KYBERNETIKA — VOLUME 17 (1981), NUMBER 5

INVARIANTS AND CANONICAL FORMS FOR LINEAR MULTIVARIABLE SYSTEMS UNDER THE ACTION OF ORTHOGONAL TRANSFORMATION GROUPS

M. M. KONSTANTINOV, P. HR. PETKOV, N. D. CHRISTOV

Subject of the present paper is the study and construction of complete independent invariants and canonical forms for linear multivariable systems under the action of orthogonal transformation groups. Stable computational algorithms for finding the orthogonal canonical forms are presented and their numerical properties are discussed.

In view of their nice numerical properties the orthogonal canonical forms are preferable for computations. They reveal the basic invariant structure of linear multivariable systems and provide the same theoretical advantages as the canonical forms relative to general transformation groups.

1. INTRODUCTION

The problem of finding invariants and canonical forms for linear multivariable systems (LS) is of current interest for the analysis and synthesis as well as for the realization theory and identification of these systems: [1]–[8], etc. However the determination of canonical forms relative to general transformation groups (e.g. general linear or feedback groups) is quite unsatisfactory from computational point of view (see [9] for more details). That is why the determination of canonical forms relative to smaller groups is of great importance, the more so, as such canonical forms provide the same advantages in the synthesis of LS. Unfortunately there are no investigations up to this moment devoted to the rigorous study of this subject.

In [9]–[11] the so-called serial canonical form (SCF) of LS is introduced and applied to the solution of the general problem of synthesis of LS with complete and incomplete state feedback. In [9] the SCF is defined relative to the orthogonal group $\mathfrak{O}(n)$; however for multi-input (multi-output) systems this variant of SCF is in fact quasicanonical.

In this paper a complete independent invariant of LS relative to $\mathfrak{O}(n)$ and $\mathfrak{O}^+(n)$ ($\mathfrak{O}^+(n) \subset \mathfrak{O}(n)$ denotes the group of pure rotations) is defined and stable computational algorithms for finding the corresponding canonical forms are presented. The orthogonal canonical forms reveal the basic invariant structure of LS (e.g. the Kronecker indices) and have the same theoretical advantages as the canonical forms relative to general transformation groups, being as well favorable for computations.

Further on the following abbreviations are used: $\mathscr{R}_m^n$ – the space of real $n \times m$ matrices ($\mathscr{R}_1^n = \mathscr{R}^n$, $\mathscr{R}_m^1 = \mathscr{R}_m$, $\mathscr{R}_1^1 = \mathscr{R}$); $\boldsymbol{I}_n \in \mathscr{R}_n^n$ – the unit matrix; $\boldsymbol{A}^T \in \mathscr{R}_n^m$ – the transpose of $\boldsymbol{A} \in \mathscr{R}_m^n$; $\mathfrak{GL}(n)$, $\mathfrak{O}(n)$ and $\mathfrak{O}^+(n)$ – the groups of matrices $\boldsymbol{P} \in \mathscr{R}_n^n$ with $\det \boldsymbol{P} \neq 0$, $\boldsymbol{P}^T\boldsymbol{P} = \boldsymbol{I}_n$ and $\boldsymbol{P}^T\boldsymbol{P} = \boldsymbol{I}_n$, $\det \boldsymbol{P} = 1$ resp.

2. ABSTRACT INVARIANTS AND CANONICAL FORMS

Let X be a set and $\mathfrak{G}$ a group of automorphisms $\mathfrak{g}, \mathfrak{h}, \ldots,$ on X. The transformation group $\mathfrak{G}$ defines an equivalence relation E on $X : xEy \Leftrightarrow y = \mathfrak{g}(x)$ for some $\mathfrak{g} \in \mathfrak{G}$. The set $E_x = \{\mathfrak{g}(x) : \mathfrak{g} \in \mathfrak{G}\}$ is called an orbit of $x \in X$. The set X is the union of all disjoint orbits: $X = \bigcup E_x$.

If J is any set then a function $f : X \to J$ is said to be an invariant (for the equivalence relation E, or relative to the group $\mathfrak{G}$) if xEy implies $f(x) = f(y)$. If in addition $f(x) = f(y) \Rightarrow xEy$, the function f is said to be a complete invariant. In what follows we shall deal only with surjective invariants, replacing (if necessary) J by $f(X)$.

A subset $\tilde{X} \subset X$ is called a set of canonical forms, or a canonical set, if for any $x \in X$ there exists exactly one $\tilde{x} \in \tilde{X}$ such that $xE\tilde{x}$, i.e. if $\tilde{X} \cap E_x = \{\tilde{x}\}$. Denote by C_X the set of all canonical sets in X.

The subgroup $\mathfrak{G}_x = \{\mathfrak{g} \in \mathfrak{G} : \mathfrak{g}(x) = x\}$ is said to be the stabilizator for $x \in X$. The element x is unstable if $\mathfrak{G}_x = \{\mathfrak{e}\}$, where $\mathfrak{e} \in \mathfrak{G}$ is the identity. Denote by $U_X = \{x \in X : \mathfrak{G}_x = \{\mathfrak{e}\}\}$ the subset of unstable elements in X. It can be shown that if xEy then $x \in U_X \Leftrightarrow y \in U_X$. Indeed, let $x \in U_X$ and $y = \mathfrak{h}(x)$. Then $\mathfrak{g}(y) = y$ leads to $\mathfrak{g} \circ \mathfrak{h}(x) = \mathfrak{h}(x)$, $\mathfrak{h}^{-1} \circ \mathfrak{g} \circ \mathfrak{h}(x) = x$ and $\mathfrak{h}^{-1} \circ \mathfrak{g} \circ \mathfrak{h} = \mathfrak{e}$, i.e. $\mathfrak{g} = \mathfrak{e}$ and $\mathfrak{G}_y = \{\mathfrak{e}\}$. A simple corollary is that $\tilde{X} \in C_X$ and $\tilde{X} \subset U_X$ imply $X = U_X$.

The following propositions are essentially used later:

Proposition 1. Let $\tilde{X} \in C_X$. Then the following two statements are equivalent:

(i) $X = U_X$;

(ii) for each $x \in X$ there exists an unique $\mathfrak{s} \in \mathfrak{G}$ such that $\mathfrak{s}(x) \in \tilde{X}$.

Proof. (i) $\Rightarrow$ (ii). Let $\mathfrak{g}(x) = \mathfrak{h}(x) = \tilde{x} \in \tilde{X}$. Then $\mathfrak{h} \circ \mathfrak{g}^{-1}(\tilde{x}) = \tilde{x}$ and $\mathfrak{h} \circ \mathfrak{g}^{-1} = \mathfrak{e}$, i.e. $\mathfrak{g} = \mathfrak{h} = \mathfrak{s}$.

(ii) $\Rightarrow$ (i). If $\mathfrak{g}(x) = x$ and $\mathfrak{h}(x) = \tilde{x} \in \tilde{X}$ then $\mathfrak{h} \circ \mathfrak{g}(x) = \tilde{x}$. Now (ii) yields $\mathfrak{h} = \mathfrak{s}$, $\mathfrak{g} \circ \mathfrak{h} = \mathfrak{s}$ and $\mathfrak{g} = \mathfrak{e}$. Hence $\mathfrak{G}_x = \{\mathfrak{e}\}$ for any $x \in X$. □

Proposition 2. Let $\tilde{X} \subset X$ and $X = U_X$. Then $\tilde{X} \in C_X$ iff

(j) $\tilde{X} \cap E_x \neq \emptyset$ for each $x \in X$;

(jj) for each $y \in \tilde{X}$ and $\mathfrak{g} \in \mathfrak{G}$ the inclusion $\mathfrak{g}(y) \in \tilde{X}$ implies $\mathfrak{g} = \mathfrak{e}$.

Proof. "Only if" statement. Let $\tilde{X} \in C_X$. Then (j) is valid by definition. Let now $\mathfrak{g} \neq \mathfrak{e}$ and $y, \mathfrak{g}(y) \in \tilde{X}$. Since $\mathfrak{G}_y = \{\mathfrak{e}\}$ then $\mathfrak{g}(y) \neq y$ and hence $\tilde{X}$ contains at least two different elements from one orbit. The contradition shows that $\mathfrak{g} = \mathfrak{e}$ and (jj) holds true.

"If" statement. According to (j) $\tilde{X}$ contains at least one element from every orbit and this element is unique in view of (jj). $\square$

In what follows we assume that the elements x of X are sets of real matrices. Then X is isomorphic to certain subset of a finite-dimensional space over $\mathscr{R}$.

3. ORTHOGONAL INVARIANTS AND CANONICAL FORMS FOR LINEAR MULTIVARIABLE SYSTEMS

Consider the completely controllable LS

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\mathbf{x}(t) = \mathbf{A}\,\mathbf{x}(t) + \mathbf{B}\,\mathbf{u}(t)\,, \tag{1}$$

where $\mathbf{x}(t) \in \mathscr{R}^n$, $\mathbf{u}(t) \in \mathscr{R}^m$ and $\mathbf{A} \in \mathscr{R}^n_n$, $\mathbf{B} \in \mathscr{R}^n_m$. The system (1) will be identified with the matrix pair $\mathbf{S} = [\mathbf{A}, \mathbf{B}) \in \mathscr{L}_m(n)$. Here $\mathscr{L}_m(n) \subset \mathscr{R}^n_n \times \mathscr{R}^n_m$ is the set of all pairs $\mathbf{S}$ with rank $\mathbf{C}_n(\mathbf{S}) = n$,

$$\mathbf{C}_n(\mathbf{S}) = [\mathbf{B} \mid \mathbf{AB} \mid \dots \mid \mathbf{A}^{n-1}\mathbf{B}] \in \mathscr{R}^n_{nm}\,.$$

Let $\mathfrak{G}$ be a transformation group with exact matrix representation $\mathfrak{G} \to \mathfrak{F} \subset \subset \mathfrak{GL}(n)$ whose action on $\mathscr{L}_m(n)$ is given by

$$\mathfrak{g}[\mathbf{A}, \mathbf{B}) = [\mathbf{P}^{-1}\mathbf{A}\mathbf{P}, \mathbf{P}^{-1}\mathbf{B})\,; \quad \mathfrak{g} \in \mathfrak{G}\,, \quad \mathbf{P} = \mathbf{P}(\mathfrak{g}) \in \mathfrak{F}\,,$$

and corresponds to the linear transformation $\mathbf{x}(t) = \mathbf{P}\,\mathbf{x}'(t)$ in the state space $\mathscr{R}^n$. Further on we shall identify $\mathfrak{G}$ with the matrix group $\mathfrak{F}$.

An invariant f for LS relative to $\mathfrak{F}$ can be determined as the product of an integer valued and a real valued functions in the following way. Let $\mathbf{Z}(n)$ be the set of ordered additive partitionings $z = (n_1, \dots, n_m)$ of $n : n_1 + \dots + n_m = n$, $n_i \geqq 0$. Then the invariant f can be chosen in the form $f = (a, r_a)$, where the first projection $a : \mathscr{L}_m(n) \to \to \mathbf{Z}(n)$ is surjective and is said to be arithmetic invariant [8]. On each subset $a^{-1}(z) \subset \mathscr{L}_m(n)$ we define an integer $N = N(z)$ and a rational invariant function $r_a : a^{-1}(z) \to \mathbf{R}^{\sim}$, where $\mathbf{R}^{\sim}$ is isomorphic to $\mathscr{R}^N$. Then the image $\mathbf{J}$ of f is the set of pairs $(z, \mathbf{v})$ with $z \in \mathbf{Z}(n)$, $\mathbf{v} \in \mathbf{R}^{\sim}$, and

$$f : \mathbf{S} \mapsto (a(\mathbf{S}), r_{a(\mathbf{S})}(\mathbf{S}))\,.$$

If the invariant f is complete and $\mathbf{R}^{\sim} \subset \mathscr{L}_m(n)$ then $\mathbf{R}^{\sim}$ is a canonical set for $a^{-1}(z)$, where z is a fixed member of $\mathbf{Z}(n)$. Moreover, in this case the invariant f is independent in the sense of the definitions from [4], [5], [7], [8].

In the construction of invariants and canonical forms an alternative approach is possible: For z fixed first find $N = N(z)$ and construct an appropriate set $\mathscr{L}_m^{\sim}(n) \subset \subset \mathscr{L}_m(n)$ (isomorphic to $\mathscr{R}^N$) which satisfies the conditions of Proposition 2, and clarify that every $\mathbf{S} \in \mathscr{L}_m(n)$ is unstable. Then $\mathscr{L}_m^{\sim}(n)$ is a canonical set and the non-fixed entries of $\mathbf{S}^{\sim} \in \mathscr{L}_m^{\sim}(n)$ constitute an independent invariant relative to $\mathfrak{F}$. Adding the map a which produces the Kronecker indices of $\mathbf{S}$ one obtains a complete independent invariant.

The problems of construction of canonical forms relative to $\mathfrak{GL}(n)$ (i.e. in the case $\mathfrak{F} = \mathfrak{GL}(n)$) have been studied by many authors: e.g. in [1]–[4] for the Brunovsky-Luenberger canonical form, and in [10]–[12] for the SCF.

However the methods for obtaining canonical forms relative to the general linear group $\mathfrak{GL}(n)$ are numerically unstable since they correspond to transformation of a matrix in its Frobenius form [9]. Most of these methods require the determination of the controllability matrix $\mathbf{C}_n(\mathbf{S})$. This matrix however is ill-conditioned since its columns tend to become linearly dependent for large n [16].

The algorithms proposed in [17] are free from this disadvantage since they are based on elementary transformations of $\mathbf{A}, \mathbf{B}$. Unfortunately these algorithms are also numerically unstable as a result of setting unit elements in prescribed positions in the canonical form.

The above numerical difficulties are common for all canonical forms relative to general transformation groups. Thus it is necessary to develop a theory and computational methods for obtaining canonical forms of LS relative to smaller groups which are more favorable from computational point of view. Typical representatives here are the orthogonal groups $\mathfrak{O}(n)$ and $\mathfrak{O}^+(n)$ which act on $\mathscr{L}_m(n)$ according

$$\mathfrak{g}[\mathbf{A}, \mathbf{B}) = [\mathbf{P}^T\mathbf{AP}, \mathbf{P}^T\mathbf{B}); \; \mathbf{P} = \mathbf{P}(\mathfrak{g}) \in \mathfrak{O}(n) \; (\mathbf{P} = \mathbf{P}(\mathfrak{g}) \in \mathfrak{O}^+(n)) .$$

In this section complete independent orthogonal invariants and canonical sets for LS are defined and numerically stable algorithms for transformation into the orthogonal canonical forms are derived.

Let $z = (n_1, \ldots, n_m)$ be the set of ordered Kronecker indices for the pair $\mathbf{S} = [\mathbf{A}, \mathbf{B})$. Denote by $(m_0, m_1, \ldots, m_p)$ the set of unordered conjugate indices $m = m_0 \geqq m_1 \geqq \ldots \geqq m_p \geqq 1$, $m_1 + \ldots + m_p = n$, where m_i is the number of n_j's that are $\geqq i$:

$$m_i = \operatorname{rank} \mathbf{C}_i(\mathbf{S}) - \operatorname{rank} \mathbf{C}_{i-1}(\mathbf{S}) , \quad i \geqq 1 ; \quad \mathbf{C}_0(\mathbf{S}) = 0 .$$

Each z uniquely determines $p + 1$ unordered sets

$$\mathbf{t}^{(k)} = \{j : n_j \geqq k\} = (t_1^{(k)}, \ldots, t_{m_k}^{(k)}) , \quad t_1^{(k)} < \ldots < t_{m_k}^{(k)}$$

such that $\mathbf{t}^{(p)} \subset \mathbf{t}^{(p-1)} \subset \ldots \subset \mathbf{t}^{(0)}$.

Consider the following procedure which transforms the sets $\boldsymbol{t}^{(k)}$ into another collection of sets

$$\boldsymbol{q}^{(k)} = (q_1^{(k)}, \ldots, q_{m_k}^{(k)}) ; \quad k = 1, \ldots, p .$$

For $k = 1$ let $\boldsymbol{q}^{(1)} = \boldsymbol{t}^{(1)}$ and let p_1 be a permutation such that $p_1(t_s^{(1)}) = s$; $s = 1, \ldots, m_1$.

If the sets $\boldsymbol{q}^{(1)}, \ldots, \boldsymbol{q}^{(k)}$ and the permutations $p_1, \ldots, p_{k-1}$ are already determined, define the permutation p_k such that

$$p_k(q_s^{(k)}) = s ; \quad s = 1, \ldots, m_k ,$$

and set

$$\boldsymbol{q}^{(k+1)} = (q_1^{(k+1)}, \ldots, q_{m_{k+1}}^{(k+1)}) ,$$

$$q_s^{(k+1)} = p_k \circ p_{k-1} \circ \ldots \circ p_1(t_s^{(k+1)}) ; \quad s = 1, \ldots, m_{k+1} .$$

Let

$$\boldsymbol{B} = [\boldsymbol{b}_1 \mid \ldots \mid \boldsymbol{b}_m] , \quad \boldsymbol{b}_i \in \mathscr{R}^n$$

and denote by

$$\boldsymbol{C}(\boldsymbol{S}) = [\boldsymbol{b}_{q_1^{(1)}} \mid \ldots \mid \boldsymbol{b}_{q_{m_1}^{(1)}} \mid \boldsymbol{A}\boldsymbol{b}_{q_1^{(2)}} \mid \ldots \mid \boldsymbol{A}\boldsymbol{b}_{q_{m_2}^{(2)}} \mid \ldots$$
$$\ldots \mid \boldsymbol{A}^{p-1}\boldsymbol{b}_{q_1^{(p)}} \mid \ldots \mid \boldsymbol{A}^{p-1}\boldsymbol{b}_{q_{m_p}^{(p)}}] \in \mathscr{R}_n^n$$

the matrix formed by the first n linearly independent columns of $\boldsymbol{C}_n(\boldsymbol{S})$. Then the set $\mathscr{L}_m(n)$ can be represented as the union

$$\mathscr{L}_m(n) = \mathscr{L}_m^+(n) \cup \mathscr{L}_m^-(n)$$

of the disjoint sets

$$\mathscr{L}_m^+(n) = \{\boldsymbol{S} : \det \boldsymbol{C}(\boldsymbol{S}) > 0\}$$

$$\mathscr{L}_m^-(n) = \{\boldsymbol{S} : \det \boldsymbol{C}(\boldsymbol{S}) < 0\} .$$

For $\boldsymbol{z}$ fixed let $\mathscr{L}_m^{\sim +}(n) \subset \mathscr{L}_m^+(n)$ be the set of pairs $\boldsymbol{S}^{\sim} = [\boldsymbol{A}^{\sim}, \boldsymbol{B}^{\sim})$ such that

$$\boldsymbol{A}^{\sim} = \begin{bmatrix} \boldsymbol{G}_1 & \boldsymbol{H}_1 & & & \\ \boldsymbol{D}_2 & \boldsymbol{G}_2 & \boldsymbol{H}_2 & & \\ & \ddots & \ddots & & \\ 0 & & \boldsymbol{D}_{p-1} & \boldsymbol{G}_{p-1} & \boldsymbol{H}_{p-1} \\ & & & \boldsymbol{D}_p & \boldsymbol{G}_p \end{bmatrix} , \quad \boldsymbol{B}^{\sim} = \begin{bmatrix} \boldsymbol{D}_1 \\ \\ 0 \\ \\ \end{bmatrix} ,$$

where

$$\boldsymbol{G}_i \in \mathscr{R}_{m_i}^{m_i} , \quad \boldsymbol{H}_i \in \mathscr{R}_{s_i}^{m_i} \quad (s_i = n - m_1 - \ldots - m_i)$$

and the matrices

$$\boldsymbol{D}_i = \left[D_{jk}^{(i)}\right] \in \mathscr{R}^{m_i}_{m_{i-1}} ,$$

$$\boldsymbol{D}_i = \begin{array}{c} \begin{array}{cccccc} q_1^{(i)} & q_2^{(i)} & \dots & q_{m_i-1}^{(i)} & q_{m_i}^{(i)} \\ \downarrow & \downarrow & & \downarrow & \downarrow \end{array} \\ \left[\begin{array}{ccccc} \boxed{+} & & & & \\ & \boxed{+} & & \sim & \\ & & \ddots & & \\ & 0 & & \boxed{+} & \\ & & & & \boxed{+} \end{array}\right] \end{array}$$

have the property

$$D_{jk}^{(i)} = 0 \quad \text{for} \quad k < q_j^{(k)}$$
$$d_j^{(i)} = D_{j,q_j^{(i)}}^{(i)} > 0$$
$$D_{jk}^{(i)} - \text{non-prespecified for} \quad k > q_j^{(k)} .$$

Introduce also the set $\tilde{\mathscr{L}}_m^-(n) \subset \mathscr{L}_m^-(n)$ which differs from $\tilde{\mathscr{L}}_m^+(n)$ only by the assumption $d_p^{(p)} < 0$ (instead of $d_p^{(p)} > 0$). Let finally

$$\tilde{\mathscr{L}}_m(n) = \tilde{\mathscr{L}}_m^+(n) \cup \tilde{\mathscr{L}}_m^-(n) .$$

Now we are in position to formulate our main result:

Theorem.

1°. Table 1 defines the canonical sets for the action of orthogonal groups on sets of completely controllable LS.

Table 1

Set of LS	$\mathscr{L}_m^+(n)$	$\mathscr{L}_m^-(n)$	$\mathscr{L}_m(n)$	$\mathscr{L}_m(n)$
Transformation group	$\mathfrak{O}^+(n)$	$\mathfrak{O}^+(n)$	$\mathfrak{O}^+(n)$	$\mathfrak{O}(n)$
Canonical set	$\tilde{\mathscr{L}}_m^+(n)$	$\tilde{\mathscr{L}}_m^-(n)$	$\tilde{\mathscr{L}}_m(n)$	$\tilde{\mathscr{L}}_m^+(n)$

2°. Let X, $\mathfrak{G}$ and $\tilde{X}$ denote any set, transformation group and canonical set as described in Table 1. Then for each $\boldsymbol{S} \in X$ there exists an unique $g \in \mathfrak{G}$ such that $g(\boldsymbol{S}) = \boldsymbol{S}^\sim \in \tilde{X}$.

3°. The non-prespecified as zero entries of $\boldsymbol{S}^\sim$ (with regard to their position in $\boldsymbol{S}^\sim$) constitute a complete independent rational invariant for the action of $\mathfrak{G}$ on X.

Remark 1. Note that only the group $\mathfrak{O}^+(n)$ can be defined on $\mathscr{L}_m^+(n)$ or $\mathscr{L}_m^-(n)$ since the members of $\mathfrak{O}(n) \setminus \mathfrak{O}^+(n) = \mathfrak{O}^-(n)$ are not automorphisms on these sets.

Remark 2. When z is not preliminary fixed, the complete independent invariant consists both the surjective arithmetic invariant $a : X \to Z(n)$ and the independent rational invariant described in 3°.

The proof is organized as follows:

a) We first show that all elements of X are unstable. Hence if $\tilde{X}$ is a canonical set then in view of Proposition 1 the transformation $\mathbf{S} \mapsto \mathfrak{g}(\mathbf{S}) = \mathbf{S}^{\sim} \in \tilde{X}$ is unique.

b) After choosing an appropriate $\tilde{X} \subset X$ we prove that the conditions of Proposition 2 take place. At the same time we present a stable computational algorithm for obtaining the orthogonal serial canonical form (OSCF) $\mathbf{S}^{\sim}$ of $\mathbf{S}$.

Proof. Consider the set $\mathscr{L}_m(n)$ under the action of $\mathfrak{O}(n)$.

a) Since $\mathbf{C}(\mathfrak{g}(\mathbf{S})) = \mathbf{P}^T\mathbf{C}(\mathbf{S})$ we see that $\mathfrak{g}(\mathbf{S}) = \mathbf{S}$ imply $\mathbf{P} = \mathbf{I}_n$. This is a well known fact which follows from the complete controllability of $\mathbf{S}$.

b) We shall describe a numerically stable algorithm for transformation of $[\mathbf{A}, \mathbf{B})$ into the OSCF $[\mathbf{A}^{\sim}, \mathbf{B}^{\sim}) \in \mathscr{L}_m^{\sim +}(n)$ which is based on a sequence of QR-decompositions.

Step 1. Let $[\mathbf{A}, \mathbf{B}) = [\mathbf{A}_1, \mathbf{B}_1)$ and

$$\mathbf{B}_1 = \mathbf{U}_1 \left[\begin{array}{c} \mathbf{D}_1 \\ \hline 0 \end{array}\right]; \quad \mathbf{U}_1 \in \mathfrak{O}(n)\,, \quad \mathbf{D}_1 \in \mathscr{R}_m^{m_1}$$

be the QR-decomposition of $\mathbf{B}_1$ (a detailed consideration of the QR-decomposition technique may be found in [18], [19]). Then

$$\mathbf{P}_1^T\mathbf{A}_1\mathbf{P}_1 = \left[\begin{array}{c|c} \mathbf{G}_1^{\wedge} & \mathbf{H}_1^{\wedge} \\ \hline \mathbf{B}_2 & \mathbf{A}_2 \end{array}\right], \quad \mathbf{P}_1^T\mathbf{B}_1 = \mathbf{B}^{\sim} = \left[\begin{array}{c} \mathbf{D}_1 \\ \hline 0 \end{array}\right],$$

where $\mathbf{P}_1 = \mathbf{U}_1$, $\mathbf{G}_1^{\wedge} \in \mathscr{R}_{m_1}^{m_1}$, $[\mathbf{A}_2, \mathbf{B}_2) \in \mathscr{L}_{m_1}(s_1)$ and rank $\mathbf{B}_2 = m_2$.

Step 2. Let

$$\mathbf{B}_2 = \mathbf{U}_2 \left[\begin{array}{c} \mathbf{D}_2 \\ \hline 0 \end{array}\right]; \quad \mathbf{U}_2 \in \mathfrak{O}(s_1)\,, \quad \mathbf{D}_2 \in \mathscr{R}_{m_1}^{m_2}$$

be the QR-decomposition of $\mathbf{B}_2$. Then

$$(\mathbf{P}_1\mathbf{P}_2)^T\,\mathbf{A}_1\mathbf{P}_1\mathbf{P}_2 = \left[\begin{array}{c|cc} \mathbf{G}'_1 & \mathbf{H}'_1 & \\ \hline \mathbf{D}_2 & \mathbf{G}_2^{\wedge} & \mathbf{H}_2^{\wedge} \\ \hline 0 & \mathbf{B}_3 & \mathbf{A}_3 \end{array}\right]$$

and $(\mathbf{P}_1\mathbf{P}_2)^T\,\mathbf{B}_1 = \mathbf{B}^{\sim}$, where $\mathbf{P}_2 = \operatorname{diag}(\mathbf{I}_{m_1}, \mathbf{U}_2)$, $[\mathbf{A}_3, \mathbf{B}_3) \in \mathscr{L}_{m_2}(s_2)$ and rank $\mathbf{B}_3 =$ $= m_3$.

Step k. Extending this procedure for $k \geqq 3$ we obtain the pair $[\mathbf{A}_k, \mathbf{B}_k) \in$ $\in \mathscr{L}_{m_{k-1}}(s_{k-1})$ with rank $\mathbf{B}_k = m_k$ at the $(k-1)$-th step. Let

$$\mathbf{B}_k = \mathbf{U}_k \left[\begin{array}{c} \mathbf{D}_k \\ \hline 0 \end{array}\right]; \quad \mathbf{U}_k \in \mathfrak{O}(s_{k-1})\,, \quad \mathbf{D}_k \in \mathscr{R}_{m_{k-1}}^{m_k}$$

and set

$$\mathbf{P}_k = \operatorname{diag}(\mathbf{I}_{m_1+\ldots+m_{k-1}}, \mathbf{U}_k).$$

Final step. Denoting $\mathbf{P} = \mathbf{P}_1\mathbf{P}_2 \ldots \mathbf{P}_p$ one gets

$$[\mathbf{P}^T\mathbf{A}\mathbf{P}, \mathbf{P}^T\mathbf{B}) = [\mathbf{A}^\sim, \mathbf{B}^\sim) \in \mathscr{L}_m^{\sim +}(n)$$

(if necessary, $\mathbf{P}$ can be multiplied by a diagonal matrix with ± 1 diagonal entries in order to obtain $d_j^{(i)} > 0$; $i = 0, \ldots, p$; $j = 1, \ldots, m_i$).

If only the group $\mathfrak{O}^+(n)$ is considered, and if the obtained matrix $\mathbf{P} \in \mathfrak{O}^-(n)$, then the matrix $\operatorname{diag}(\mathbf{I}_{n-1}, -1)\,\mathbf{P}$ transforms $[\mathbf{A}, \mathbf{B})$ into $[\mathbf{A}^\sim, \mathbf{B}^\sim) \in \mathscr{L}_m^\sim(n)$.

It is easy to show that $\mathbf{Q} \in \mathfrak{O}(n)$ and $[\mathbf{Q}^T\mathbf{A}^\sim\mathbf{Q}, \mathbf{Q}^T\mathbf{B}^\sim) \in \mathscr{L}_m^{\sim +}(n)$ imply $\mathbf{Q} = \mathbf{I}_n$. Thus both conditions of Proposition 2 are valid and the theorem is proved. □

Of a special interest is the case $m = 1$. Here

$$\mathscr{L}_1^\sim(n) = \mathscr{H}(n) \times \mathscr{V}, \quad \mathscr{L}_1^{\sim +}(n) = \mathscr{H}^+(n) \times \mathscr{V}, \quad \mathscr{L}_1^{\sim -}(n) = \mathscr{H}^-(n) \times \mathscr{V},$$

where:

$\mathscr{V} \subset \mathscr{R}^n$ is the set of vectors $\mathbf{v} = [v_1 \mid 0 \mid \ldots \mid 0]^T$ with $v_1 > 0$;

$\mathscr{H}^+(n) \subset \mathscr{R}_n^n$ is the set of upper Hessenberg matrices $\mathbf{H} = [H_{ij}]$

(i.e. $H_{ij} = 0$ for $i > j + 1$) such that $H_{i,i-1} > 0$, $i = 2, \ldots, n$;

$\mathscr{H}^-(n) = \operatorname{diag}(\mathbf{I}_{n-1}, -1)\,\mathscr{H}^+(n)$ and

$\mathscr{L}_1^\sim(n) = \mathscr{L}_1^{\sim +}(n) \cup \mathscr{L}_1^{\sim -}(n)$.

The reduction of $\mathbf{S} \in \mathscr{L}_1(n)$ into the OSCF $\mathbf{S}^\sim$ can be accomplished also by $n - 1$ elementary Householder reflections $\mathbf{P}_1, \ldots, \mathbf{P}_{n-1}$, where

$$\begin{aligned} &\mathbf{P}_1 = \mathbf{H}_n(\mathbf{B}_1) \\ &\vdots \\ &\mathbf{P}_k = \operatorname{diag}(\mathbf{I}_{k-1}, \mathbf{H}_{n-k+1}(\mathbf{B}_k)), \quad k = 2, \ldots, n-1, \end{aligned}$$

and $\mathbf{H}_n(\mathbf{B})$ denotes the symmetric Householder matrix

$$2\mathbf{B}^\wedge\mathbf{B}^{\wedge T}/\mathbf{B}^{\wedge T}\mathbf{B}^\wedge - \mathbf{I}_n; \quad \mathbf{B}^\wedge = \mathbf{B} + \sqrt{(\mathbf{B}^T\mathbf{B})}\,[1\,0 \ldots 0]^T, \quad \mathbf{B} \in \mathscr{R}^n.$$

Note that in both the single-input ($m = 1$) and the multi-input ($m > 1$) cases the matrix $\mathbf{C}(\mathbf{S}^\sim) = \mathbf{P}^T\mathbf{C}(\mathbf{S})$ of the first n linearly independent columns of the controllability matrix $\mathbf{C}_n(\mathbf{S}^\sim) = \mathbf{P}^T\mathbf{C}_n(\mathbf{S})$, is an upper triangular matrix with non-zero diagonal entries. Hence the transformation into the OSCF is equivalent theoretically (but not numerically!) to a standard Gramm-Schmidt orthogonalization of $\mathbf{C}(\mathbf{S})$.

Let us now consider the number $N = N(z)$ of rational invariants for the multi-input case.

It may be shown that

$$N(z) = n^2/2 + n + L + M/2 - I,$$

where

$$L = \sum_{i=0}^{p-1} m_i m_{i+1}, \quad M = \sum_{i=1}^{p} m_i^2, \quad I = \sum_{i=1}^{p} \sum_{s=1}^{m_p} q_s^{(i)}.$$

The exact bounds for $N(z)$ are

$$N(z_1) \leqq N(z) \leqq N(z_2),$$

where

$N(z_1) = n(n+1)/2 + L - (mn - M), \quad z_1 = (0, \ldots, 0, n_{m-m_1+1}, \ldots, n_m),$
$n_{m-m_1+1} \leqq \ldots \leqq n_m;$
$N(z_2) = n(n+1)/2 + L, \quad z_2 = (n_1, \ldots, n_{m_1}, 0, \ldots, 0),$
$n_1 \geqq \ldots \geqq n_{m_1}.$

Note that in the generic case $m = m_1 = \ldots = m_{p-1}$ one has $n = (p-1)m + m_p$, and

$$N(z) = n(n+1)/2 + mn.$$

It is interesting to compare $N(z)$ with the number $N_{GL}(z)$ of the independent rational invariants relative to $\mathfrak{GL}(n)$.

It follows from the results in [4] that

$$N_{GL}(z) = Q + R,$$

where

$$R = \sum_{i=1}^{m} \sum_{j=1}^{m} \min(n_i, n_j)$$

and Q is the number of pairs (n_i, n_j) with $j < i$ and $n_j > n_i$.

On the other hand it can be shown that $R = M$. Hence

$$N_{GL}(z) = Q + M.$$

Straighforward calculations give

$$N(z) = N_{GL}(z) + n(n+1)/2.$$

Therefore

$$Q = (n - M)/2 + L - I$$

and

$$N_{GL}(z) = (n + M)/2 + L - I.$$

Note finally that the above results can be extended directly to the observable LS

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\mathbf{x}(t) = \boldsymbol{A}\,\mathbf{x}(t)$$

$$\mathbf{y}(t) = \mathbf{C}\,\mathbf{x}(t),$$

where $\mathbf{x}(t) \in \mathscr{R}^n$, $\mathbf{y}(t) \in \mathscr{R}^r$ and $\boldsymbol{A} \in \mathscr{R}_n^n$, $\mathbf{C} \in \mathscr{R}_n^r$.

4. NUMERICAL PROPERTIES OF THE ALGORITHMS FOR COMPUTING THE ORTHOGONAL SERIAL CANONICAL FORMS

In contrast to the canonical forms relative to general transformation groups, the OSCF can be obtained by stable computational algorithms.

The computational algorithm for transformation of multi-input LS into the OSCF is based on QR-decomposition of the matrices $\mathbf{B}_k$ at each step. This may be accomplished by the subroutine SQRDC from LINPACK [19]. It permits pivoting which in turn makes possible to obtain the decomposition of $\mathbf{B}_k$ in the form

$$\mathbf{B}_k = \mathbf{U}_k \left[\begin{array}{c|c} \mathbf{D}_{11} & \mathbf{D}_{12} \\ \hline \multicolumn{2}{c}{0} \end{array}\right] \mathbf{E}\,,$$

where $\mathbf{D}_{11}$ is upper triangular, and $\mathbf{E}$ is a permutation matrix. The latter reflects the moving of the columns of $\mathbf{B}_k$ during the decomposition. The determination of rank $\mathbf{B}_k$ (i.e. the dimension of $\mathbf{D}_{11}$) is discussed in details in [19]. Thus the obtained matrices $\mathbf{A}^{\sim}, \mathbf{B}^{\sim}$ are exact for systems whose matrices are $\mathbf{A} + \varepsilon\|\mathbf{A}\|$, $\mathbf{B} + \varepsilon\|\mathbf{B}\|$, where ε is a matrix with elements small multiple of the order of the relative machine precision in the floating point arithmetic of the computer used. Hence the above computational algorithm is numerically stable.

In the single-input case the reduction of LS into OSCF can be done also by a sequence of Householder reflections. First the vector $\mathbf{B}$ is transformed by one elementary reflection $\mathbf{P}_1$ and then the matrix $\mathbf{P}_1\mathbf{A}\mathbf{P}_1$ is reduced to upper Hessenberg form by $n - 2$ reflections. This can be realized, for example, using the subroutine ORTHES from EISPACK [18]. To accumulate the successive transformation matrices the subroutine ORTRAN [18] can be utilized. This reduction is also numerically stable.

In this way the computational algorithms for finding the OSCF have very good numerical properties and may be realized on the basis of the modern numerical linear algebra algorithms.

5. CONCLUSIONS

The paper is devoted to the study and construction of invariants and canonical forms for linear multivariable systems under the action of orthogonal transformation groups. This is motivated by the fact that the canonical forms relative to general transformation groups can not be obtained in a numerically stable way. On the other hand the problem of determining invariants and canonical forms for the action of various transformation groups is of independent theoretical interest.

Up to this moment, however, there are no rigorous investigations on orthogonal invariants and canonical forms in spite of the fact that the members of orthogonal groups are most favorable for computations. Moreover, the orthogonal canonical forms have the same advantages in the analysis and synthesis of linear systems as the canonical forms relative to the general groups.

In the present paper complete independent invariants of linear multivariable systems are determined for the action of orthogonal transformation groups. The corresponding orthogonal canonical forms are defined and stable computational algorithms for their construction are developed. The program realization of these algorithms involves the modern numerical linear algebra software.

The number of independent rational orthogonal invariants is given, and explicit expression for the number of independent rational invariants relative to the general linear group is presented.

(Received October 6, 1980.)

REFERENCES

[1] D. G. Luenberger: Canonical forms for linear multivariable systems. IEEE Trans. Autom. Control *AC-12* (1967), 290—293.

[2] P. Brunovský: A classification of linear controllable systems. Kybernetika *3* (1970), 173—187.

[3] R. E. Kalman: Kronecker invariants and feedback. Proc. NRL-MRC Conf. (L. Weiss Ed.), Academic Press, N.Y. 1972, pp. 459—471.

[4] V. M. Popov: Invariant description of linear time-invariant controllable systems. SIAM J. Control *10* (1972), 252—264.

[5] M. J. Denham: Canonical forms for the identification of linear multivariable systems. IEEE Trans. Autom. Control *AC-19* (1974), 646—656.

[6] K. Glover, J. C. Willems: Parametrization of linear dynamical systems: Canonical forms and identifiability. IEEE Trans. Autom. Control *AC-19* (1974), 640—646.

[7] J. Rissanen: Basis of invariants and canonical forms for linear dynamic systems. Automatica *10* (1974), 175—182.

[8] O. H. Bosgra, A. J. J. Van der Weiden: Input-output invariants for linear multivariable systems. IEEE Trans. Autom. Control *AC-25* (1980), 20—36.

[9] M. M. Konstantinov, P. Hr. Petkov, N. D. Christov: Synthesis of linear systems with desired equivalent form. J. Comp. & Appl. Math. *6* (1980), 27—35.

[10] M. M. Konstantinov, S. P. Patarinski, P. Hr. Petkov, N. D. Christov: Synthesis of linear control systems with prescribed dynamics. Proc. First Internat. Conf. on Math. Modeling, Aug. 1977 St. Louis (Miss.) vol. 3, pp. 1639—1648.

[11] M. M. Konstantinov, P. Hr. Petkov, N. D. Christov: The serial canonical form and synthesis of linear multivariable systems. Podstawy Sterowania *9* (1979), 295—308.

[12] M. M. Konstantinov, P. Hr. Petkov, N. D. Christov: Synthesis of linear multivariable systems with prescribed equivalent form. Systems Sci. *5* (1979), 381—394.

[13] D. Jordan, B. Sridhar: An efficient algorithm for calculation of Luenberger canonical form. IEEE Trans. Autom. Control *AC-18* (1973), 292—295.

[14] K. B. Datta: An algorithm to compute canonical forms in multivariable control systems. IEEE Trans. Autom. Control *AC-22* (1977), *129—132*.

[15] J. Hickin, N. K. Sinha: An efficient algorithm for transformation to canonical forms. IEEE Trans. Autom. Control *AC-22* (1977), 652—653.

[16] E. J. Davison, W. Gesing, S. H. Wang: An algorithm for obtaining the minimal realization of a linear time-invariant system and determining if a system is stabilizable-detectable. Proc. 1977 IEEE Conf. on Dec. & Control, pp. 777—781.

[17] J. D. Aplevich: Direct computation of canonical forms for linear systems by elementary matrix operations. IEEE Trans. Autom. Control *AC-19* (1974), 124—126.
[18] B. T. Smith, J. M. Boyle, J. J. Dongarra, B. S. Garbow, Y. Ikebe, V. C. Klema, C. B. Moler: Matrix Eigensystem Routines-EISPACK Guide. Springer-Verlag, Berlin 1976.
[19] J. J. Dongarra, J. R. Bunch, C. B. Moler, G. W. Stewart: LINPACK Users' Guide. SIAM, Philadelphia 1979.

Dr. Mihail M. Konstantinov, Institute of Engineering Cybernetics and Robotics, Akad. G. Bonchev Str., Bl. 12, 1113 Sofia. Bulgaria.
Dr. Petko Hr. Petkov, Dr. Nicolai D. Christov, Department of Automatics, Higher Institute of Mechanical and Electrical Engineering, 1156 Sofia. Bulgaria.