Laboratory. During January to June 1963 he was a Visiting Scholar at the University of California, Berkeley; during 1969 to 1970 he was a Guggenheim Fellow and a Visiting Professor at the Indian Institute of Science, Bangalore. In 1972 he was a UK Science Research Council Fellow at Imperial College, London, England. His research interests are generally in statistical data processing in communications and control. He is Editor of a Prentice-Hall series of books on information and system sciences.

Dr. Kailath is a member of the Institute of Mathematical Statistics, Commission VI of URSI (International Scientific Radio Union), Sigma Xi, and several other scientific societies. He is a member of the Editorial Board of the IEEE PRESS, an Associate Editor of the IEEE TRANSACTIONS ON INFORMATION THEORY, and a member of the Administrative Committees of the IEEE Information Theory Group and the IEEE Control Systems Society.

Parametrizations of Linear Dynamical Systems: Canonical Forms and Identifiability

KEITH GLOVER, MEMBER, IEEE, AND JAN C. WILLEMS, MEMBER, IEEE

Abstract—We consider the problem of what parametrizations of linear dynamical systems are appropriate for identification (i.e., so that the identification problem has a unique solution, and all systems of a particular class can be represented). Canonical forms for controllable linear systems under similarity transformation are considered and it is shown that their use in identification may cause numerical difficulties, and an alternate approach is proposed which avoids these difficulties. Then it is assumed that the system matrices are parametrized by some unknown parameters from *a priori* system knowledge. The identiability of such an arbitrary parametrization is then considered in several situations. Assuming that the system transfer function can be identified asymptotically, conditions are derived for local and global identifiability. Finally, conditions for identifiability from the output spectral density are given for a system driven by unobserved white noise.

I. Introduction

CONSIDER the standard linear dynamical systems:

$$\frac{dx(t)}{dt} = Ax(t) + Bu(t), y(t) = Cx(t) + Du(t) \quad (1)$$

or

$$x(k+1) = Ax(k) + Bu(k), y(k) = Cx(k) + Du(k), \quad (2)$$

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, $y \in \mathbb{R}^p$.

A parametrization of the system matrices (A,B,C,D) is then a $\mathcal{C}'$ (i.e., continuously differentiable on Ω) function $(A,B,C,D)(\alpha)$: $\Omega \subset \mathbb{R}^q \rightarrow \mathbb{R}^{n(n+m+p)+mp}$. That is the system matrices are parametrized by the unknown parameters α. In the context of identifying such dynamical systems the following two properties of a parametrization are desirable.

Property 1: The parametrization should be identifiable in some sense.

Property 2: All systems in an appropriate class can be represented by the parametrization.

Canonical forms are one approach to this parametrization problem and have been suggested by several authors (e.g., Mayne [9], Weinert and Anton [15]). In Section II we will give some general comments on using canonical forms.

Canonical parametrizations are useful (and necessary) when there is very little *a priori* system knowledge except, perhaps, the system order. An alternative approach can be used when there is sufficient *a priori* information from, for example, physical considerations, to write down the system matrices as functions of relatively few unknown parameters, α, as $(A,B,C,D)(\alpha)$. The advantages of such models are that the prior knowledge is conveniently summarized and the resulting state variables and parameters have a physical interpretation. Property 2 above is then automatically satisfied since the prior knowledge determines the class of systems of interest, and hence all systems of order n need not be represented as is the case for canonical forms. Therefore for these parametrizations identifiability is the prime concern and conditions for identifiability are given in Section III.

In Section IV the identifiability of parametrizations of systems driven by unobserved white noise is considered.

II. Remarks on Using Canonical Forms in Identification

The canonical forms normally considered for identification are those of the triple (A,B,C) of given dimension,

Manuscript received January 16, 1974; revised July 29, 1974. Portions of the results in this paper were presented at the 3rd International Federation of Automatic Control Symposium, The Hague, The Netherlands, June 1973. This paper is based in part on the Ph.D. dissertation by K. Glover submitted to the University of Massachusetts Institute of Technology, Cambridge, Mass.

K. Glover is with the Department of Electrical Engineering Systems, University of Southern California, Los Angeles, Calif. 90007.

J. C. Willems is with the Mathematical Institute, University of Groningen, Groningen, The Netherlands.

under a similarity transformation, $(A,B,C) \xrightarrow{T} (TAT^{-1}, TB, CT^{-1})$, with the assumption of controllability. Similarity transformations are considered since for minimal systems all input–output equivalent systems are related by a similarity transformation. This problem has been considered by Popov [12], Mayne [9], and others including Denham [3], to which the reader is referred to for definitions and properties of canonical forms, invariants, completeness, and independence.

Canonical forms for multivariable systems necessarily consist of several separate parametrizations each one determined by a set of indices (e.g., the Kronecker invariants, minimal indices). In all the available canonical forms there is one parametrization which has the maximum number of free parameters, and will be generic in that almost all controllable systems of order n can be represented using this parametrization. However to represent the boundary of this first parametrization many more parametrizations may be required, each containing fewer degrees of freedom (see Denham [3]). Identifying a system from noisy data to be in the nongeneric form is generally not a statistically well-posed problem, since it will require estimating a determinant to be exactly 0. Further, numerical difficulties are likely to occur when representing systems near the boundary of one parametrization.

Example

To justify these general remarks for a particular canonical form we now give details of canonical form given implicitly in [12], and illustrate how the above phenomenon may occur.

Define the controllability matrix,

$$W(A,B) \triangleq [B, AB, \cdots, A^{n-1}B],$$

let

$$\kappa = (n_1, n_2, \cdots, n_m)$$

and define,

$$P(A,B,\kappa) \triangleq [b_1, Ab_1, \cdots, A^{n_1-1}b_1, \cdots, b_m, Ab_m, \cdots, A^{n_m-1}b_m].$$

Then κ is called the set of Kronecker invariants if the columns of $P(A,B,\kappa)$ are the first columns of $W(A,B)$, beginning at the left, to form a basis for $\mathbb{R}^n$.

A canonical form for the controllable triple (A,B,C) with rank $B = m$, is now given by a family of parametrizations, one for each set of Kronecker invariants, κ, such that $n_1 + n_2 + \cdots + n_m = n$ and $n_i > 0$ for $i = 1,2,\cdots,m$.

$$\tilde{A} = \begin{bmatrix} A_{11} & A_{12} \cdots A_{1m} \\ A_{21} & \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ A_{m1} \cdots & A_{mm} \end{bmatrix} \qquad \tilde{B} = \begin{bmatrix} B_{11} & B_{12} \cdots B_{1m} \\ B_{21} & \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ B_{m1} \cdots & B_{mm} \end{bmatrix}.$$

$\tilde{C}$: completely free. In here,

$$\underset{(n_j \times n_i)}{A_{ji}} = \begin{cases} \left[\begin{array}{c|c} 0_{1,n_i-1} & \alpha_{ii0} \\ \hline & \alpha_{ii1} \\ I_{n_i-1} & \cdot \\ & \cdot \\ & \cdot \\ & \alpha_{iin_i-1} \end{array}\right] & \text{for } j = i, \\[2ex] \left[\begin{array}{c|c} & \alpha_{ij0} \\ & \alpha_{ij1} \\ & \cdot \\ 0_{n_j,n_i-1} & \cdot \\ & \cdot \\ & \alpha_{ijk} \\ & 0 \\ & \cdot \\ & \cdot \\ & \cdot \\ & 0 \end{array}\right] & \begin{array}{l} \text{for } j \neq i, \\ k = \min(n_i, n_j - 1), \ \text{for } j < 1, \\ k = \min(n_i, n_j) - 1, \ \text{for } j > i, \end{array} \end{cases}$$

$$\underset{(n_j \times 1)}{B_{ji}} = \begin{cases} 0_{n_j,1} & j \neq i \quad n_i \neq 0, \\ \begin{bmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix} & j = i \quad n_i \neq 0. \end{cases}$$

Furthermore, given any triple (A,B,C) the transformation taking it to canonical form is given by $T = P^{-1}(A,B,\kappa)$.

The parametrization for the case when the first n columns of $W(A,B)$ are independent is in fact generic. If the Kronecker invariants for this case are denoted $\hat{\kappa} = (\hat{n}_1, \hat{n}_2, \cdots, \hat{n}_m)$ then,

$$n_j = \begin{cases} \left\lceil \dfrac{n}{m} \right\rceil & j = 1,2,\cdots, \\ \left\lceil \dfrac{n}{m} \right\rceil - 1 & j = l+1,\cdots,m, \end{cases}$$

where

$$l = m - \left(m\left\lceil \frac{n}{m} \right\rceil - n\right).$$

To show that this parametrization is generic it is sufficient to show that the set, $S_{\hat{\kappa}}$, of triples (A,B,C) which can be represented by this parametrization forms an open dense subset of the set of controllable triples (A,B,C). $S_{\hat{\kappa}}$ is open since it is the inverse image of an open set under a continuous function given by

$$S_{\hat{\kappa}} = \left\{(A,B,C) \middle| \det\left[W(A,B)\begin{bmatrix} I \\ 0 \end{bmatrix}\right] \neq 0\right\}.$$

To show that $S_{\hat{\kappa}}$ is dense consider any triple $(A,B,C) \notin S_{\hat{\kappa}}$, add $\epsilon\,(\hat{A},\hat{B},\hat{C})$ where $(A,B,C) \in S_{\hat{\kappa}}$, then it is straightforward to show that $(A + \epsilon\,\hat{A}, B + \epsilon\,\hat{B}, C + \epsilon\,\hat{C}) \in S_{\hat{\kappa}}$ for all $\epsilon > 0$ and sufficiently small. Hence the triple (A,B,C) can be approached arbitrarily closely by an element of $S_{\hat{\kappa}}$. Note however that as $\epsilon \to 0$ the matrix $T^{-1} = P(A,B,\hat{\kappa})$ approaches singularity, therefore, the representation of $(A + \epsilon \hat{A}, B + \epsilon \hat{B}, C + \epsilon \hat{C})$ in the canonical form, given by

$$(T(A + \epsilon \hat{A})T^{-1}, T(B + \epsilon \hat{B}), (C + \epsilon \hat{C})T^{-1})$$

will generally have some elements in the A matrix tend to infinity as $\epsilon \to 0$. Hence, the representation will become numerically very sensitive as $\epsilon \to 0$ in that some entries of the A matrix will become arbitrarily large cancelled by some in the C matrix becoming very small. ■

The problem with using a true canonical form is due to the requirement that no system has more than one representative in the canonical form. Therefore no two parametrizations may "overlap" and hence the boundary between two parametrizations will necessarily give problems.

For identification it is not essential that every system have a unique representation, rather it is sufficient that the free parameters in any of the separate parametrizations are identifiable, and that every system can be represented by at least one parametrization. For example both these requirements are satisfied in the Luenberger type I pseudocanonical form [8], which is the same as the above canonical form except that for $\kappa = (k_1,k_2,\cdots,k_m)$, α_{ijk} is defined for $k = 0,1,\cdots,k_j - 1$, and the set κ is any set such that $\det(P(A,B,\kappa)) \neq 0$. This set of parametrizations is not truly canonical since in general a particular system could be represented by several different parametrizations. However within any single parametrization a system will not have more than one representation. In fact every parametrization will be generic. Therefore it is just necessary to find a set of integers κ, such that $\det[P(A,B,\kappa)]$ is sufficiently large and then the corresponding parametrization will represent all systems of order n, in a large neighborhood of (A,B,C). This is now statistically and numerically well posed and a procedure for this is given in Glover [4].

It is therefore concluded that in identification, the use of canonical forms (which have often been derived for other purposes) may not be advisable due to numerical and statistical problems. It is instead recommended that sets of parametrizations be used that are individually identifiable and generic, and such that any system of order n can be represented by at least one of the parametrizations.

III. Identifiability from the Transfer Function

Identifiability of a set of parameters roughly means that parameter estimates can be determined that are asymptotically exact. Identifiability will then depend on the data available, and in this section we assume that there is sufficiently good data to asymptotically identify the transfer function of the system but that no more information can be obtained (e.g., the response to a known initial condition is not available). There are many situations where the above is a valid assumption, and in such cases the natural definition of local identifiability is the following.

Definition 1: Let $(A,B,C,D)(\alpha)$: $\Omega \subset \mathbb{R}^q \to \mathbb{R}^N$ ($N = n(n + m + p) + mp$), be a parametrization of the system matrices (A,B,C,D) of a linear dynamical system such as (1) or (2). This parametrization is said to be *locally identifiable from the transfer function* at the point $\alpha \in \Omega$ if there exists $\epsilon > 0$ such that,

1) $\quad \|\alpha - \hat{\alpha}\| < \epsilon, \ \|\beta - \hat{\alpha}\| < \epsilon, \ \alpha, \beta \in \Omega,$

and

2) $$C(\alpha)(Is - A(\alpha))^{-1}B(\alpha) + D(\alpha) = C(\beta)(Is - A(\beta))^{-1}B(\beta) + D(\beta) \quad \text{for all } s \in \mathbb{C}(s \neq \lambda(A(\alpha)),\lambda(A(\beta)))$$

together imply $\alpha = \beta$. ■

In other words, in a neighborhood of $\hat{\alpha}$, there are no two systems with distinct parameters, which have the same transfer function. This definition is similar to the definition of "nondegeneracy" as given by Kalman [7]. Definition 1 is equivalent to requiring that the map from the parameters, α, into the Markov parameters is locally one-to-one. The following lemma gives a standard result on injective maps and is a direct consequence of the rank theorem (Narasimhan [10]).

Lemma 1: Let Ω be an open set in $\mathbb{R}^n$ and $f: \Omega \to \mathbb{R}^m$ be a c^k map with $k \geq 1$. Then if $(\partial f(x)/\partial x)$ has constant rank r in a neighborhood of $\hat{x}$, f is locally injective at $\hat{x}$ if and only if $r = n$.

The condition of Lemma 1 could be applied to the map from the unknown parameters, α, to the Markov parameters, $D(\alpha)$ and $C(\alpha)A^k(\alpha)B(\alpha)$ $k = 0,1,\cdots$, to give conditions for local identifiability. This is similar to sensitivity analysis and is given in Glover [4] (see also Cruz [2]). A more elegant condition can be obtained if it is known that the system is of minimal order in which case we know that all equivalent systems are related by similarity transformations. Theorem 1 now gives conditions for local identifiability of minimal systems.

Theorem 1: Let $(A,B,C,D)(\alpha):\Omega \subset \mathbb{R}^q \to \mathbb{R}^N$ (with Ω an open subset of $\mathbb{R}^q$) be a c' (i.e., continuously differentiable on Ω) parametrization of the system matrices (A,B,C,D) and suppose $(A,B,C,D)(\hat{\alpha})$ is minimal. Then

1) $(A,B,C,D)(\alpha)$ is locally identifiable from the transfer function at $\alpha = \alpha$ if and only if $F:GL(n) \times \Omega \to \mathbb{R}^N$ is locally injective at $T = I$ and $\alpha = \hat{\alpha}$, where

$$F(T,\alpha) \triangleq (TA(\alpha)T^{-1}, TB(\alpha), C(\alpha)T^{-1}, D(\alpha)).$$

2) If rank $\partial F(I,\alpha)/\partial(T,\alpha) = r$ for all α in some neighborhood of $\hat{\alpha}$, then $(A,B,C,D)(\alpha)$ is locally identifiable from the transfer function at $\alpha = \hat{\alpha}$ if and only if $r = n^2 + q$, or, equivalently if and only if $\det[X'(\hat{\alpha})X(\hat{\alpha})] \neq 0$, where

$$X(\alpha) = \left[\frac{\partial \tilde{F}}{\partial T}(I,\alpha); \frac{\partial \tilde{F}}{\partial \alpha}(I,\alpha)\right] = \left[\begin{array}{c|c} I_n \otimes A'(\alpha) - A(\alpha) \otimes I_n & \\ I_n \otimes B'(\alpha) & \\ - C(\alpha) \otimes I_n & M(\alpha) \\ O_{mp,n^2} & \end{array}\right].$$

$\tilde{F}$ is a reordering of F defined by

$$\bar{F}(T,\alpha) = \begin{bmatrix} \overline{\overline{TA(\alpha)T^{-1}}} \\ \overline{\overline{TB(\alpha)}} \\ \overline{\overline{C(\alpha)T^{-1}}} \\ \overline{\overline{D(\alpha)}} \end{bmatrix},$$

and

$$M(\alpha) = \begin{bmatrix} \frac{\partial \overline{\overline{A}}}{\partial \alpha}(\alpha) \\ \frac{\partial \overline{\overline{B}}}{\partial \alpha}(\alpha) \\ \frac{\partial \overline{\overline{C}}}{\partial \alpha}(\alpha) \\ \frac{\partial \overline{\overline{D}}}{\partial \alpha}(\alpha) \end{bmatrix}.$$

In here if $X' = [x_1x_2\cdots x_n]$ with $x_i \in \mathbb{R}^m$, then $\overline{\overline{X}}$ is the $nm \times 1$ vector given by $\overline{\overline{X}}' = [x_1',x_2',\cdots,x_n']$. Also $\otimes$ denotes Kronecker product (see Pease [11]).

Proof:

1) The necessity is straightforward and sufficiency is proved as follows.

First note that since $(A,B,C,D)(\alpha)$ is minimal there exists a neighborhood $W \subset \Omega$ of such that $(A,B,C,D)(\alpha)$ is also minimal for all $\alpha \in W$. (Since minimal systems form an open set in parameter space and the parametrization is assumed to be continuous.) Therefore when restricted to W all equivalent systems are related by a similarity transformation. Therefore the parametrization is locally identifiable if F is injective when restricted to $GL(n) \times V$, where $V \subset W$ is any open set containing $\hat{\alpha}$.

In order to prove the result we will prove the contrapositive. Assume therefore that the parametrization is not locally identifiable, then for all $\epsilon > 0$ there exists $T_\epsilon, S_\epsilon \in GL(n)$, $\alpha_\epsilon \neq \beta_\epsilon \in N_\epsilon(\hat{\alpha}) \subset W$ such that $F(T_\epsilon, \alpha_\epsilon) = F(S_\epsilon, \beta_\epsilon)$. Therefore we have that

$$S_\epsilon^{-1}T_\epsilon = W(\beta_\epsilon)W'(\alpha_\epsilon)[W(\alpha_\epsilon)W'(\alpha_\epsilon)]^{-1}$$

where

$$W(\alpha) = [B(\alpha),A(\alpha)B(\alpha),\cdots,A^{n-1}(\alpha)B(\alpha)].$$

$S_\epsilon^{-1}T_\epsilon$ is therefore a continuous function of $(\alpha_\epsilon,\beta_\epsilon)$ since $W(\alpha)$ has full rank for all $\alpha \in W$ by the reachability assumption. Therefore $\|S_\epsilon^{-1}T_\epsilon - I\|$ can be made arbitrarily small by taking ϵ sufficiently small and $F(S_\epsilon^{-1}T_\epsilon, \alpha_\epsilon) = F(I,\beta_\epsilon)$. Hence there does not exist a neighborhood of (I,α) in which F is injective, and thus F is not locally injective.

2) To prove this we use part (1) and Lemma 1.

The Jacobian matrix of $\tilde{F}$ can readily be computed as,

$$\left[\frac{\partial \tilde{F}}{\partial T}(T,\alpha), \frac{\partial \tilde{F}}{\partial \alpha}(T,\alpha)\right] = \begin{bmatrix} T \otimes T^{-1'} & 0 & 0 & 0 \\ 0 & T \otimes I & 0 & 0 \\ 0 & 0 & I \otimes T^{-1'} & 0 \\ 0 & 0 & 0 & I \otimes I \end{bmatrix} X(\alpha) \cdot \begin{bmatrix} T^{-1} \otimes I & 0 \\ 0 & I_q \end{bmatrix}.$$

Therefore since $T \in GL(n)$ the rank of the Jacobian of F at (T,α) is equal to the rank of $X(\alpha)$, hence the assumption that rank $X(\alpha) = r$ for all $\alpha \in N_\epsilon(\hat{\alpha})$ implies that rank $(\partial\tilde{F}(T,\alpha))/(\partial(T,\alpha)) = r$ for all (T,α) in some neighborhood of $(I,\hat{\alpha})$. Therefore the assumptions of Corollary 1 are valid and the result follows immediately. ■

Theorem 1 gives a comparatively simple test for local identifiability which is significantly simpler than the methods based on the information matrix (e.g., Rothenberg [13], Tse [14]). The information matrix is a quite general approach and indeed also gives approximations to the covariance of the parameter estimates, however for the problem considered here it gives computationally difficult tests. The condition of Theorem 1 is complementary to other approaches in that it examines the identifiability of the parametrization alone and should be checked before the inputs and observations are considered.

As stated, Theorem 1 requires the evaluation of the determinant of a sparse matrix of size $n^2 + q$, however this can be reduced if $(A,B,C,D)(\alpha)$ is transformed into some standard form such as the Luenberger type I form [8]. The matrix $M(\alpha)$ will also have to be transformed, but then some simple operations will reduce the problem to finding the rank of an $(n(m + p) + mp) \times q$ matrix that is essentially the smallest possible for this problem.

If a parametrization is locally identifiable, this ensures that any consistent algorithm which minimizes some cost function over the parameters will be well posed and have a unique solution in some neighborhood of the nominal values. Furthermore, if a parametrization is locally identifiable for all values of $\alpha \in \Omega$ then an algorithm will always be well posed but may converge to one of several solutions depending on the initial parameter estimates and the actual data received. This is the problem of global identifiability which will now be discussed.

Global Identifiability

Disadvantages of the concept of local identifiability are that the nominal values, $\hat{\alpha}$, must be known and the size of the neighborhood of $\hat{\alpha}$ is in general not easily found. It is thus desirable to attempt to generalize the result of the theorem to global identifiability.

Definition 2: Let $(A,B,C,D)(\alpha):\Omega \subset \mathbb{R}^q \to \mathbb{R}^{n(n+m+p)+mp}$ be a parametrization of the system matrices (A,B,C,D). This parametrization is said to be *globally identifiable from the transfer function* if,

1) $C(\alpha)(Is - A(\alpha))^{-1}B(\alpha) + D(\alpha) = C(\beta)(Is - A(\beta))^{-1}B(\beta) + D(\beta)$ for all $s \in \mathbb{C}$ $(s \neq \lambda(A(\alpha)),\lambda(A(\beta)))$, and

2) $(A,B,C,D)(\alpha)$ is minimal,

together imply that $\alpha = \beta$. ■

Condition (2) could be deleted in the above definition but then the definition would be very restrictive since most useful parametrizations admit multiple representations of nonminimal systems.

The following proposition gives a sufficient condition for global identifiability, when the parametrization is affine, (i.e., a linear map plus an offset).

Proposition 1: An affine parametrization $(A,B,C,D)(\alpha):\Omega \subset \mathbb{R}^q \to \mathbb{R}^{n(n+m+p)+mp}$ is globally identifiable if

$\det[Y'(\alpha,\beta)Y(\alpha,\beta)] \neq 0$ for all $\alpha,\beta \in \Omega$, where

$$Y(\alpha,\beta) = \begin{bmatrix} Z(\alpha,\beta) & 0 & M \\ 0 & Z(\beta,\alpha) & -M \end{bmatrix},$$

$$Z(\alpha,\beta) = \begin{bmatrix} I \otimes A'(\alpha) - A(\beta) \otimes I \\ I \otimes B'(\alpha) \\ -C(\beta) \otimes I \\ 0 \end{bmatrix},$$

and M is defined as given in Theorem 1.

Proof: Since we are only concerned with minimal systems global identifiability is implied if the equations $TA(\alpha) = A(\beta)T$, $TB(\alpha) = B(\beta)$, $C(\alpha) = C(\beta)T$, $D(\alpha) = D(\beta)$, have a unique solution for all $\alpha,\beta \in \Omega$ and $T \in GL(n)$. Let q_1 and q_2 be the vectors formed by listing respectively the elements of $(T - I)$ and $(T^{-1} - I)$ by rows. Then it is easily verified that the above equations are equivalent to $[q_1', q_2', \alpha' - \beta']Y'(\alpha,\beta) = 0$, since $(A,B,C,D)(\alpha)$ is affine. Therefore since $\det(Y'(\alpha,\beta)Y(\alpha,\beta)) \neq 0$ the nullspace of $Y(\alpha,\beta) = N(Y(\alpha,\beta)) = \{0\}$ and the result is thus verified. ■

Remarks

1) A somewhat more restrictive sufficient condition for global identifiability is that $N(Z(\alpha,\beta),M) = \{0\}$ for all $\alpha,\beta \in \Omega$. We remark that this condition is in fact satisfied by the canonical forms given in [9] and [12].

2) Note the similarity between the condition in Remark 1 and the condition of Theorem 1. However local identifiability for all $\alpha \in \Omega$ does not in general imply global identifiability. An open conjecture is that local identifiability for all $\alpha \in \mathcal{R}^{n(m+p)}$ implies global identifiability when the parametrization is affine, and $D = 0$.

IV. Identifiability from the Output Spectral Density

In this section we will consider the identifiability of a continuous time linear stationary system under the following assumptions.

Assumption 1: The input $u(t)$ is not observed directly, but is assumed to be a white noise process with $E(u(t)u'(T)) = I\delta(t - T)$.

Assumption 2: The matrix A is asymptotically stable, (i.e., the eigenvalues of A are strictly in the left half-plane).

Assumption 3: The system has reached steady state when the observations begin (i.e., the output process $y(t)$ is a stationary process).

Assumption 4: The system to be identified is globally minimal, i.e., the dimension of the state is less than or equal to that of any other system having the same output spectral density when driven by white noise for example there are no "all-pass" factors (see Anderson [1]).

Under these assumptions the most information that may be obtained from the output observations is the output spectral density, $\Phi(s) = G(s)G'(-s)$, where $G(s) = C(Is - A)^{-1}B + D$. This motivates the following definition.

Definition 3: Let $(A,B,C,D)(\alpha):\Omega \subset \mathcal{R}^q \to \mathcal{R}^{n(n+m+p)+mp}$ be a parametrization of the system matrices (A,B,C,D). This parametrization is said to be *locally identifiable from its output spectral density* at the point $\hat{\alpha} \in \Omega$ if there exists an $\epsilon > 0$ such that

1) $$\|\alpha - \hat{\alpha}\| < \epsilon,\ \|\beta - \hat{\alpha}\| < \epsilon;\ \alpha,\beta \in \Omega,$$

and

2) $$G(s,\alpha)G'(-s,\alpha) = G(s,\beta)G'(-s,\beta) \quad \text{for all } s \in \mathbb{C},\ s \neq \lambda(A(\alpha)),\lambda(A(\beta)),$$

together imply $\alpha = \beta$, where

$$G(s,\alpha) = C(\alpha)(Is - A(\alpha))^{-1}B(\alpha) + D(\alpha). \quad ■$$

Determining a transfer function, $G(s)$, from $\Phi(s) = G(s)G'(-s)$ is called the spectral factorization problem (Youla [17]), or the inverse problem of stationary covariance generation (Anderson [1]). In order to consider identifiability we need to characterize all equivalent solutions to the spectral factorization problem. Now if the transfer function satisfies the additional "minimum phase" condition that the rank $G(s) = m$ for all s such that $\mathrm{Re}(s) > 0$ then from Youla's results all systems equivalent to (A,B,C,D) are given by $(TAT^{-1}, TBU, CT^{-1}, DU)$ where $T \in GL(n)$ and U is orthogonal, $UU' = I$. In this case the identifiability question is very similar to that for the transfer function case. However if the system is not known to be minimum phase the characterization of equivalent systems is more difficult as given by the following lemma.

Lemma 2: If (A_1,B_1,C_1,D_1) and (A_2,B_2,C_2,D_2) are globally minimal systems then

$$G_1(s)G_1'(-s) = G_2(s)G_2'(-s)$$

(where $G_i(s) = C_i(Is - A_i)^{-1}B_i + D_i$, $i = 1,2$) if and only if there exists $T \in GL(n)$ and $Q = Q'$ such that

$$\left.\begin{aligned} A_1 &= TA_2T^{-1},\ C_1 = C_2T^{-1} \\ QA_1' + A_1Q &= -B_1B_1' + TB_2B_2'T' \\ QC_1' &= -B_1D_1' + TB_2D_2',\ D_1D_1' = D_2D_2'. \end{aligned}\right\} \quad (3)$$

Further if D_1D_1' is nonsingular the above is equivalent to there being a similarity transformation between the Kalman filters of the two systems, or equivalently of their innovations representations. (Kailath and Geesey [5], [6].) ■

This lemma is a straightforward consequence of Lemma 2 in Anderson [1], and standard filtering arguments (see also Willems [16]).

The local identifiability question is now equivalent to there being a locally unique solution to (3) with $(A_1,B_1,C_1,D_1) = (A,B,C,D)(\alpha)$ and $(A_2,B_2,C_2,D_2) = (A,B,C,D)(\beta)$ (i.e., $\alpha = \beta$, $Q = 0$, $T = I$). The following theorem gives conditions for local identifiability and can be proven in an analogous manner to Theorem 1.

Theorem 2: Let $(A,B,C,D)(\alpha):\Omega \subset \mathcal{R}^q \to \mathcal{R}^{n(n+m+p)+mp}$ (with Ω an open set in $\mathcal{R}^q$) be a c' parametrization of the system matrices (A,B,C,D) of the continuous time system (1) satisfying Assumptions 1–4. Then this parametrization is locally identifiable from its output spectral density at $\boldsymbol{\alpha} \in \Omega$, if the following linear equations in $(\delta\boldsymbol{\beta},\delta B,\delta D,\delta Q)$, have a unique solution (i.e., zero):

1) $\delta Q = \delta Q'$,
2) $(\hat{A}\delta Q + \delta B\hat{B}' - \delta T\hat{B}\hat{B}') + (\hat{A}\delta Q + \delta B\hat{B}' - \delta T\hat{B}\hat{B}')' = 0$,
3) $\delta Q\hat{C}' = -\delta B\hat{D}' - \hat{B}\delta D' + \delta T\hat{B}\hat{D}'$,
4) $\delta D\hat{D}' + \hat{D}\delta D' = 0$,
5) $$\begin{bmatrix} \overline{\overline{\delta T\hat{A} - \hat{A}\delta T}} \\ \overline{\overline{\delta B}} \\ \overline{\overline{-\hat{C}\delta T}} \\ \overline{\overline{\delta D}} \end{bmatrix} = M(\hat{\boldsymbol{\alpha}})\delta\boldsymbol{\beta}$$

where $M(\boldsymbol{\alpha})$ is defined in Theorem 1, and $(\hat{A},\hat{B},\hat{C},\hat{D}) = (A,B,C,D)(\hat{\boldsymbol{\alpha}})$. ■

In general fewer parameters can be identified than when input observations are also allowed and thus $G(s)$ rather than $\Phi(s) = G(s)G'(-s)$ is observed. In fact if $m = p$ the number of identifiable parameters is bounded by $2np + \frac{1}{2}p(p+1)$ which is $\frac{1}{2}p(p-1)$ less than when input observations are allowed.

Simplified conditions could be obtained if there are the minimum number of white noise inputs, in which case $Q = 0$ (locally) is implied without restrictions on $B(\boldsymbol{\beta})$ and T (assuming the minimal polynomial of the inverse system matrix $(A - BD^{-1}C)$ equals its characteristic polynomial —see Willems [16]).

Analogous results for all the above results for discrete time systems can be found in [4].

V. Conclusions

This paper has considered some of the problems associated with parameterizing linear dynamical systems for identification. The discussion in Section II on canonical forms illustrates that the requirement that no system have more than one representation in the canonical form is not necessary for identification parametrizations and in fact can cause numerical difficulties. A modified approach using globally identifiable parametrizations is suggested which avoids these problems.

Sections III and IV consider the identifiability of arbitrary parametrizations that would be obtained from the *a priori* information on the system. It is recommended that the resulting conditions for local (and global) identifiability are tested prior to the application of any identification algorithm to ensure that the parametrization is well chosen.

Finally it is remarked that extensions of these results to partial identifiability and identification in the presence of feedback can be found in [4].

References

[1] B. D. O. Anderson, "The inverse problem of stationary covariance generation," *J. Statist. Phys.*, vol. 1, pp. 133–147, 1969.

[2] J. B. Cruz, *System Sensitivity Analysis*. London, England: Dowden, Hutchinson, and Ross, 1973.

[3] M. J. Denham, "Canonical forms for the identification of multivariable linear systems," this issue, pp. 646–656.

[4] K. Glover, "Structural aspects of system identification," Ph.D. dissertation, Dep. Elec. Eng., Electron. Syst. Lab., Mass. Inst. Tech., Cambridge, Mass., Rep. ESL-R-516.

[5] T. Kailath and R. A. Geesey, "An innovations approach to least squares estimation—Part IV: Recursive estimation given lumped covariance functions," *IEEE Trans. Automat. Contr.*, vol. AC-16, pp. 720–726, Dec. 1971.

[6] ——, "An innovations approach to least squares estimation—Part V: Innovations representations and recursive estimation in colored noise," *IEEE Trans. Automat. Contr.*, vol. AC-18, pp. 435–453, Oct. 1973.

[7] R. E. Kalman, "On structural properties of linear, constant, multivariable systems," in *Proc. 3rd Int. Fed. Automat. Contr. Congr.*, London, England, Paper 6.A., 1966.

[8] D. G. Luenberger, "Canonical forms for linear multivariable systems," *IEEE Trans. Automat. Contr.*, vol. AC-12, pp. 290–293, June 1967.

[9] D. Q. Mayne, "A canonical model for identification of multivariable linear systems," *IEEE Trans. Automat. Contr.*, vol. AC-17, pp. 728–729, Oct. 1972.

[10] R. Narasimham, *Analysis on Real and Complex Manifolds*. Amsterdam, The Netherlands: North-Holland.

[11] M. C. Pease, *Methods in Matrix Algebra*. New York: Academic, 1965.

[12] V. M. Popov, "Invariant description of linear time-invariant controllable systems," *SIAM J. Contr.*, vol. 10, no. 2, pp. 252–264, 1972.

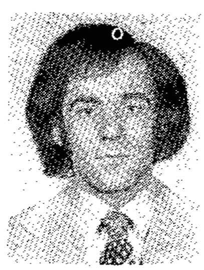
[13] T. J. Rothenberg, "Identification in parametric models," *Econometrica*, vol. 39, no. 3, pp. 577–591, 1971.

[14] E. Tse, "Information matrix and local identifiability of parameters," in *1973 Joint Automatic Control Conf., Preprints*.

[15] H. Winert and J. Anton, "Canonical forms for multivariable system identification," in *Proc. IEEE Conf. Decision and Control*, 1972, pp. 37–39.

[16] J. C. Willems, "Least squares stationary optimal control and the algebraic Riccati equation," *IEEE Trans. Automat. Contr.*, vol. AC-16, pp. 621–634, Dec. 1971.

[17] D. C. Youla, "On the factorization of rational matrices," *IRE Trans. Inform. Theory*, vol. IT-7, pp. 172–189, July 1961.



Keith Glover (S'71–M'73) was born in Bromley, Kent, England, on April 23, 1946. He received the B.Sc.(Eng.) degree in electrical engineering from the Imperial College of Science and Technology, London University, London, England, in 1967, and the S.M., E.E., and Ph.D. degrees in electrical engineering from the Massachusetts Institute of Technology, Cambridge, in 1971, 1971, and 1973, respectively.

From 1967 to 1969 he worked on the development of digital communication equipment at the Marconi Company, Chelmsford, Essex, England, and he was a Kennedy Memorial Fellow at M.I.T. from 1969 to 1971. He is currently an Assistant Professor of Electrical Engineering at the University of Southern California, Los Angeles. His present research interests are in system identification and linear system theory.

Jan C. Willems (S'66–M'68) was born in Bruges, Belgium, in 1939. He graduated in electrical and mechanical engineering from the University of Ghent, Belgium, in 1963, received the M.S. degree in electrical engineering from the University of Rhode Island, Kingston, in 1965, and the Ph.D. degree in electrical engineering from Massachusetts Institute of Technology, Cambridge, in 1968.

From June 1968 he was an Assistant Professor in the Department of Electrical Engineering at M.I.T. until in February 1973 when he was appointed

to his present position of Professor of Systems and Control with the Mathematical Institute of the University of Groningen, Groningen, The Netherlands. He has held visiting appointments with the Department of Applied Mathematics and Theoretical Physics of the University of Cambridge, Cambridge, England, with the Center of Mathematical System Theory of the Department of Mathematics at the University of Florida, and with the Division of Engineering and Applied Physics of Harvard University. His area of main interest is in system theory. He is the author of *The Analysis of Feedback Systems* (M.I.T. Press, 1971) and was Associate Editor for Stability, Nonlinear Systems, and Distributed Systems of the IEEE Transactions of Automatic Control in 1971 to 1973. He is a member of the Society for Industrial and Applied Mathematics, and the Dutch General Systems Society.

Canonical Forms for the Identification of Multivariable Linear Systems

MICHAEL J. DENHAM, STUDENT MEMBER, IEEE

Abstract—**The advantage of using a unique parameterization in a numerical procedure for the identification of a system from operating records has been well established. In this paper several sets of canonical forms are described for state space models of deterministic multivariable linear systems; the members of these sets having therefore the required uniqueness property within the equivalence classes of minimal realizations of the system. In the identification of a stochastic system, it is shown how the problem depends also upon determining a unique factorization of the spectral density matrix of the system, and the sets of canonical forms obtained for the deterministic system are extended to this case.**

I. Introduction

A BASIC requirement of any successful identification algorithm is that it should lead to consistent estimates of the parameters of the unknown system. However, systems which have weighting functions that are close in some sense can have widely differing state space representations, with the result that, in general, the maximum likelihood estimates of the parameters of the state space model are not consistent. This is even true when there is only a single output, but in that case, if the state dimension is known, a unique canonical form can be specified and consistency can be established. This paper is essentially concerned with a review of some of the canonical forms which can be specified in the multivariable case.

In Section II, the necessary definitions for an accurate description of what is meant by a canonical form are given, culminating in the basic concept of a canonical form as the unique member of an equivalence class of a given set. In Section III, these definitions are applied in the case of the equivalence classes of minimal realizations of transfer functions. In this way, we can determine what properties the state space model must satisfy to be a unique minimal realization of the given input/output map. Several such canonical forms are described and the essential properties of completeness and independence of the parameters are established. The implications of determining these canonical forms from the Hankel matrix formed from the system Markov matrices (the minimal realization problem) are also discussed.

In Section IV, the particular problems raised in the identification of stochastic systems by nonuniqueness of the model are resolved. This is done by determining a set of canonical forms for the equivalence classes of stable, minimum phase factors of the spectral density matrix $\Phi(z)$ for a discrete time system.

Finally, in Section V, some of the implications of using canonical forms in identification algorithms are discussed, including the question of consistency of the estimated parameters.

II. Notation and General Definitions

Consider any set X. We can define an equivalence relation E on X and denote the equivalence of two elements $x, y \in X$ by xEy. We shall now relate some important definitions concerning the set X and its equivalence relation E [1].

Definition 1: A function $f: X \rightarrow S$ for some set S is an *invariant* for the equivalence relation E if, for any $x,y \in X$, then

$$xEy \Longrightarrow f(x) = f(y).$$

It is easy to see therefore that the equivalence relation E generates, for each $x \in X$, a disjoint set of *equivalence classes* or *orbits* in X which we will denote as

$$E(x) = \{y : yEx, \quad \text{for} \quad x,y \in X\}$$

The set of all such equivalence classes (i.e., for all $x \in X$) is called a *quotient set* or *orbit space*, and is denoted by X/E.

Manuscript received January 16, 1974; revised July 1, 1974. This work was supported in part by the United Kingdom Science Research Council.

The author is with the Department of Computing and Control, Imperial College of Science and Technology, London, England.