controllable (observable) system may be expressed as an interconnected set of sub-systems in a manner which separates the structure and parameters of the system. This structure differs from previously used canonical matrix forms (e.g., the Jordan form) in that at no time is one worried about obtaining the matrix eigenvalues and determining their multiplicity.

It is worthwhile to point out that since the controllability condition for the differential system $\dot{\underline{x}} = A\underline{x} + B\underline{u}$ is the same as for the discrete system considered here, the canonical form for a differential system is identical to the one derived in this paper.

The proof of the theorems is constructive and can be used as an algorithm to actually obtain the proper canonical form from any original system description. For single input or single output systems, the standard form obtained by this procedure becomes the "phase-variable" canonical form, which has been the subject of recent interest.[3-7] For this special case, the procedure simplifies considerably and has been reported previously.[8,9]

APPENDIX I: EXAMPLES

Example 1

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0.1 & 1 & 0 \\ .5 & 0 & 0 & 1 \\ 0 & 0 & .2 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad H = [1 \ 0 \ 0 \ 0] .$$

$n = 4; \ r = 2; \ m = 1$.

Let us obtain first the control canonic form. The array (3.1) becomes

$$\underline{b}_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad A\underline{b}_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad A^2\underline{b}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ .2 \end{bmatrix} \quad A^3\underline{b}_1$$

$$\underline{b}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad A\underline{b}_2 = \begin{bmatrix} 1 \\ .1 \\ 0 \\ 0 \end{bmatrix} \quad A^2\underline{b}_2 = \begin{bmatrix} .1 \\ .01 \\ .5 \\ 0 \end{bmatrix} \quad A^3\underline{b}_2 \ .$$

The independent vectors are $\underline{b}_1$; $\underline{b}_2$; $A\underline{b}_1$; $A\underline{b}_2$. Hence, $p_1 = 2$, $p_2 = 2$, and

$$A^{p_1}\underline{b}_1 = A^2\underline{b}_1 = .2\underline{b}_1 + \underline{b}_2$$

$$A^{p_2}\underline{b}_2 = A^2\underline{b}_2 = .5A\underline{b}_1 + .1A\underline{b}_2$$

analogous to (3.2). The only non-zero α's are

$$\alpha_{10}(1) = 0.2 \qquad \alpha_{11}(2) = 0.5$$

$$\alpha_{20}(1) = 1.0 \qquad \alpha_{21}(2) = 0.1 \ .$$

From (3.7), the vectors $\underline{v}_{ij}$ become

$$\underline{v}_{0,1} = \underline{b}_1 - \sum_{i=1}^{2} \alpha_{i,2}(1)\underline{b}_i = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \underline{b}_1$$

$$\underline{v}_{1,1} = A\underline{b}_1 - \sum_{j=1}^{2}\sum_{i=1}^{2} \alpha_{ij}(1)A^{j-1}\underline{b}_i = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

$$\underline{v}_{0,2} = \underline{b}_2 - \sum_{i=1}^{2} \alpha_{i,2}(2)\underline{b}_i = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \underline{b}_2$$

$$\underline{v}_{1,2} = A\underline{b}_2 - \sum_{j=1}^{2}\sum_{i=1}^{2} \alpha_{ij}(2)A^{j-1}\underline{b}_i = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -.5 \end{bmatrix} .$$

Since $\underline{b}_i = \underline{v}_{0,i}$,

$$\beta_{11} = 1; \ \beta_{21} = 0; \ \beta_{22} = 1,$$

from (3.13). Thus, from (3.14)

$$\gamma_{p_k-m,j}(k) = \alpha_{j,p_k-m}(k), \quad k = 1, 2; \ m = 0, 1; \ j = 1, 2.$$

From (3.17), the transforming matrix T_1 becomes

$$T_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -.5 & 1 & 0 \end{bmatrix} .$$

In $\hat{A}_{11}$, $Z_1 = 0$ and $Z_2 = 0$, while in $\hat{A}_{12}$, $\underline{e}_1 = 1$ and $\underline{e}_2 = 1$. Thus, the canonical matrix $\hat{A}$ becomes from (3.25)

$$\hat{A} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \gamma_{0,1}(1) & \gamma_{0,1}(2) & \gamma_{1,1}(1) & \gamma_{1,1}(2) \\ \gamma_{0,2}(1) & \gamma_{0,2}(2) & \gamma_{1,2}(1) & \gamma_{1,2}(2) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ .2 & 0 & 0 & .5 \\ 1 & 0 & 0 & .1 \end{bmatrix}$$

while from (3.28),

$$\hat{B} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Now to obtain the observable canonic form, consider the vectors

$$\underline{h}'_1 = [1, 0, 0, 0]$$

$$\underline{h}'_1 A = [0, 1, 0, 0]$$

$$\underline{h}'_1 A^2 = [0, .1, 1, 0]$$

$$\underline{h}'_1 A^3 = [.5, .01, .1, 1]$$

$$\underline{h}'_1 A^4 = [.05, .501, .21, .1] .$$

The first four are seen to be independent, while the last may be expressed as

$$\underline{h}'_1 A^4 = .48\underline{h}'_1 A + .2\underline{h}'_1 A^2 + .1\underline{h}'_1 A^3 ,$$

which defines

$$\delta_{10}(1) = 0;\ \delta_{11}(1) = .48;\ \delta_{12}(1) = .2;\ \delta_{13}(1) = .1$$

from (4.5). It is seen also that $q_1 = 4$. From (4.7),

$$\underline{t}'_{0,1} = \underline{h}'_1 = [1\ 0\ 0\ 0]$$

$$\underline{t}'_{1,1} = \underline{h}'_1 A - \sum_{j=3}^{3} \delta_{1j}(1)\underline{h}'_1 A^{j-3} = [-.1\ 1\ 0\ 0]$$

$$\underline{t}'_{2,1} = \underline{h}'_1 A^2 - \sum_{j=2}^{3} \delta_{1j}(1)\underline{h}'_1 A^{j-2} = [-.2\ 0\ 1\ 0]$$

$$\underline{t}'_{3,1} = \underline{h}'_1 A^3 - \sum_{j=1}^{3} \delta_{1j}(1)\underline{h}'_1 A^{j-1} = [.02\ -.2\ 0\ 1] .$$

Then, using (4.17),

$$T_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -.1 & 1 & 0 & 0 \\ -.2 & 0 & 1 & 0 \\ .02 & -.2 & 0 & 1 \end{bmatrix}.$$

From (4.13) and (4.14), $\mu_{11} = 1$. Thus,

$$\nu_{4-\ell,1}(1) = \delta_{1,4-\ell}(1) .$$

The subpartitions of $\tilde{A}$ become

$$\hat{A}_{11} = \nu_{3,1}(1) = \delta_{1,3}(1) = .1$$

$$\hat{A}_{21} = \begin{bmatrix} \nu_{2,1}(1) \\ \nu_{1,1}(1) \\ \nu_{0,1}(1) \end{bmatrix} = \begin{bmatrix} .2 \\ .48 \\ 0 \end{bmatrix}$$

$$\tilde{A}_{22} = Z_1 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix},\quad A_{12} = [1\ 0\ 0] .$$

Thus,

$$\tilde{A} = \left[\begin{array}{c|ccc} .1 & 1 & 0 & 0 \\ \hline .2 & 0 & 1 & 0 \\ .48 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right],$$

from (4.25), and $\tilde{H} = [1 \ \ 0 \ \ 0 \ \ 0]$ from (4.29).

Example 2

$$A = \begin{bmatrix} .7 & 1 & 0 & 0 & 0 & 0 \\ 0 & .1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & .5 & 0 & 0 \\ .2 & 0 & 0 & 0 & -.1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 2. & 0 \end{bmatrix},$$

$n = 6; \ r = 3.$

The vector array is shown in Figure 2. From Figure 2, $p_1 = 1$; $p_2 = 3$; $p_3 = 2$. Equation (3.2) gives the relations

$$A\underline{b}_1 = .5\underline{b}_1$$

$$A^3\underline{b}_2 = .08A\underline{b}_2 + A^2\underline{b}_2 + .014\underline{b}_3 - .31A\underline{b}_3$$

$$A^2\underline{b}_3 = .1A\underline{b}_2 + .7A\underline{b}_3 - .07\underline{b}_3,$$

defining

$\alpha_{10}(1) = .5$

$\alpha_{2,1}(2) = .08$

$\alpha_{2,2}(2) = 1.$

$\alpha_{3,0}(2) = .014$

$\alpha_{3,1}(2) = -.31$

$\alpha_{2,1}(3) = .1$

$\alpha_{3,0}(3) = -.07$

$\alpha_{3,1}(3) = .7$.

$$\underline{b}_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad A\underline{b}_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ .5 \\ 0 \\ 0 \end{bmatrix} \quad A^2\underline{b}_1 = \begin{bmatrix} \\ \\ \\ \\ \\ \\ \end{bmatrix} \quad A^3\underline{b}_1 = \begin{bmatrix} \\ \\ \\ \\ \\ \\ \end{bmatrix} \quad \cdots$$

$$\underline{b}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 2 \end{bmatrix} \quad A\underline{b}_2 = \begin{bmatrix} 1 \\ .1 \\ 0 \\ 0 \\ 2. \\ 0 \end{bmatrix} \quad A^2\underline{b}_2 = \begin{bmatrix} .8 \\ .01 \\ 2. \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad A^3\underline{b}_2 = \begin{bmatrix} .57 \\ .001 \\ 2. \\ 0 \\ .16 \\ 0 \end{bmatrix} \quad \cdots$$

$$\underline{b}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad A\underline{b}_3 = \begin{bmatrix} 1 \\ .1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad A^2\underline{b}_3 = \begin{bmatrix} .8 \\ .01 \\ 0 \\ 0 \\ .2 \\ 0 \end{bmatrix} \quad A^3\underline{b}_3 = \begin{bmatrix} \\ \\ \\ \\ \\ \\ \end{bmatrix} \quad \cdots$$

Figure 2 Vector Array for Example 2.

All other $\alpha_{ij}(k)$ are zero. From (3.7), one calculates the vectors

$$\underline{v}_{0,1} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \underline{b}_1; \qquad \underline{v}_{0,2} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 2. \end{bmatrix} = \underline{b}_2;$$

$$\underline{v}_{0,3} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \underline{b}_3; \qquad \underline{v}_{1,2} = \begin{bmatrix} 1. \\ -.9 \\ 0 \\ 0 \\ 2. \\ -2. \end{bmatrix};$$

$$\underline{v}_{2,2} = \begin{bmatrix} -.2 \\ .14 \\ 2. \\ 0 \\ -2. \\ -.16 \end{bmatrix}; \qquad \underline{v}_{1,3} = \begin{bmatrix} 1. \\ -.7 \\ 0 \\ 0 \\ 0 \\ -.2 \end{bmatrix}$$

giving the transformation

$$T_1 = [\underline{v}_{2,2};\ \underline{v}_{1,2};\ \underline{v}_{1,3};\ \underline{v}_{0,1};\ \underline{v}_{0,2};\ \underline{v}_{0,3}]$$

$$= \begin{bmatrix} -.2 & 1. & 1. & 0 & 0 & 0 \\ .14 & -.9 & -.7 & 0 & 1 & 1 \\ 2. & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ -2. & 2. & 0 & 0 & 0 & 0 \\ -.16 & -2. & -.2 & 0 & 2. & 0 \end{bmatrix}.$$

Note that the matrix V_1 is deleted, since $p_1 = 1$.

From (3.13) and (3.14),

$$\gamma_{p_k-m,j}(k) = \alpha_{j,p_k-m}(k); \quad m = 1, 2, \ldots, p_k; \quad k = 1, 2, 3; \quad j = 1, 2, 3.$$

The submatrices become

$$\hat{A}_{11} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad \hat{A}_{12} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$\hat{A}_{21} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & .08 & 0 \\ .014 & -.31 & -.07 \end{bmatrix}, \qquad \hat{A}_{22} = \begin{bmatrix} .5 & 0 & 0 \\ 0 & 1. & .1 \\ 0 & 0 & .7 \end{bmatrix}.$$

Note that the terms Z_1, $\underline{e}_1$, and $\underline{e}'_{11}$ are deleted, since $p_1 = 1$.

Finally,

$$\hat{A} = \left[\begin{array}{ccc|ccc} 0 & 1. & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & .5 & 0 & 0 \\ 0 & .08 & 0 & 0 & 1. & .1 \\ .014 & -.31 & -.07 & 0 & 0 & .7 \end{array}\right],$$

and

$$\hat{B} = \left[\begin{array}{ccc} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \hline 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right].$$

APPENDIX II. Proof that the Transformation T_1 Is Non-Singular

Recall that T_1 is defined by the vectors

$$\{\underline{v}_{ij}\}, \quad i = 0, 1, \ldots, p_{j-1}; \quad j = 1, 2, \ldots, r.$$

It is known that T_1 is non-singular if, and only if, the vectors $\{\underline{v}_{ij}\}$ are linearly independent. Further, recall that the sequence $\{\underline{v}_{m,k}\}$ is generated by

$$\underline{v}_{m,k} = A^m \underline{b}_k - \sum_{j=p_k-m}^{p_k} \sum_{i=1}^{r} \alpha_{ij}(k) A^{j-p_k+m} \underline{b}_i \qquad (3.7)$$

where the second term involves the independent vectors in the set

$$\underline{b}_1;\ \underline{b}_2;\ \ldots;\ \underline{b}_r;\ A\underline{b}_1;\ \ldots;\ A\underline{b}_r;\ \ldots;\ A^m\underline{b}_1;\ \ldots;\ A^m\underline{b}_{k-1}$$

by the definition of $\alpha_{ij}(k)$ in (3.4). Since, by construction, $A^m\underline{b}_k$ is independent, it is clear that each $\underline{v}_{m,k}$ is independent of the $\underline{v}_{ij}$'s previously generated. This guarantees that the $\underline{v}_{m,k}$'s are all independent.

Q. E. D.

REFERENCES

1. K. J. Åström, R. W. Koepcke, F. Tung: "On the Control of Linear Discrete Dynamic Systems with Quadratic Loss," IBM Research Report RJ-222, September 10, 1962.

2. R. E. Kalman, Y. C. Ho, K. S. Narenda: "Controllability of Linear Dynamical Systems," in Contributions to Differential Equations, Vol. 1, No. 2, pp. 189-213, Interscience, New York, 1963.

3. C. D. Johnson and W. M. Wonham: "A Note on the Transformation to Canonical (Phase-Variable) Form," IEEE Trans. on Auto. Control (Short Papers), Vol. AC-9, pp. 312-313, July 1964.

4. I. H. Mufti: "On the Reduction of a System to Canonical (Phase-Variable) Form," IEEE Trans. on Auto. Control (correspondence), Vol. AC-10, pp. 206-207, April 1965.

5. M. R. Chidambara: "The Transformation to (Phase-Variable) Canonical Form," IEEE Trans. on Auto. Control (correspondence), Vol. AC-10, pp. 492-495, October 1965.

6. L. M. Silverman: "Transformation of Time-Variable Systems to Canonical (Phase-Variable) Form," IEEE Trans. on Auto. Control, Vol. AC-11, April 1966.

7. R. W. Bass and I. Gura: "Linear Canonical Forms for Controllable Systems," Hughes Aircraft Company, February 1965.

8. W. G. Tuel, Jr.: "On the Transformation to (Phase-Variable) Canonical Form," IEEE Trans. on Auto. Control (correspondence), Vol. AC-11, July 1966.

9. J. Rissanen: "Control System Synthesis by Analog Computer Based on the Generalized Feedback Concept," Proceedings of "Analog Computation Applied to the Study of Chemical Processes," International Seminar, Brussels, November 1960.