Automatica, Vol. 25, No. 1, pp. 87–95, 1989
Printed in Great Britain.

0005-1098/89 $3.00 + 0.00
Pergamon Press plc

Properties of the Parametrization of Monic ARMA Systems*

MANFRED DEISTLER† and MICHEL GEVERS‡

New properties of the parametrization of multi-input, multi-output monic ARMA systems with prescribed column degrees include a minimal cover for the set of all ARMA systems in terms of this parametrization.

Key Words—Parametrization; ARMA systems; identification.

Abstract—Some properties of the parametrization of (multi-input, multi-output) monic ARMA systems with prescribed column degrees are considered. In particular we give a minimal cover for the set of transfer functions corresponding to all ARMA systems in terms of this parametrization, we investigate the relation between the Kronecker indices and the prescribed column degrees and we investigate the boundary of a certain parameter space and of the corresponding set of transfer functions.

1. INTRODUCTION

FOR THE PURPOSE of identification usually the set of transfer functions corresponding to all ARMA systems with given input and output dimensions is broken up into parts in order to ensure a convenient parametrization for each part. By "convenient" we mean that each part should be described by an identifiable and finite-dimensional parameter space, and that the parametrization should be continuous. In addition, the set consisting of all these parts should be a cover for the set of transfer functions of all ARMA systems. A desirable property is that this cover should be nonredundant in a sense which will be explained later.

Especially in the multi-output case, several different ARMA parametrizations are used. Perhaps the most common are the overlapping parametrization of the manifold $M(n)$ of all transfer functions of order (or McMillan degree) n and the so-called Guidorzi or Echelon canonical forms (Kalman, 1974; Hazewinkel and Kalman, 1976; Guidorzi, 1981; Deistler, 1985; Gevers and Wertz, 1987). These have first been derived for the state space representations, where the McMillan degree is a natural quantity, being the minimal dimension of the state vector. ARMA systems derived from these state space parametrizations have several disadvantages. They are not necessarily *monic*; by monic we mean that the coefficient of the present output is the identity matrix. In addition, they are more naturally described in polynomials of the forward than of the backward shift operator. Of course, it is always possible to rewrite these ARMA models using polynomials in the backward shift operator, but by doing this some of the structural properties are lost (Deistler and Hannan, 1981; Gevers, 1986).

In this paper we consider a genuine ARMA parametrization and its relation to the parametrizations mentioned above. Unless otherwise stated, our ARMA models will be described using polynomials in the backward shift operator. The ARMA system will be monic (in the sense described above) and the maximum lags in each variable will be prescribed. With the addition of rank and coprimeness conditions, Hannan (1971) showed that this class of monic ARMA systems is identifiable. Properties of this class of monic ARMA models were described in Deistler *et al.* (1978), Deistler (1983) and Gevers (1986).

All such classes of ARMA systems, where the column degrees are allowed to vary over all non-negative integers, obviously define a cover for the set of all transfer functions. However, this cover in general contains too many elements. This means that in an identification procedure, where the column degrees are

* Received 13 April 1987; revised 15 March 1988; received in final form 23 May 1988. The original version of this paper was presented at the 8th IFAC Symposium on Identification and System Parameter Estimation which was held in Beijing, People's Republic of China during August 1988. The Published Proceedings of this IFAC meeting may be ordered from: Pergamon Press plc, Headington Hill Hall, Oxford OX3 0BW, England. This paper was recommended for publication in revised form by Associate Editor V. Kucera under the direction of Editor H. Kwakernaak.

† Institute of Econometrics, Operations Research and System Theory, University of Technology Vienna, Argentinierstraße 8, 1040 Vienna, Austria.

‡ Université Catholique de Louvain, Automatique et Analyse des Systèmes, Bâtiment Maxwell, Place du Levant, B-1348 Louvain-la-Neuve, Belgium.

estimated as a first step, the search would have to be performed over too many sets (i.e. too many different model structures). Therefore, from a practical point of view, there is a strong incentive to find a subcover containing fewer sets. Our main contribution in this paper is to produce such subcovers first for all ARMA systems, and next for a case where the maximum lag is bounded. The latter case is of importance because such a bound is often either known to exist, or has to be imposed in algorithms. These results are presented in Section 4.

The problem that a given cover is too rich (i.e. contains unnecessarily many elements) does not arise in the cases of overlapping parametrizations and Echelon forms mentioned above. Since the number of columns in an ARMA system is in general larger than the number of rows, it is quite clear that a cover prescribed in terms of all possible column degrees has in general "many more" elements than a cover prescribed in terms of all possible row degrees.

Section 5 deals with the relations between our monic ARMA parametrization and Echelon canonical forms. In Section 6 the boundary of a certain set of transfer functions is investigated; the characterization of these boundary points turns out to be important to understand the behaviour of identification procedures. The material of Sections 4, 5 and 6 is new to the best of our knowledge.

In Sections 2 and 3, we first establish the notation and recall some basic results concerning monic ARMA systems and Echelon forms.

2. NOTATIONS AND PRELIMINARIES

By an ARMA system we mean a (vector) difference equation

$$a(D)y(t) = b(D)e(t) \qquad (2.1.\text{a})$$

where $y(t)$ is the s-dimensional output, $e(t)$ is the m-dimensional input, D is the backward shift operator on $\mathbb{Z}[D(y(t) \mid t \in \mathbb{Z}) = (y(t)-1 \mid \in \mathbb{Z})]$ or a complex variable and $a(D)$, $b(D)$ are polynomial matrices:

$$a(D) = \sum_{j=0}^{p} A_j D^j, \quad A_j \in \mathbb{R}^{s\times s} \qquad (2.1.\text{b})$$

$$b(D) = \sum_{j=0}^{q} B_j D^j, \quad B_j \in \mathbb{R}^{s\times m}. \qquad (2.1.\text{c})$$

If $A_0 = \mathrm{I}$, the system will be called monic. Throughout this paper we shall denote $u = \max(p, q)$, and we shall make the following assumptions.

Assumption 1. det $a(D) \neq 0$.

Assumption 2. The transfer function

$$k(D) = a^{-1}(D)b(D) \text{ is causal, i.e.}$$

$k(D) = \sum_{0}^{\infty} K_j D^j$ has a convergent power series expansion in a suitable neighbourhood of zero.

We shall not explicitly distinguish between observed and unobserved inputs. In particular, $e(t)$ can be unobserved white noise (this is the proper ARMA case in time series analysis) with $A_0 = B_0 = \mathrm{I}$; $e(t)$ may also be decomposed into unobserved white noise and observed inputs, which includes the ARMAX case.

As we mentioned in the Introduction, ARMA models derived from state space forms are often described using polynomials in the forward shift operator z; $[z(y(t) \mid t \in \mathbb{Z}) = (y(t+1) \mid t \in \mathbb{Z})]$:

$$p(z)y(t) = q(z)e(t) \qquad (2.2.\text{a})$$

where

$$p(z) = \sum_{j=0}^{1} P_j z^j, \quad q(z) = \sum_{j=0}^{r} Q_j z^j. \qquad (2.2.\text{b})$$

A form (2.1) can easily be obtained from (2.2) (and vice versa) as follows. Let $n_1, \ldots, n_s$ be the row degrees of the matrix $[p(z), q(z)]$; then

$$[a(D), b(D)] = \operatorname{diag}\{D^{n_1}, \ldots, D^{n_s}\}[p(z), q(z)] \qquad (2.3)$$

where $D \cdot z = 1$ (for $z \neq 0$, $D \neq 0$). It then follows that

$$k(z^{-1}) = p^{-1}(z)q(z) = \sum_{i=0}^{\infty} K_i z^{-i} \qquad (2.4)$$

is proper. If $p(z)$ and $q(z)$ are left coprime, then the McMillan degree of k is the degree of $\det p(z)$. We shall denote the McMillan degree of k by $\delta[k]$. We shall use δa for the polynomial degree of a polynomial, a polynomial vector or matrix a.

We denote by T_A the set of all ARMA systems (2.1), i.e. the set of all pairs $[a(D), b(D)]$ of polynomial matrices with s and m fixed, $p, q \in \mathbb{Z}$ arbitrary, satisfying Assumptions 1 and 2.

By $\pi: T_A \to V_A$ we denote the mapping defined by $\pi(a, b) = a^{-1} \cdot b$; V_A is the image of T_A under π (i.e. the set of all $s \times m$ rational, causal transfer functions).

Consider a subset $T \subset T_A$ of ARMA systems. Under some additional assumptions, the input–output behaviour of a system is described by its transfer function. It is therefore natural to consider the quotient space $T \mid \pi$ of π restricted to T: $T \mid \pi = \{\pi^{-1}(k) \cap T \mid k \in \pi(T)\}$. Its elements are called the *k-equivalence classes* or the classes of all observationally equivalent ARMA systems corresponding to k (in T). Clearly we want a unique description of the transfer functions by their parameters. A subset $T \subset T_A$ of ARMA systems is called *identifiable* if π restricted to T is injective, i.e. if within T, (a, b) is uniquely determined from $k = a^{-1}b$.

In addition we want T to be finite dimensional in the sense that $\delta(a, b) \leq n$, for some $n \in \mathbb{N}$ holds. Let $\tau \in \mathbb{R}^d$ be a vector of free parameters for $(a, b) \in T$. We will identify τ with $(a, b) \in T$. Let $V \subset V_A$. A mapping $\psi: V \to T \subset T_A$ such that $\pi(\psi(k)) = k$ and $\psi(V) = T$ holds is called an *ARMA parametrization* of V.

The natural topology for parameter spaces like $T \subset \mathbb{R}^d$ is the relative Euclidian topology. We endow V_A with the topology corresponding to the relative topology in the product space $(\mathbb{R}^{s\times m})^{\mathbb{Z}_+}$ of the power series coefficients $(K_i)_{i\in\mathbb{Z}_+}$. As convergence of transfer functions then corresponds to the pointwise convergence of the power series coefficients, we call this topology the pointwise topology T_{pt}. If A is a set in a topological space, its closure is denoted by $\bar{A}$.

Ideally one would like an ARMA parametrization of the set V_A of all $s \times m$ rational, causal transfer functions. As mentioned in the Introduction, there is no continuous parametrization for the whole set V_A (Hazewinkel and Kalman, 1976); in addition, the corresponding parameter space would not be finite dimensional. Therefore, one has to break V_A into bits and to consider suitable families $\{\psi_\alpha: V_\alpha \to T_\alpha \mid \alpha \in \mathrm{I}\}$ of parametrizations such that $\{V_\alpha \mid \alpha \in \mathrm{I}\}$ covers V_A and that every ψ_α is continuous (compare also Hannan and Deistler (1988)).

3. PARAMETRIZATION BY MONIC ARMA SYSTEMS AND ECHELON FORMS

Let $\alpha = (\alpha_1, \ldots, \alpha_{s+m}) \in \mathbb{Z}_+^{s+m}$ denote the specified (i.e. maximal) degrees of the columns $a_1, \ldots, a_s, b_1, \ldots, b_m$ of $[a(D), b(D)]$ and let $a_i(j), b_i(j)$ denote the ith columns of A_j and B_j, respectively. Furthermore let

$$C_\alpha = (a_1(\alpha_1), \ldots, a_s(\alpha_s), b_1(\alpha_{s+1}), \ldots, b_m(\alpha_{s+m})). \quad (3.1)$$

Notice that C_α is the $s \times (s+m)$ matrix formed by the coefficients of degree $\alpha_1, \ldots, \alpha_{s+m}$ in the columns $a_1, \ldots, a_s, b_1, \ldots, b_m$, respectively; we shall often refer to C_α as the "column-end-matrix" of (a, b). If a prescribed degree, say α_j, is larger than the actual degree of the jth column of (a, b), then the corresponding column in C is zero.

Then by T_α we denote the set of all ARMA systems (a, b) of the form (2.1) with the following properties.

(i) The column degrees are specified by α.
(ii) (a, b) is left coprime.
(iii) C_α has rank s.
(iv) $A_0 = \mathrm{I}$.

As is easily seen, T_α is identifiable, since left-multiplication of (a, b) by a non-constant unimodular matrix due to (iii) would increase at least one column degree. Let $V_\alpha = \pi(T_\alpha)$. A vector $\tau \in \mathbb{R}^{d_\alpha}$ of free parameters for T_α is defined as the vector consisting of the stacked rows of $(a_1(1), \ldots, a_1(\alpha_1), \ldots, a_s(1), \ldots, a_s(\alpha_s), b_1(0), \ldots, b_1(\alpha_{s+1}), \ldots, b_m(0), \ldots, \pi b_m(\alpha_{s+m}))$; here $d_\alpha = s(\alpha_1 + \cdots + \alpha_{s+m}) + s\cdot m$. By $\psi_\alpha: V_\alpha \to T_\alpha$ we denote the corresponding ARMA parametrization.

By $\alpha \leq \beta$ we mean $\alpha_1 \leq \beta_1, \ldots, \alpha_{s+m} \leq \beta_{s+m}$ and $\alpha < \beta$ will mean that at least one inequality holds. In addition we shall also use the notation

$$|\alpha| = \sum_{i=1}^{s+m} \alpha_i.$$

Some properties of the monic ARMA parametrizations with prescribed column degrees have been derived by Hannan (1971), Deistler *et al.* (1978), Deistler (1983) and Deistler (1985). We summarize the main properties which will be used later.

Proposition 3.1.

(1) T_α is an open and dense subset of $\mathbb{R}^{d_\alpha}$, where
$$d_\alpha = s\left(\sum_1^{s+m} \alpha_i\right) + s \cdot m.$$
(2) $\psi_\alpha: V_\alpha \to T_\alpha$ is a (T_{pt}) homeomorphism.
(3) $\pi(\bar{T}_\alpha) = \bigcup_{\beta\leq\alpha} V_\alpha$.
(4) V_α is $(T_{\text{pt}}$-) open in $\bar{V}_\alpha$.
(5) $\pi(\bar{T}_\alpha) \subset \bar{V}_\alpha$ and equality holds for $s = 1$.

By W_n, $n = (n_1, \ldots, n_s)$, we denote the set of all transfer functions $k \in V_A$ with Kronecker indices $(n_1, \ldots, n_s)$. Thus for $k \in W_n$ we have $k(z^{-1}) = p^{-1}(z)$, $q(z)$ where (p, q) is in Echelon form and thus satisfies the following (Guidorzi, 1981; Deistler, 1985; Gevers, 1986).

(i) (p, q) is left coprime.
(ii) p_{ii} are monic polynomials (in the original meaning of the word, namely, that the coefficient corresponding to the highest power equal to one) and

$$\begin{aligned} &\delta p_{ij} \leq \delta p_{ii} = n_i \quad && j < i \\ &\delta p_{ij} < \delta p_{ii} && j > i \\ &\delta p_{ji} < \delta p_{ii} && j \neq i \\ &\delta q_{ij} \leq \delta p_{ii}. \end{aligned}$$

The reversed Echelon (or Guidorzi) form is then obtained as

$$[a(D), b(D)] = \text{diag}\{D^{n_i}\}[p(z), q(z)].$$

The following proposition which combines results of Janssen (1987) and Kailath (1980, pp. 481–482), will play a key role in some of our proofs.

Proposition 3.2.
Let $[a(D), b(D)]$ be a left coprime ARMA system and $k = a^{-1}b$. Denote $c(D) = [a(D), b(D)]$. Then

(1) the McMillan degree $\delta[k]$ is the highest degree of all $s \times s$ minors of $c(D)$;

(2) $c(D)$ is row proper if and only if its row degrees are equal to the Kronecker indices of k, up to reordering;

(3) by left multiplication with a suitable unimodular matrix, $c(D)$ can be brought in a form satisfying:

(i) $c(D)$ is row proper with row degrees $k_1, \ldots, k_s$ arranged in ascending order;

(ii) for row j, $j = 1, \ldots, s$, there exists an index p_j such that c_{j,p_j} has degree k_j and is monic (i.e. the (j, p_j)th element of the matrix $c(D)$ is monic of degree k_j);

(iii) $\delta c_{ji} < k_j$ for $i > p_j$;

$\delta c_{ip_j} < k_j$ for $i \neq j$;

if $k_i = k_j$ and $i < j$, then $p_i < p_j$.

An ARMA system satisfying conditions (3) of Proposition 3.2 is said to be in Popov form. Note that the $s \times s$ submatrix corresponding to the indices $p_1, \ldots, p_s$ is both row and column proper, and its row degrees are equal to the column degrees and to the Kronecker indices, up to permutation. If we premultiply a system in Popov form by any non-singular constant matrix (and in particular by $a^{-1}(0)$), the column properness of the above mentioned submatrix is preserved.

4. A NON-REDUNDANT COVER FOR V_A

As can easily be shown, $\mathcal{G} = \{V_\alpha \mid \alpha \in \mathbb{Z}_+^{s+m}\}$ is a cover for V_A. As is immediately clear for the scalar case ($s = m = 1$), there are several proper subcovers of $\mathcal{G}$, e.g. $\mathcal{S} = \{V_{(n,n)} \mid n \in \mathbb{Z}_+\}$. One desirable property of a cover $\{V_\alpha \mid \alpha \in \mathrm{I}\}$ is its *nonredundancy* in the sense that no element V_α of $\{V_\alpha \mid \alpha \in \mathrm{I}\}$ can be eliminated without destroying the covering property, or equivalently that for every $\alpha \in \mathrm{I}$ there is a $k \in V_\alpha$ which is not contained in any of the V_β, $\beta \in \mathrm{I}$, $\beta \neq \alpha$. This is to postulate the minimality of a cover with respect to inclusion of sets.

Another desirable property of a cover is that the dimension d of the parameter space used to describe the transfer function should be small. This is clearly related to efficiency of estimators. A necessary condition that a given k is described in a parameter space T_α of minimal dimension is that α is a *minimal index* in the sense that there is no $\beta < \alpha$, $\beta \in \mathbb{Z}_+^{s+m}$, such that $k \in V_\beta$ holds. We have the following Lemma.

Lemma 4.1. An index β is minimal for $k \in V_A$ if and only if ($k \in V_\beta$ and) C_β has no zero column.

Proof. One direction is evident. Conversely, let C_β have no zero column. To preserve coprimeness, only left multiplications by unimodular transformations $u(D)$ are allowed. Since a is required to be monic, we must have $u(0) = \mathrm{I}$. Since C_β is full rank, multiplication of $[a(D), b(D)]$ by a non-constant $u(D)$ would increase at least one column degree.

As is easily seen, for $s = 1$ and for a given $k \in V_A$, the minimal index α is uniquely defined and it corresponds to the minimum of $|\alpha|$ (and thus to the minimal dimension of T_α). Neither of these statements are true for $s > 1$ as the following example indicates.

Example 4.1. Consider the case $s = m = 2$, with

$$[a(D), b(D)] = \begin{pmatrix} 1 - D^3 & D - D^2 - D^3 & 1 + D & 0 \\ D & 1 + D & 0 & 1 \end{pmatrix}.$$

Clearly $k(D) = a^{-1}(D)b(D) \in V_{(3\ 3\ 1\ 0)}$ and $\alpha = (3\ \ 3\ \ 1\ \ 0)$ is a minimal index. An alternative description for $k(D)$, however, is $k(D) = \tilde{a}^{-1}(D)\tilde{b}(D)$ with

$$[\tilde{a}(D), \tilde{b}(D)] = \begin{pmatrix} 1 & D & 1 + D & D^2 \\ D & 1 + D & 0 & 1 \end{pmatrix}.$$

Here $k \in V_{(1\ 1\ 1\ 2)}$ where $\beta = (1\ \ 1\ \ 1\ \ 2)$ is again a minimal index. Note that $|\alpha| = 7$ while $|\beta| = 5$.

Nonredundancy in general conflicts with the desirable property of describing transfer functions with a small number of parameters, as is illustrated for the case $s = m = 1$ by the following example.

Example 4.2. Recall that $\mathcal{S} = \{V_{(n,n)} \mid n \in \mathbb{Z}_+\}$ is a proper subcover of $\mathcal{G}$ for the scalar case, and even with disjoint elements. Consider the transfer function $k(D) = (1 + a_1D + a_2D^2)^{-1}(1 + b_1D) \in V_{(2,1)}$ where $(1 + a_1D + a_2D^2)$ and $(1 + b_1D)$ have no common zero and a_2, b_1 are both nonzero. If we choose the cover $\mathcal{S}$, then $k \in V_{(2,2)}$ and its description requires one more real-valued parameter. On the other hand, the cover $\mathcal{G}$ contains more elements than $\mathcal{S}$, but we can always describe k by a minimal index. Note also that for this particular example, $k \in V_{(2,j)}$, $j \geq 1$ and $k \in V_{(j,1)}$, $j \geq 2$.

Note that in the case $s > 1$, $\mathcal{S}_1 = \{V_{(n,n,\ldots,n)} \mid n \in \mathbb{Z}_+\}$ is not a cover for V_A. This makes the problem of finding a non-redundant subcover of $\mathcal{G}$ nontrivial. We now focus on this problem.

Theorem 4.1. Let

$$\mathrm{I}(u) = \{\alpha \in \mathbb{Z}_+^{s+m} \mid \alpha_i \leq u; \quad i = 1, \ldots, s \text{ where at least one equality holds, and } \alpha_i = u; \quad i = s+1, \ldots, m+s\}.$$

Then $\mathcal{S} = \{V_\alpha \mid \alpha \in \mathrm{I}(u), \quad u \in \mathbb{Z}_+\}$ is a non-redundant cover for V_A.

Proof. First, let us show the covering property: let $k \in V_A$ be an arbitrary transfer function and let $k(D) = \tilde{a}(D)^{-1}\tilde{b}(D)$ where $[\tilde{a}(D), \tilde{b}(D)]$ is in Hermite's form (see e.g. Barnett (1971)) and thus $(\tilde{a}, \tilde{b})$ is left coprime and $\tilde{a}$ is upper triangular and column proper. Since $k(D)$ is causal, $\tilde{a}(0)$ is nonsingular and thus $(a, b) = \tilde{a}^{-1}(0)(\tilde{a}, \tilde{b})$ is monic, left coprime and has the same actual column degrees and the same linear dependence relations between the columns in the column-end-matrix as $(\tilde{a}, \tilde{b})$; in particular also a is column proper. Let n_i denote the degree of the ith column of (a, b) and let $u = \max n_i$; $i = 1, \ldots, s+m$. Then we define α as follows. If there is an index j, $1 \leq j \leq s$ such that $n_j = u$, then $\alpha = (\alpha_1, \ldots, \alpha_{s+m})$ is defined by $\alpha_i = n_i$, $i = 1, \ldots, s$; $\alpha_i = u$, $i = s+1, \ldots, s+m$. If this is not the case we proceed as follows. Let j, $s+1 \leq j \leq s+m$ be such that $n_j = u$ and thus $b_{j-s}(u)$ (i.e. the highest degree coefficient of b_{j-s}) has at least one non-zero element, in position 1 say. Then we define $\alpha_i = n_i$, $i \neq 1$, $i = 1, \ldots, s$; $\alpha_i = u$ for $i = 1$ and $i = s+1, \ldots, s+m$. Clearly then $\alpha \in \mathrm{I}(u)$, and by construction, since a is upper triangular and column proper, C_α has rank s. Thus $k \in V_\alpha$ and thus $\mathcal{S}$ is a cover for V_A.

In order to prove nonredundancy, we show that for each $V_\alpha \in \mathcal{S}$ there exists a $k \in V_\alpha$ which does not belong to any other $V_\beta \in \mathcal{S}$. Let

$$[a(D), b(D)] = \begin{pmatrix} 1+D^{\alpha_1} & & 0 & 1 \ldots\ldots 1 \\ & 1+2\cdot D^{\alpha_2} & & \vdots \qquad \vdots \\ 0 & & 1+s\cdot D^{\alpha_s} & 1 \ldots\ldots 1 \end{pmatrix}.$$

Clearly (a, b) is left coprime, a is monic, for $\alpha = (\alpha_1, \ldots, \alpha_s, u, \ldots, u)$, $u = \max \alpha_i$, $i = 1, \ldots, s$, C_α has rank s and $\alpha \in \mathrm{I}(u)$. Let v be any unimodular matrix with $v(0) = \mathrm{I}$ and let $l_1, \ldots, l_s$ denote its column degrees. We consider $(\hat{a}, \hat{b}) = v(a, b)$; clearly then the column degrees of $\hat{a}$ are $\alpha_1 + l_1, \ldots, \alpha_s + l_s$ and the column-end-matrix of $\hat{a}$, corresponding to $(\alpha_1 + l_1, \ldots, \alpha_s + l_s)$, shows the identical linear dependence relations as the column-end-matrix of v, corresponding to $(l_1, \ldots, l_s)$, and thus is singular, unless $l_1 = \cdots = l_s = 0$. Since the latter case means $v = \mathrm{I}$, without restriction of generality we assume $l_j > 0$ for at least one j. The degrees of the columns of b are smaller than or equal to $l = \max l_j$, $j = 1, \ldots, s$. If this maximum is not attained or if $l < \max(\alpha_i + l_i)$, $i = 1, \ldots, s$ then for every $\beta \in \mathrm{I}(n)$, $n \in \mathbb{Z}_+$, such that $(\hat{a}, \hat{b}) \in \bar{T}_\beta$, C_β has rank smaller than s. If at least one of the columns of $\hat{b}$ has degree l and $l = \max(\alpha_i + l_i)$, $i = 1, \ldots, s$, then the coefficients of these columns corresponding to power l are linear combinations of the columns of the column-end-matrix of v and thus again C_β has rank smaller than s for every $\beta \in \mathrm{I}(n)$, $n \in \mathbb{Z}_+$, such that $(\hat{a}, \hat{b}) \in \bar{T}_\beta$. Thus for such a $(a, b) \in T_\alpha$ every unimodular left transformation with $v(0) = \mathrm{I}$ such that $v(a, b) \in T_\beta$, $\beta \in \mathrm{I}(n)$, for some n must be the identity matrix. As is clear from the construction of (a, b), we cannot have $(a, b) \in T_\beta$, $\beta \in \mathrm{I}(u)$, $u < \max \alpha_i$, $i = 1, \ldots, s$. If $u = \max \alpha_i$, $i = 1, \ldots, s$ and $\beta \in \mathrm{I}(u)$, then $(a, b) \in T_\beta$ implies $\beta = \alpha$ and if $u > \max \alpha_i$ then C_β has rank less than s for all $\beta \in \mathrm{I}(u)$ such that $(a, b) \in T_\beta$. Thus we have shown nonredundancy.

Comment 4.1. Clearly for every $k \in V_\alpha$, $\alpha \in \mathrm{I}(u)$, there are infinitely many $\beta \in \mathbb{Z}_+^{s+m}$, $\beta \geq \alpha$ such that $k \in V_\beta$ (compare Example 4.2), but none of these β are contained in a $\mathrm{I}(u)$ for some $u \in \mathbb{Z}_+$.

In many situations, either there is *a priori* information on an upper bound for the maximum lag, or if we want to estimate α, estimation procedures like AIC or BIC demand the prescription of such an upper bound (see Hannan and Kavalieris (1984)). If n is such an upper bound, the next example shows that $\{V_\alpha \mid \alpha \in \mathrm{I}(u),\ u \leq n\}$ is not a cover for $\bigcup_{\alpha_j \leq n} V_\alpha$.

Example 4.3. Consider $s = 2$, $m = 1$ and $k = a^{-1}b$

$$a(D) = \begin{pmatrix} 1+D & D \\ 0 & 1 \end{pmatrix}, \quad b(D) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Clearly $k \in V_{(1\ 1\ 0)}$, but it can easily be checked that k is not contained in any of the V_α, $\alpha = (0\ \ 0\ \ 0), (0\ \ 1\ \ 1), (1\ \ 0\ \ 1), (1\ \ 1\ \ 1)$.

This demonstrates that in the case where an upper bound for the maximum lag is prescribed, we have to search for another cover.

Theorem 4.2. Let

$$J(u)=\{\alpha\in\mathbb{Z}_+^{s+m}\mid \alpha_i\le u,\quad i=1,\dots,s+m,$$
$$\text{where at least } m+1 \text{ equalities hold}\}.$$

Then

$$\mathcal{T}_n=\{V_\alpha \mid \alpha\in J(u),\quad u\le n\}$$

has the following property

$$\bigcup_{V_\alpha\in\mathcal{T}_n} V_\alpha=\bigcup_{\alpha\le(n,\dots,n)} V_\alpha.$$

Proof. One inclusion is obvious. It remains to show that every $k\in V_\alpha$, $\alpha_i\le n$, $i=1,\dots,s+m$ is contained in a suitable V_β, $\beta\in J(u)$, $u\le n$. Consider such a k and the corresponding $(a,b)\in T_\alpha$. First we transform (a,b) to a row proper $(\tilde a,\tilde b)$ by a left unimodular transformation. It follows from Theorem 6.3–13 p. 387 of Kailath (1980) applied to the inverse transformation that the row degrees of $(\tilde a,\tilde b)$ are all less than or equal to n. Taking in particular the Popov form (see Proposition 3.2) $\beta=(\beta_1,\dots,\beta_{s+m})$ can be specified as follows: let $\beta_{p_j}=k_j$, $j=1,\dots,s$ and let $\beta_j=\max_{i=1,\dots,s} k_i$ for the other indices. Then C_β has full rank and thus $k\in V_\beta$.

Although the cover $\mathcal{T}_n$ is not nonredundant, it still leads to a considerable reduction of the number of subsets when compared with $\{V_\alpha \mid \alpha_i\le n\}$. This has important practical applications in identification. We illustrate with the following example.

Example 4.4. Let $s=m=2$ and assume that $n=3$ is an upper bound for the highest lag. Then if we try to identify k using the cover $\mathcal{T}_3$, we have to search through 28 subsets, whereas $\{V_\alpha \mid \alpha_i\le 3,\quad i=1,\dots,4\}$ contains 256 subsets.

5. CONNECTION BETWEEN KRONECKER INDICES AND THE COLUMN DEGREES OF MONIC ARMA MODELS

In this section we describe a non-redundant cover $\{V_\alpha \mid \alpha\in I_n\}$ for the set of all systems with prescribed (but arbitrary) Kronecker indices $n=(n_1,\dots,n_s)$. We shall assume throughout that all Kronecker indices are nonzero.

First let us consider the case where all indices are equal. We denote by $W(u)$ the set W_n for $n=(u,u,\dots,u)$ and $V(u)$ the set V_α for $\alpha=(u,\dots,u)$.

Lemma 5.1. (The Mountain Lemma.*)

$$V(u)=W(u).$$

Proof. (1) Let $k\in V(u)$ with $k(D)=a^{-1}(D)\cdot b(D)$; $(a,b)\in T_{(u,\dots,u)}$. Then (a,b) are left coprime, $A_0=I$, the prescribed row degrees of (a,b) are $(u,\dots,u)$ and $[A_u,B_u]$ has rank s. Therefore $[a(D),b(D)]$ is in reversed Echelon form and $k\in W(u)$.

(2) Conversely let $k\in W(u)$ and let $k(z)=p^{-1}(z)\cdot q(z)$ be in Echelon form for k. Define $[a(D),b(D)]=D^u[p(z),q(z)]$. Then $[p(z),q(z)]$ is coprime and $[A_0,B_0]=[P_u,Q_u]$, where P_u is nonsingular in the Echelon form. Also, $[A_u,B_u]=[p(0),q(0)]$ has full rank by coprimeness of (p,q). Therefore $k\in V(u)$.

Now we consider arbitrary Kronecker indices. The main result of this section is as follows.

Theorem 5.1. (The Melk Theorem.†) Consider W_n with $n=(n_1,\dots,n_s)$ arbitrary. Then

(i) $W_n\subset\bigcup_{\alpha\in I_n} V_\alpha$ where $I_n=\{\alpha=(\alpha_1,\dots,\alpha_{s+m})\mid \alpha$ contains $n_1,\dots,n_s$; the remaining m elements of α are equal to $\max n_i\}$.

(ii) If $s\le m$ the cover $\{V_\alpha \mid \alpha\in I_n\}$ is nonredundant. $\{V_\alpha \mid \alpha\in I_n\}$ is even a cover for $\bigcup_{n\in P_n} W_n$, where P_n is the index set corresponding to all permutations of $n_1,\dots,n_s$.

Proof. Part (i). Consider an arbitrary $k\in W_n$. Then we can find a corresponding $[a(D),b(D)]$ that satisfies the conditions of Proposition 3.2, i.e. $[a(D),b(D)]$ is in Popov form. In particular, there is a subset of s columns of index $p_1,\dots,p_s$, say whose column degrees are equal to $n_1,\dots,n_s$, and which form a column proper matrix. Let $u=\max n_i$; then the other column degrees are smaller than or equal to u. Denote $[\tilde a(D),\tilde b(D)]=a^{-1}(0)[a(D),b(D)]$. Since the column degrees of $[\tilde a(D),\tilde b(D)]$ are the same as those of $[a(D),b(D)]$, $[\tilde a(D),\tilde b(D)]\in T_\alpha$ for $\alpha\in I_n$.

To prove that the cover is nonredundant for $s\le m$, we show that for each $\beta\in I_n$ there is at least one $k\in W_n$ which belongs to this V_β and to no other V_α in $\{V_\alpha \mid \alpha\in I_n\}$. Let $\beta\in I_n$ and assume that l of the first s indices and q of the last m are strictly smaller than $u=\max\beta_i$. Note that $l+q\le s-1$. Now construct $[a(D),b(D)]$ as follows.

Step 1. For each $\beta_i<u$, $i=1,\dots,s$, set the ith element of the corresponding column of $a(D)$ equal to $1+D^{\beta_i}$ and set all other elements of that column equal to zero.

Step 2. In the first $s-q-l$ remaining columns of $a(D)$ (corresponding to $\beta_i=u$) set the

* So called because the result was obtained when both authors were on top of an Austrian mountain.

† So called because it was discovered on the way to the famous Melk monastery.

diagonal element equal to $1+D^u$ and the other elements to zero. Hence $a(D)$ will have $1+D^u$ in at least one row.

Step 3. In the remaining q columns of $a(D)$, set the diagonal element equal to 1, the jth element equal to D^u and the other elements equal to zero.

Step 4. For each of the q indices $\beta_i < u$, $i = s+1, \ldots, s+m$, set the corresponding column of $b(D)$ equal to $[0, \ldots, 0, D^{\beta_i}, 0, \ldots, 0]'$, where the elements D^{β_i} are in the rows in which $a(D)$ has 1 on the diagonal and in such a way that only one such term D^{β_i} appears in any given row.

Step 5. In the remaining $m-q$ columns of $b(D)$ set the jth element equal to D^u. In $s-q-1$ of these columns set one element equal to 1 in such a way that each row of $[a(D), b(D)]$, except the jth row, contains one element equal to 1. Set all remaining elements in these columns equal to zero.

Because this construction is rather complicated, we illustrate with an example. Let $s=m=4$, and $\beta=(3, 1, 3, 3, 3, 3, 2, 3)$; then $u=3$, $l=1$, $q=1$, $s-l-q=2$, $m-q=3$. Therefore

$$[a(D), b(D)] = \begin{pmatrix} 1+D^3 & 0 & 0 & D^3 & D^3 & D^3 & 0 & D^3 \\ 0 & 1+D & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1+D^3 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & D^2 & 0 \end{pmatrix}.$$

With the above construction it is easy to see that (a, b) are left coprime with full rank C_β and $k(D) = a^{-1}(D) \cdot b(D) \in W_n$, where n contains the $l+q$ indices $\beta_i < u$, while the remaining n_j are equal to u. Now suppose $k \in V_\alpha$, $\alpha \in \mathrm{I}_n$, $\alpha \neq \beta$. Then there must exist a unimodular transformation $u(D)$ that decreases at least one column degree of (a, b) while keeping the maximal degree equal to u, since $\alpha \in \mathrm{I}_n$. The only columns whose degrees can be decreased are the columns that contain the elements 1 and D^u. Since, by construction, each row that contains an element 1 also contains a term D^{β_i}, $\beta_i \geq 1$, any unimodular transformation that decreases the column degrees would produce another column degree strictly larger than u elsewhere. Therefore α would not be in I_n.

Part (ii). The covering property for $\bigcup_{n \in P_n} W_n$ follows immediately from part (i).

Comment 5.1. The non-redundancy property of the cover probably also holds when $s > m$, but another construction is needed for (a, b) in order to guarantee coprimeness.

Comment 5.2. Theorem 5.1 shows that in order to describe the set W_n by ARMA Popov forms (see Proposition 3.2) for $s \leq m$ every possible square matrix of $[a(D), b(D)]$ has to be considered as a row proper candidate in the transformation used in that Proposition.

Theorem 5.1 gives a cover in terms of monic ARMA models for the sets W_n of systems with prescribed Kronecker indices. Conversely given $k \in V_\alpha$ one would like to know what the possible Kronecker indices or at least the possible McMillan degrees of k are. The following result gives a partial answer in the form of lower and upper bounds on $\delta[k]$, $k \in V_\alpha$ as a function of the column-end-matrix C_α.

Theorem 5.2. Let $(a, b) \in T_\alpha$, α arbitrary, with column-end-matrix C_α and let $u = \max \alpha_i$. Let γ be the largest possible sum of column degrees of any s columns of (a, b), let σ be the largest possible sum of column degrees of s columns of (a, b) where the corresponding columns of C_α are linearly independent, and let ρ be the rank of $[A_u \quad B_u]$. Then, with $k = a^{-1}b$, we have

$$\sigma \leq \delta[k] \leq \max\{(u-1)s + \rho, \gamma\}.$$

Proof. This is an easy consequence of Corollary 3.2, part (b) of Janssen (1987) [see Proposition 3.2(1)].

6. THE BOUNDARY OF $V(u)$

As has been mentioned already, for $s > 1$, $\{V(u) \mid u \in \mathbb{Z}^+\}$ is not a cover for V_A. However, from Proposition 3.1 we see that $\{\bar{V}(u) \mid u \in \mathbb{Z}_+\}$ is a cover for V_A and, moreover, from the results in Hannan and Kavalieris (1984) we see that the smallest u such that $k \in \bar{V}(u)$ holds, can be consistently estimated by an identification procedure, e.g. obtained from minimizing BIC. Let $M(j) \subset V_A$ denote the manifold of all transfer functions of order j. As is well known, for $s > 1$, $M(j)$ in general cannot be described by one coordinate system and the relation $\overline{M(j)} = \bigcup_{i \le j} M(i)$ holds (see, e.g. Hazewinkel and Kalman (1976) and Deistler and Hannan (1981)). Thus it is clear that also $\bar{V}(u) = \overline{M(s \cdot u)}$ cannot be described by one coordinate system and therefore the question of a suitable cover for $\bar{V}(u) - V(u)$ arises. $\bar{V}(u)$ may be partitioned into three parts (see, e.g. Deistler (1983)).

(i) $V(u)$ which is the set of all transfer functions which are parametrized by $T(u) = T_{(u,\dots,u)}$.

(ii) $\pi(\bar{T}(u)) - V(u) = \pi(\bar{T}(u) - T(u))$
$$= \bigcup_{\beta \le (u,\dots,u)} V_\beta - V(u);$$
i.e. the set of all transfer functions which can be described by equivalence classes on the boundary of $T(u)$ but not within $T(u)$ (see Proposition 3.1).

(iii) $\bar{V}(u) - \pi(\bar{T}(u))$; this set is void for $s = 1$; it corresponds to the point of infinity in the parameter space $T(u)$ (see Proposition 3.1) for $s > 1$.

The investigation of the boundary points in (ii) and (iii) is particularly important to understand the properties of identification procedures when a parametrization of $V(u)$ is considered and when the true transfer function is in $\pi(\bar{T}(u) - T(u))$ or in $\bar{V}(u) - \pi(\bar{T}(u))$. This has been discussed in Deistler *et al.* (1978), Deistler and Hannan (1981) and Deistler (1983).

In the next Theorem we give a characterization of the sets $\pi(\bar{T}(u) - T(u))$ and $\bar{V}(u) - \pi(\bar{T}(u))$ in terms of Kronecker indices.

Theorem 6.1.

(1) $$\bigcup_{\beta \le (u,\dots,u)} V_\beta - V(u) = \bigcup_{n < (u,\dots,u)} W_n.$$

(2) $$\bar{V}(u) - \pi(\bar{T}(u)) = \bigcup_{\substack{|n| \le s \cdot u \\ n \nleq (u,\dots,u)}} W_n.$$

Proof. Let $k \in W_n$, $n < (u, \dots, u)$; then all column degrees of the corresponding reversed Echelon form $(\hat{a}, \hat{b})$ are smaller than or equal to u and clearly the same holds for $(a, b) = \hat{a}(0)^{-1}(\hat{a}, \hat{b})$ (where a is monic). Now since (a, b) is left coprime, the rank of (A_u, B_u) must be smaller than s, since otherwise $k \in V(u) = W(u)$ and thus, by Proposition 3.1, $k \in V_\beta$, $\beta < (u, \dots, u)$. Conversely, let $(a, b) \in \bar{T}(u) - T(u)$ and write
$$[p(z), q(z)] = z^u[a(D), b(D)].$$
Clearly the degree of (p, q) does not exceed u. By Theorem 1 in Deistler *et al.* (1978), (p, q) can be made left coprime without increasing its degree. As can easily be shown (p, q) can also be made row proper by left multiplication with unimodular matrices without increasing its degree. Since for any left coprime and row proper (p, q), the row degrees are given up to permutation by Kronecker indices and since $(a, b) \notin V_u = W_u$, part (1) has been proved. Part (2) follows from part (1), since $\bar{V}_u = \bar{W}_u$ holds.

Comment 6.1. By the theorem above, the class of transfer functions corresponding to non-trivial equivalence classes (i.e. which are not singletons) on the boundary of $V_u = W_u$ is the same for the two parametrizations considered. Note that $\bigcup_{\beta < (u,\dots,u)} V_\beta \supset \bigcup_{\beta \le (u,\dots,u)} V_\beta - V_u$ holds, where in general the inclusion is proper.

7. CONCLUSION

We have presented a number of new properties of monic ARMA models with prescribed column degrees. We believe that the main contributions of our paper are as follows.

First we have given a non-redundant cover for the set of all transfer functions in terms of these monic ARMA models; in addition we also propose a cover for the case where the maximum lag has been prescribed. Our results here allow for a very significant reduction in the number of coordinates for which a search must be performed during an identification procedure.

Secondly, we have examined some connections between the prescribed column degrees of these monic ARMA models and the Kronecker indices of the corresponding transfer functions. We have given a cover in terms of monic ARMA models with prescribed column degrees for the set of all transfer functions with given Kronecker indices. Conversely, we have given upper and lower bounds for the McMillan degree of a monic ARMA system with prescribed column degrees in terms of these column degrees and of the corresponding column-end-matrix. These connections between column degrees and Kronecker indices and/or McMillan degrees are rather complicated. Moreover, it is likely that an exact calculation of the McMillan degree can only be based on a complete knowledge of the coefficient matrices and not just on the column degrees or on the column-end-matrix.

Finally, we have characterized the boundary of the set of monic ARMA systems of prescribed column degrees, but only in the simple case where all these column degrees are identical. The characterization for the points at infinity of the parameter space is in terms of Kronecker indices only. A characterization in terms of monic ARMA systems, and for sets of arbitrary column degrees, is still not available. It appears to be a very hard problem, but a practically important one if one wants to understand the properties of the identification procedure.

Acknowledgements—This research was supported by the Austrian Science Foundation (Fonds zur Förderung der wissenschaftlichen Forschung), project No. S 32/02.

REFERENCES

Barnett, S. (1971). *Matrices in Control Theory.* Van Nostrand Reinhold, London.

Deistler, M. (1983). The properties of the parametrization of ARMAX systems and their relevance for structural estimation. *Econometrica,* **51,** 1187–1207.

Deistler, M. (1985). General structure and parametrization of ARMA and state space systems and its relation to statistical problems. In E. J. Hannan, P. R. Krishnaiah and M. M. Rao (Eds), *Handbook of Statistics,* Vol. 5, pp. 257–277. North-Holland, Amsterdam.

Deistler, M., W. Dunsmuir and E. J. Hannan (1978). Vector linear time series models; corrections and extensions. *Adv. Appl. Probab.,* **10,** 360–372.

Deistler, M. and E. J. Hannan (1981). Some properties of the parametrization of ARMA systems with unknown order. *J. Multivariate Anal.,* **11,** 474–484.

Gevers, M. (1986). ARMA models, their Kronecker indices and their McMillan degree. *Int. J. Control,* **43,** 1745–1761.

Gevers, M. and V. Wertz (1987). Techniques for the selection of identifiable parametrizations for multivariable linear systems. In C. T. Leondes (Ed.), *Control and Dynamic Systems,* Vol. XXIV, pp. 35–86. Academic Press, Orlando, FL.

Guidorzi, R. (1981). Invariants and canonical forms for systems of structural and parametric identification. *Automatica,* **17,** 117–133.

Hannan, E. J. (1971). The identification problem for multiple equation systems with moving average errors. *Econometrica,* **39,** 751–756.

Hannan, E. J. and M. Deistler (1988). *The Statistical Theory of Linear Systems.* Wiley, New York.

Hannan, E. J. and L. Kavalieris (1984). Multivariate linear time series models. *Adv. Appl. Probab.,* **16,** 492–561.

Hazewinkel, M. and R. E. Kalman (1976). On invariance canonical forms and moduli for linear constant finite dimensional dynamic systems. In *Lecture Notes Econ.-Math. System Theory,* Vol. 131, pp. 48–60. Springer, Berlin.

Janssen, P. (1987). General results on the McMillan degree and the Kronecker indices of ARMA- and MFD-models. *Int. J. Control.*

Kailath, T. (1980). *Linear Systems.* Prentice-Hall, Englewood Cliffs, NJ.

Kalman, R. E. (1974). Algebraic geometric description of the class of linear systems of constant dimension. 8th Annual Princeton Conference on Information Sciences and Systems. Princeton, NJ.