The page matrix : an excellent tool for noise filtering of Markov parameters, order testing and realization

Eindhoven
University of Technology
the Netherlands

Department of
Electrical Engineering

The Page Matrix: An excellent Tool for Noise Filtering of Markov Parameters, Order Testing and Realization.

By
A.A.H. Damen, P.M.J. Van den Hof and
A.K. Hajdasiński

EUT Report 82-E-127
ISBN 90-6144-127-7
ISSN 0167-9708
August 1982

Eindhoven University of Technology Research Reports

EINDHOVEN UNIVERSITY OF TECHNOLOGY

Department of Electrical Engineering

Eindhoven The Netherlands

THE PAGE MATRIX: An excellent tool for noise filtering of Markov parameters, order testing and realization.

By

A.A.H. Damen

P.M.J. Van den Hof

A.K. Hajdasiński

EUT Report 82-E-127

ISBN 90-6144-127-7

ISSN 0167-9708

Eindhoven

August 1982

CIP-gegevens

Damen, A.A.H.

The Page matrix: an excellent tool for noise filtering
of Markov parameters, order testing and realization /
A.A.H. Damen, P.M.J. Van den Hof, A.K. Hajdasiński. -
Eindhoven: University of technology. -
(Eindhoven university of technology research reports;
82-E-127)
Met lit. opg., reg.
ISBN 90-6144-127-7
ISSN 0167-9708
SISO 656.2   UDC 519.71
Trefw.: meettechniek / regeltechniek.

CONTENTS

A note on this report

Chapter one of this report presents the most relevant information, but mainly as a result of one crucial theorem. The proof of this theorem is quite extensive and given in chapter two as a separate part.

ABSTRACT

When the Ho-Kalman algorithm is applied to a truncated series of noisy Markov parameters, the Hankel matrix used has either the improper rank or it lacks the Hankel structure. Furthermore the Markov parameters are not processed with a constant weighting factor, which implies that the noise filtering is inadequate.
In this paper we propose to use an alternative matrix: the Page matrix. It is shown that this method is much better suited for handling the noisy Markov parameters . This holds with respect to three aspects: order testing, noise filtering and realization. Even in the deterministic case, the Page matrix offers the advantage of a considerable reduction in computation.

KEYWORDS: multivariable systems, stochastic systems, identification, parameter estimation, system order reduction, noise filtering, realization, Hankel matrix.

Damen, A.A.H., P.M.J. Van den Hof and A.K. Hajdasiński
THE PAGE MATRIX: An excellent tool for noise filtering of Markov parameters, order testing and realization.
Department of Electrical Engineering, Eindhoven University of Technology, 1982.
EUT Report 82-E-127

Address of the authors:

Group Measurement and Control,
Department of Electrical Engineering,
Eindhoven University of Technology,
P.O. Box 513,
5600 MB EINDHOVEN,
The Netherlands

A.K. Hajdasiński is on the same address, but employed by "Samenwerkingsorgaan KHT-THE" (Catholic University Tilburg - Eindhoven University of Technology).

Chapter 1. OUTLINE OF THE PROBLEMS AND INTRODUCTION OF THE PAGE MATRIX

1.0 INTRODUCTION

The minimum realization for a sufficiently long truncated series of deterministic Markov parameters offers no problems. This has been shown by Ho and Kalman [1]. In section 1.1 we will summarize this algorithm. However, a fully satisfactory solution for the noisy case has not yet been proposed. Of course, the Ho-Kalman algorithm is being applied in a modified way for the noisy case, but the results are rather questionable. This is quite understandable because, theoretically, the processing of the noise is fundamentally wrong. This inadequacy will be elucidated in section 1.2.
In our opinion, the use of the Page matrix, which will be introduced in section 1.3, may overcome the majority of the problems caused by the use of the Hankel matrix. This will be emphasized in section 1.4, where we will compare the Hankel matrix algorithm with the one based on the Page matrix.

1.1. THE HO-KALMAN ALGORITHM FOR DETERMINISTIC DATA

As a short recapitulation and in order to define a notation, we will briefly sketch the Ho-Kalman algorithm. This algorithm was introduced in 1966 [1]. It constructs a minimum realization (A,B,C) of a linear, time-invariant, state space model, given a noise-free Hankel matrix with the correct size for the system.
Let a truncated series of Markov parameters be denoted by:

$$M_1, M_2, M_3, \ldots , M_L \quad \text{where } M_i = CA^{i-1}B. \tag{1.1}$$

Then a Hankel matrix and its decomposition can be written as:

$$H = \begin{bmatrix} M_1 & M_2 & M_3 & \cdot & \cdot & M_{L/2} \\ M_2 & M_3 & M_4 & & & M_{L/2+1} \\ \vdots & \vdots & \vdots & & & \vdots \\ M_{L/2} & \cdot & \cdot & \cdot & \cdot & M_{L-1} \end{bmatrix} =$$

$$= \Gamma.\Delta = \begin{bmatrix} C \\ CA \\ \cdot \\ \cdot \\ CA^{L/2-1} \end{bmatrix} \begin{bmatrix} B & AB & A^2B & A^3B & \cdot & \cdot & \cdot & A^{L/2-1}B \end{bmatrix} \tag{1.2}$$

If the number of Markov parameters L exceeds twice the dimension n of the minimum realization of the system, and if we are dealing with a completely observable and completely controllable system we are sure that the rank of H equals n. Apart from this, under these conditions, any such decomposition into Γ and Δ of minimal dimension n and full rank will lead to a minimum realization (A,B,C). The complete set of all possible (Γ,Δ) together then produces the complete equivalence class of the system under study. It is well known that the elements of both the set (Γ,Δ) and the set (A,B,C) can be transformed into each other by means of a non-singular square matrix T of dimension n.

The triplet (A,B,C) can be obtained from (Γ,Δ) as follows: the matrices B and C can be recognized as the first blocks in Δ and Γ respectively. In order to obtain matrix A, we need a shifted matrix, which we indicate by an arrow. A vertically pointing arrow indicates a shift of one block row, whereas a horizontally pointing arrow denotes a shift of one block column. From this it is clear that

$$H\uparrow = \overleftarrow{H} = \Gamma A \Delta \tag{1.3}$$

and as Γ and Δ have a maximum rank n, we may write:

$$A = \Gamma^{+} H\uparrow \Delta^{+} \tag{1.4}$$

where + stands for pseudo inverse.

Note that we needed the extra Markov parameter M_L to construct the shifted Hankel matrix.

An alternative procedure is provided by either the 'extended observability' matrix Γ or the 'extended controllability' matrix Δ, because:

$$\Gamma\uparrow = \Gamma A \qquad \text{or} \qquad A\Delta = \overleftarrow{\Delta} \tag{1.5}$$

Consequently this yields:

$$A = \Gamma^{+}\Gamma\uparrow \qquad \text{or} \qquad A = \overleftarrow{\Delta}\,\Delta^{+} \tag{1.6}$$

As we do not have any information on what to insert in the latter blocks of Γ or Δ during the shift operation, we have to apply a reduced version of Γ and Δ.

Finally, a numerically stable decomposition of H is offered by the singular value decomposition, given by

$$H = U\ D\ V^{T} \qquad \dim(H) = g \cdot l \tag{1.7}$$

where $D = \text{diag}\ (\delta_1,\ \delta_2,\ \delta_3, \ldots\ldots, \delta_s)$

$$\delta_1 > \delta_2 > \delta_3 > \ldots\ldots > \delta_n > 0$$

$$\delta_{n+1} = \delta_{n+2} = \ldots\ldots = \delta_s = 0$$

$$s = \min\ (g,l)$$

$$U^{T}U = I_s \qquad \text{and} \qquad V^{T}V = I_s$$

Because rank(H) = n we may rewrite this as:

$$H = H_n = U_n\ D_n\ V_n^{T} \tag{1.8}$$

where for H_n only the first n non-zero singular values are used in D_n, and the corresponding singular vectors (i.e. columns) in U_n and V_n.

If we distribute the singular values in a balanced way among Γ and Δ we get:

$$\Gamma = U_n \, D_n^{1/2} \qquad \text{and} \qquad \Delta = D_n^{1/2} \, V_n \tag{1.9}$$

1.2. APPLICATION OF THE HO-KALMAN ALGORITHM FOR STOCHASTIC DATA

In 1974 Zeiger and McEwen [2] suggested using the Ho-Kalman algorithm with the singular value decomposition with noisy data. In that case, all singular values up to and including $\hat{\delta}_s$ will be non-zero. (The distinction with the deterministic case is denoted by the circumflex $\hat{}$).
It is easy to verify that, in cases where the Markov parameters are contaminated with SWAYING noise, we may write (see appendix):

$$E\{\hat{\delta}_i^2\} = \delta_i^2 + \sigma^2 . \max(1,g) \tag{1.10}$$

where E stands for expectation.
SWAYING noise is defined as follows:

$$\hat{M}_i \,(a,b) = M_i(a,b) + \xi_{iab} \tag{1.11}$$

where $M_i(a,b)$ is element a,b in matrix M_i and ξ_{iab} is the corresponding additive noise. This noise is assumed to be stationary (S), white (W) (zero mean), additive (A), signal-independent (Y), inter-independent (among channels) (I), with non-changing global variance σ^2 (NG):

$$E\{\xi_{iab}\} = 0 \qquad i,a,b \in \mathbb{N} \tag{1.12}$$

$$E\{M_i(a,b)\;\xi_{jcd}\} = 0 \qquad i,a,b,j,c,d \in \mathbb{N} \tag{1.13}$$

$$E\{\xi_{iab}\xi_{jcd}\} = \begin{matrix} 0 & i,a,b \neq j,c,d \\ \sigma^2 & i,a,b = j,c,a \end{matrix} \tag{1.14}$$

This increase in singular values is reflected in Fig. 1.

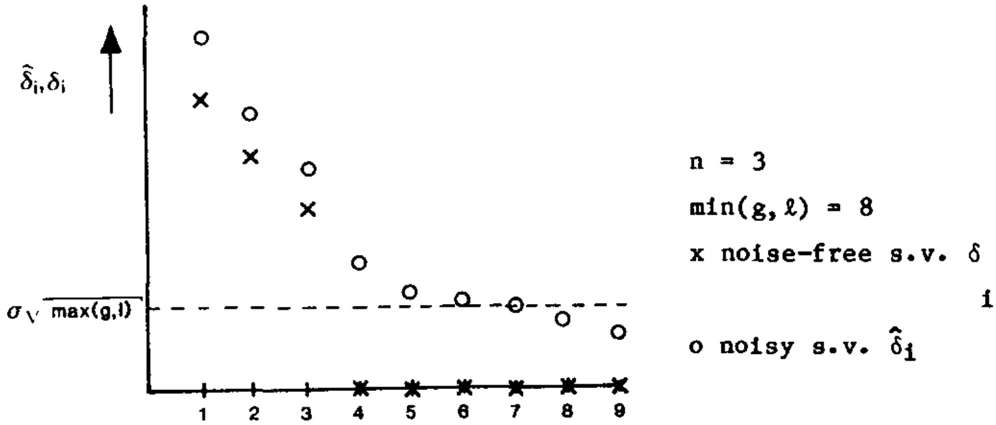

Fig. 1   Singular values of the Hankel matrix

Based upon knowledge of the noise level in the singular value, one may decide upon the dimension n of the system and approximate the Hankel matrix in a least squares sense by

$$H_n = U_n D_n V_n^T \tag{1.15}$$

(see e.g. [3]).

In this way we perform a noise filtering on the Hankel matrix and implicitly use the singular values for the order testing, before we apply the Ho-Kalman algorithm.
Then, there exist several possibilities for applying the formulae of the previous section 1.1 in order to find a realization. These possibilities are compared in [4].
For the shifted Hankel matrix we may either use the original H [5], or the approximating H_n [6]. Alternatively, the Γ and Δ matrices may be used [7] to evaluate A.
However, all methods using the Ho-Kalman algorithm for noisy Markov parameters show severe drawbacks, which are mainly due to the least squares approximation of the Hankel matrix. Although the approximating Hankel matrix H_n may have the proper rank, it lacks the necessary Hankel structure. Therefore it does not provide

directly unique Markov parameters. From another point of view, the Markov parameters, that constitute the Hankel matrix, are weighted by an isosceles triangular function with the top on L/2 for $M_{L/2}$. This is due to the fact that, depending on their index, the Markov parameters appear more frequently in the Hankel matrix.

Statistical considerations are also difficult to make, as the noise on the entries in the Hankel matrix is not independent, but exactly the same noise data appear frequently in several entries.
Because of all these drawbacks when using the Hankel matrix, we will now introduce an alternative matrix.

1.3. INTRODUCTION OF THE PAGE MATRIX

In order to overcome the problems caused by the special block structure of the Hankel matrix, a trivial matrix is introduced here, which is constructed from the Markov parameters in the most natural way.
Similarly to filling a page with characters, it follows that this matrix should be called the Page matrix. It is defined as:

$$P = \begin{bmatrix} M_1 & M_2 & M_3 & \cdots & M_\mu \\ M_{\mu+1} & \cdot & \cdot & \cdot & M_{2\mu} \\ \cdot & & & & \cdot \\ M_{(\eta-1)\mu+1} & & \cdot & \cdot & M_{\eta\mu} \end{bmatrix} \qquad \dim(P) = h \times m \qquad (1.16)$$

As ever , some nations show a different behaviour in this respect, and consequently we may also define a Chinese Page matrix C_P, where the Markov parameters are ordered column-wise.
We will not analyse this Chinese Page matrix meticulously here, as this analysis is completely dual to the normal Page matrix.

In the purely deterministic case, the Page matrix can be decomposed in a manner similar to the Hankel matrix:

$$P = \begin{bmatrix} C \\ CA^{\mu} \\ \cdot \\ \cdot \\ CA^{(\eta-1)\mu} \end{bmatrix} \cdot \begin{bmatrix} B & AB & A^2B & \cdot & \cdot & A^{\mu-1}B \end{bmatrix} \qquad (1.17)$$

$$P = \Gamma_{\mu} \cdot \Delta$$

Whereas the Hankel matrix is the product of the extended observability matrix Γ and the extended controllability matrix Δ of the system (A,B,C), the Page matrix is the product of the extended observability matrix of the system (A^{μ},B,C) and the extended controllability matrix of the system (A,B,C).
Next we will state the crucial theorem in this context:

<u>Theorem:</u>
If the dimensions of the Page matrix are chosen large enough : $\mu, \eta \geq n$(= the dimension of the system), and if (C,A^{μ}) is a completely observable couple, it holds that

$$\text{rank } P = n \qquad (1.18)$$

and any decomposition in Γ_{μ} and Δ of minimum dimension n will lead to a minimum realization.

The formal proof of this theorem is straightforward and given in chapter 2. Furthermore in chapter 2 all conditions for the observability of the couple (C,A^{μ}) are stated.
As a short outline we give the following summary.

Because exclusively completely observable and controllable systems (A,B,C) are considered here, non-observability of the system (A^{μ},B,C) is a situation that is quite exceptional. It can only happen if distinct poles in A happen to be non-distinct ones in A^{μ}.
Let the position of a pole of system matrix A in the complex z-plane be characterized by the radius r and the argument ϕ.
Then the corresponding pole for system matrix A^{μ} is characterized

by r^{μ} and $\mu\phi$. Consequently, original poles of equal radius may coincide in system matrix A^{μ} in the case of $\mu\phi_i = \mu\phi_j + 2k\pi$, $k \in \mathbb{Z}$. Then the non-observability might occur if the corresponding columns of C_J are dependent, where (C_J, A_J^{μ}) is the Jordan canonical form of (C,A^{μ}).
It is obvious that this can be avoided by requiring that for all poles $\mu\,|\phi| < \pi$ holds ,which, for example, implies a sufficiently high sampling rate.

Finally,as a result of non-observability in (C,A^{μ}), the possible multiple poles in the origin of the z-plane remain. These poles occur in the case of delays and impulse responses of finite length. This is discussed and elaborated upon in chapter 2.
From here on we will assume (C,A^{μ}) to be completely observable.
In the dual case of the Chinese Page matrix (which could be used in cases of a suspected failure of the Page matrix), the assumption of complete controllability of (A^{μ},B) is made.
Under these conditions the rank of the Page matrix simply defines the order, and the realization (A,B,C) can be obtained similarly as for the Hankel matrix, apart from the fact that the proper shifted matrices have to be used. The Page matrix has to be shifted one block to the left, and the Chinese Page matrix one block upwards.
For the use of formula (1.5) the Page matrix provides the proper Δ and the Chinese Page matrix the proper Γ.

1.4. A COMPARISON OF THE HANKEL AND PAGE MATRIX

In the deterministic case, both the Hankel and the Page matrices can be used to obtain a minimum realization.
As indicated, the Page matrix may fail in some exceptional cases, but the size of the Page matrix is much smaller (h.m = L-1) than the size of the Hankel matrix (l.g = $L^2/4$). Consequently, the reduction in the computational effort is considerable.

Nevertheless, the superiority of the Page matrix is much more significant for the noisy case. In the Page matrix all Markov parameters appear only once, which means that, when reducing the rank with the aid of singular value decomposition, there is an equally balanced filtering over the parameters. Moreover, a Page matrix of reduced rank provides a unique sequence of Markov parameters.
So the noise filtering by means of the singular value decomposition of the Page matrix is simply a least squares approximation of the Markov parameters with a fixed (or estimated) dimension of the system.

This noise filtering operation provides us with a criterion for the optimum size of the Page matrix.
For a given number L of available Markov parameters, the block dimensions of the Page matrix can be chosen in different combinations as long as $\eta\mu = L-1$.
If we assume that the Markov parameters in the Page matrix are disturbed with SWAYING noise, the total expected noise energy in P (in expectation) can be written as:

$$h \cdot m \cdot \sigma^2 \tag{1.19}$$

Because of the character of the noise, the expected noise energy will, if applying the singular value decomposition to P, be equally distributed over all squared singular values (see Appendix). When the rank of the Page matrix is reduced to n, by setting min(h,m)-n singular values equal to zero, the expected noise reduction equals:

$$\frac{n \cdot \max(h,m) \cdot \sigma^2}{h \cdot m \cdot \sigma^2} = \frac{n}{\min(h,m)} \tag{1.20}$$

For optimizing this noise reduction, min(h,m) has to be maximized. This implies that we have to choose P as close to square as possible.

Chapter 2. VALIDITY OF THE REALIZATION PROCEDURE USING THE PAGE MATRIX

2.1 PROBLEM REDUCTION TO THE OBSERVABILITY OF SYSTEM (A^{μ},B,C)

This chapter is devoted to the central theorem used in chapter 1, where the idea of a Page matrix was introduced.
Assuming a sequence of deterministic Markov parameters $\{M_k\}_{k=1..,L}$ with a finite dimensional realization (A,B,C), the Hankel matrix can be written as:

$$H = \Gamma \cdot \Delta \qquad \dim(H)=g\times l \tag{2.1}$$

where

$$\Gamma = \begin{bmatrix} C \\ CA \\ CA^2 \\ \cdot \\ \cdot \\ CA^{\gamma-1} \end{bmatrix} \qquad \Delta = \begin{bmatrix} B & AB & A^2B & \cdot & \cdot & A^{\gamma-1}B \end{bmatrix} \tag{2.2}$$

$$\dim(\Gamma)=g\times n$$
$$\dim(\Delta)=n\times l$$

where $\gamma=L/2$.

If γ is large enough, the extended observability matrix Γ and the extended controllability matrix Δ have full rank n. The limit which we will use is $\gamma \geq n$ though theoretically the condition $\gamma \geq r$ is sufficient where r is the realizability index.
The Hankel matrix H transforms the space R_l into R_g via R_n and this is made explicit by means of Δ ($R_l \rightarrow R_n$) and Γ ($R_n \rightarrow R_g$). A lot of freedom, however, is left by not defining the base in R_n. All possibilities together define the equivalence class which is invariant under the following equivalence transformation:
If the following two sets correspond to the same Hankel matrix:

$$(\Gamma,\Delta,A,B,C) \xleftrightarrow{\text{equivalent in } H} (\Gamma^*,\Delta^*,A^*,B^*,C^*)$$

then the equivalence transformation is given by:

$$\begin{aligned} \Gamma &= \Gamma^* T^{-1} \qquad & \Delta &= T\,\Delta^* \\ C &= C^* T^{-1} \qquad & B &= T\,B^* \\ A &= T\ A^*\ T^{-1} \qquad & & \text{where } T \text{ nonsingular}, \dim(T) = n \times n \end{aligned} \tag{2.3}$$

A numerically stable solution for Γ and Δ can be found from a singular value decomposition of H :

$$H = U\,D\,V^T$$

and consequently:

$$H = (UD^{\frac{1}{2}})(D^{\frac{1}{2}}V^T) = \Gamma\ \Delta \tag{2.4}$$

so that we may choose: $\Gamma = U\,D^{\frac{1}{2}}$

$$\Delta = D^{\frac{1}{2}}\,V^T \tag{2.5}$$

(Because of the orthonormality of U and V, both matrices $(UD^{\frac{1}{2}})$ and $(D^{\frac{1}{2}}V^T)$ will have rank n.)

Based on this decomposition of H into two matrices of rank n, a minimum realization (A,B,C) can be found (see [2] and [4]) by:

$$C = E_q^{\gamma q}\ U\ \ D^{\frac{1}{2}} \tag{2.6}$$

$$B = D^{\frac{1}{2}}\ \ V^T\ \ E_{\gamma p}^{p} \tag{2.7}$$

$$A = D^{-\frac{1}{2}}\ U^T\ \overleftarrow{H}\ \ V\ \ D^{-\frac{1}{2}} \tag{2.8}$$

where

$$E_q^{\gamma q} = [\ I_q\ \ \emptyset_q\ \ \emptyset_q\ \ \cdot\ \ \cdot\ \ \emptyset_q\ ] \tag{2.9}$$

$$E_{\gamma p}^{p} = [\ I_p\ \ \emptyset_p\ \ \emptyset_p\ \ \cdot\ \ \cdot\ \ \emptyset_p\ ]^T \tag{2.10}$$

and the shifted Hankel matrix:

$$\overleftarrow{H} = \Gamma \cdot A \cdot \Delta = \begin{bmatrix} M_2 & M_3 & M_4 & \cdot & \cdot & M_{\gamma+1} \\ M_3 & M_4 & \cdot & \cdot & \cdot & M_{\gamma+2} \\ M_4 & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ M_{\gamma+1} & M_{\gamma+2} & \cdot & \cdot & \cdot & M_{2\gamma} \end{bmatrix} \tag{2.11}$$

Whether this same algorithm for finding a minimum realization may be applied to the Page matrix is dependent on the fact whether the rank of P equals the minimum dimension of the system n.
The Page matrix may be decomposed as:

$$P = \Gamma_\mu \cdot \Delta \qquad \dim(P) = h \times m \tag{2.12}$$

where

$$\Gamma_\mu = \begin{bmatrix} C \\ CA^{\mu} \\ CA^{2\mu} \\ \cdot \\ \cdot \\ CA^{(\eta-1)\mu} \end{bmatrix} \qquad \begin{aligned} \Delta &= \left[\, B \;\; AB \;\; A^2B \;\; \cdot \;\; \cdot \;\; A^{\mu-1}B \,\right] \\ \dim(\Gamma_\mu) &= h \times n \\ \dim(\Delta) &= n \times m \end{aligned} \tag{2.13}$$

It is obvious, that <u>if</u> $\eta > n$ and $\mu > n$ and <u>both</u> the system (A,B,C) is completely observable and controllable <u>and the system (A^μ,B,C) is completely observable</u>, the <u>same</u> equivalence class can be defined for P as for H, with exactly the <u>same</u> equivalence transformation.
The crucial condition is that the system (A^μ,B,C) is completely observable, because then the extended observability matrix Γ_μ has full rank n.
(In the dual case of the Chinese Page matrix, the controllability of (A^η,B,C) is required to assure the full rank of Δ).
It can be proved that the rank $\Gamma_\mu = n$ (which will be the hardest job), then the following can be stated about the rank of P; A general rule for the rank of a product of matrices XY is:
rank XY $\leq$ min(rank X, rank Y). The Sylvester's inequality (see [8] p.66) is an extension of this relation: it states that for the rank

of the product of rectangular matrices X and Y of dimensions resp. $m \times n$ and $n \times q$, it holds that:

$$\text{rank } X + \text{rank } Y - n \leq \text{rank}(XY) \leq \min(\text{rank } X, \text{rank } Y)$$

Applying this inequality to the Page matrix (eq.(2.12)), this leads to

$$n \leq \text{rank } P \leq n \tag{2.14}$$

And as the result :

$$\text{rank } P = n \tag{2.15}$$

If rank $\Gamma_\mu = n$ and consequently rank $P = n$, a singular value decomposition of P leads to a decomposition of P into two matrices each with rank n. From this decomposition a minimum realization can be found in the same way as for the Hankel matrix, i.c. using the Ho-Kalman algorithm. As for the Hankel matrix all the subsequent steps are invariant under the equivalence transformation (with nonsingular matrix T).

In the remaining part of this chapter the condition rank $P = n$ will be a subject of study. As a summary it can be stated that if we can prove that rank $\Gamma_\mu = n$, this immediately leads to the required rank condition for P. This feature will be discussed in the next section, where we will just indicate all exceptions for which rank $\Gamma_\mu \neq n$.

2.2 OBSERVABILITY OF (A,B,C)

2.2.1. INTRODUCTION

Our task is to prove that

$$\text{rank} \begin{bmatrix} C \\ CA^{\mu} \\ CA^{2\mu} \\ \cdot \\ \cdot \\ CA^{(n-1)\mu} \end{bmatrix} = n \tag{2.16}$$

given a minimum realization (A,B,C). If (A,B,C) is a minimum realization, it is proved [1] that

$$\operatorname{rank}\begin{bmatrix} C \\ CA \\ \cdot \\ \cdot \\ CA^{\eta-1} \end{bmatrix} = n \qquad (2.17)$$

in other words : (A,B,C) is a completely observable realization. Rank $\Gamma_\mu = n$ can only be assured if $\eta > n$. This is a similar condition as the condition $\mu > n$ for the extended controllability matrix Δ_η. The condition rank $\Gamma_\mu = n$ corresponds to the statement that (A^μ, B, C) is a completely observable system. The final formulation of the problem now becomes:

Given a completely observable system (A,B,C); under which conditions is the system (A^μ, B, C) also completely observable for any $\mu > 1$.

To deal with the problem of complete observability of systems a congruent definition of this feature will be introduced, taking into account the structure of the matrix A, and more specifically its eigenvalues.
Such a definition is given by Chen and Desoer [9] in the first instance for complete controllability. Their theorem will be stated here and the proof of the theorem will be given along the same lines as they did. This will be necessary for applying the theorem to the situation of the system (A^μ, B, C). For this purpose the Jordan canonical form is necessary, which will be defined next.

2.2.2. JORDAN CANONICAL FORM

Consider a system with ν different eigenvalues. The system can always be represented in its Jordan canonical form, in which there are no two Jordan blocks associated with the same eigenvalue. This Jordan form can be written as follows:

$$A_J = \begin{bmatrix} A_1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & A_2 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & & \cdot & \cdot \\ 0 & \cdot & & \cdot & \cdot & A_\nu \end{bmatrix}_{(n\times n)} , \quad C_J = \underset{(q\times n)}{[\, C_1 \; C_2 \; \cdot \; \cdot \; C_\nu \,]} \quad (2.18)$$

where A_i and C_i will denote all Jordan blocks associated with eigenvalue λ_i.

Every Jordan block can be represented by a numer of Jordan cages, again ordered in a block diagonal way:

$$A_i = \begin{bmatrix} A_{i1} & & & & \\ & A_{i2} & & \emptyset & \\ & & \cdot & & \\ & \emptyset & & \cdot & \\ & & & & A_{ir(i)} \end{bmatrix}_{(n_i\times n_i)} , \quad C_i = \underset{(q\times n)}{[\, c_{i1} \; c_{i2} \; \cdot \; \cdot \; c_{ir(i)} \,]} \quad (2.19)$$

With every eigenvalue λ_i there are associated r(i) Jordan cages. This Jordan cages have the following form:

$$A_{ij} = \begin{bmatrix} \lambda_i & 1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & \lambda_i & 1 & 0 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & 0 & \cdot & & & \cdot & 1 \\ 0 & \cdot & \cdot & \cdot & \cdot & & \lambda_i \end{bmatrix}_{(n_{ij}\times n_{ij})} , \quad C_{ij} = \underset{(q\times n_{ij})}{[\, \underline{c}_{1ij} \; \underline{c}_{2ij} \cdot \cdot \, \underline{c}_{nij} \,]} \quad (2.20)$$

where $\underline{c}_{1ij}$ and $\underline{c}_{nij}$ are the first and the last column of C_{ij}, respectively.

The matrix A in the Jordan canonical form is completely defined by all numbers $\lambda_i \in \mathbb{C}$,and $n_{ij}, j=1..r(i)$, $i=1...\nu$.

λ_i, $i=1..\nu$ define the values of the diagonal elements,

$n_{ij}, i=1...\nu, j=1...r(i)$ define the structure of matrix A

$$\sum_{j=1}^{r(i)} n_{ij} = \text{the multiplicity of "pole" } \lambda_i$$

For the remainder of this chapter we assume,that (A,B,C) has been brought into a Jordan canonical form (A_J,B_J,C_J), so that the index J will be dropped.

2.2.3. A CRITERION FOR OBSERVABILITY OF (A,B,C)

The theorem of Chen and Desoer [9] now states the following: The system (A,B,C) is completely observable if and only if the condition E holds, where

E: for each $i=1,2,\ldots\nu$, the set of r(i) q-dimensional column vectors $\underline{c}_{1i1},\underline{c}_{1i2},\ldots\underline{c}_{1ir(i)}$ is a linearly independent set.

Note that these are the columns of C corresponding to <u>that</u> (first) state in each cage, which is independent on <u>all</u> other states in that cage.
In the dual case of controllability ,those rows of B have to form an independent set which correspond to <u>that</u> (last) state in each cage, which is independent on (but influencing) all other states in that cage. (see [9])
For poles with multiplicity one (single poles) the set consists of just one element (row or column) and independence then means, that this isn't a zero vector. At least one input should influence the corresponding state or in the dual case at least from one output one should be able to observe the corresponding state.

To prove this theorem there has to be demonstrated that this property of A and C fits with the definition of Kalman:

$$\text{rank } \Gamma = n \tag{2.21}$$

For the clearness of this text there will be defined:

$$\Gamma(A,C,n) = \begin{bmatrix} C \\ CA \\ CA^2 \\ \cdot \\ \cdot \\ CA^{n-1} \end{bmatrix} \tag{2.22}$$

To prove the equivalence relation

$$\text{condition } E \leftrightarrow \text{rank } \Gamma(A,C,n) = n \quad (2.23)$$

we will need three assertions that will be stated first:

Assertion 1: For any integer n, any complex λ, and any matrices A and C having proper dimensions

$$\text{rank } \Gamma(A,C,n) = \text{rank } \Gamma(A-\lambda I,C,n) \quad (2.24)$$

Proof : $C(A-\lambda I)^i$ can always be written as a linear combination of CA^i, CA^{i-1} . . ., C, and therefore it holds that $\Gamma(A-\lambda I,C,n)$ is obtainable from $\Gamma(A,C,n)$ by applying to it a sequence of elementary row operations. More specifically there exists a nonsingular matrix Q(hxh) in such a way that

$$\Gamma(A-\lambda I,C,n) = Q\Gamma(A,C,n) \quad (2.25)$$

According to the Sylvester inequality ([8] p.18) it can be written:

$$\text{rank}\,\Gamma(A,C,n)+h-h \leq \text{rank}\,\Gamma(A-\lambda I,C,n) \leq \text{rank}\,\Gamma(A,C,n) \quad (2.26)$$

Therefore it follows that eq. (2.24) holds.

In words : the rank of the observability matrix will not change when all diagonal elements of A are increased or decreased with the same constant $\lambda \in \mathbb{C}$ (which simply means a change of the origin in the complex z-plane).

Assertion 2: rank $\Gamma(A_i,C_i,s)$ = rank $\Gamma(A_i,C_i,\overline{n_i})$ for all $s > \overline{n_i}$, where $\overline{n_i} = \max\{ n_{ij}, j=1,2,\ldots r(i)\}$.

Proof: Because of the specjal Jordan structure of A_i it follows that

$$(A_i - \lambda_i I)^{\overline{n_i}} = \emptyset \tag{2.27}$$

and therefore $\qquad (A_i - \lambda_i I)^s = \emptyset \qquad$ for all $s \geq \overline{n_i}$ $\qquad$ (2.28)

Because of assertion 1:

$$\text{rank } \Gamma(A_i, C_i, s) = \text{rank } \Gamma(A_i - \lambda_i I, C_i, s) \tag{2.29}$$

With the result above this leads to

$$\text{rank } \Gamma(A_i, C_i, s) = \text{rank } \Gamma(A_i, C_i, n_i) \quad \text{for all } s \geq \overline{n_i} \tag{2.30}$$

In words: a Jordan canonical form of a system with one distinct eigenvalue, forms an observability matrix for which holds that the number of matrix products that has to be taken into account to come to the maximal rank of Γ, is determined by the dimension of the largest Jordan cage.

Assertion 3: If there exist a nonzero n-dimensional column vector $\underline{q}$ such that $\Gamma(A,C,n)\underline{q} = \underline{0}$, then for any complex λ, $\Gamma(A-\lambda I,C,n)\underline{q} = \underline{0}$

Proof: From assertion 1 it follows that

$$\Gamma(A-\lambda I,C,n) = Q\Gamma(A,C,n) \tag{2.31}$$

consequently if $\Gamma(A,C,n)\underline{q} = \underline{0}$, then for any $\lambda \in \mathbb{C}$

$$\Gamma(A-\lambda I,C,n)\underline{q} = \underline{0} \tag{2.32}$$

For the proof of the theorem, as stated at the beginning of this section, some more resources will be needed ;a schematical way of representing $\Gamma(A,C,n)$ is given next.

Because of the block diagonal structure of A, $\Gamma(A,C,n)$ can be written as:

$$\Gamma(A,C,n) = \left[\Gamma(A_1,C_1,n) \quad \Gamma(A_2,C_2,n) \quad \cdot \quad \cdot \quad \cdot \quad \Gamma(A_\nu,C_\nu,n)\right] =$$
$$= \left[\Gamma(A_{11},C_{11},n) \quad \cdot \quad \cdot \quad \cdot \quad \Gamma(A_{\nu r(\nu)},C_{\nu r(\nu)},n)\right] \quad (2.33)$$

When we take the Jordan block associated with eigenvalue λ_k, then for the Jordan cage with index j these can be written.

$$A_{kj} - \lambda_k I = \underset{(n_{kj} \times n_{kj})}{\begin{bmatrix} 0 & 1 & 0 & \cdot & \cdot & 0 \\ 0 & 0 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 1 \\ 0 & \cdot & \cdot & \cdot & \cdot & 0 \end{bmatrix}} \quad (2.34)$$

which leads to

$$\Gamma(A_{kj}-\lambda_k I, C_{kj}, n) = \begin{bmatrix} \underline{c}_{1kj} & \underline{c}_{2kj} & \cdot & \cdot & \cdot & \underline{c}_{nkj} \\ 0 & \underline{c}_{1kj} & \cdot & \cdot & \cdot & \underline{c}_{(n-1)kj} \\ \cdot & & \cdot & & & \cdot \\ \cdot & \cdot & \cdot & \cdot & & \cdot \\ 0 & \cdot & \cdot & \cdot & & \underline{c}_{1kj} \\ 0 & \cdot & \cdot & \cdot & & 0 \\ \cdot & \cdot & \cdot & \cdot & & \cdot \\ 0 & \cdot & \cdot & \cdot & & 0 \end{bmatrix} \quad (2.35)$$

Now,all instruments are available to prove the theorem.

2.2.4. PROOF OF THE THEOREM OF CHEN AND DESOER

There is going to be proved that rank $\Gamma(A,C,n) = n$ if and only if condition E, as described in section 2.2.3., holds.

1. Necessary condition:($\rightarrow$)

Suppose rank $\Gamma(A,C,n) = n$ and E does not hold for some $i=k$. This means that the set

$\underline{c}_{1k1},\underline{c}_{1k2},\ldots,\underline{c}_{1kr(k)}$ is a linearly dependent set.

Now consider the way of writing $\Gamma(A,C,n)$ and $\Gamma(A_k-\lambda_k I,C_k,n)$ as in eq.(2.33) and (2.35). If the given set of $\underline{c}$-vectors is linearly dependent, then there exists a linear combination of columns of $\Gamma(A-\lambda_k I,C,n)$ that is linear dependent. From this it follows that rank $\Gamma(A-\lambda_k I,C,n) < n$ and , by assertion 1, that rank $\Gamma(A,C,n) < n$. This is in contradiction with the hypothesis, and therefore for rank $\Gamma(A,C,n) = n$, a necessary condition is given by E.

2. Sufficient condition:($\leftarrow$)

This proof will be done in two steps.

First it will be demonstrated that if E holds, this leads to:

$$\text{rank } \Gamma(A_i,C_i,n) = n_i \qquad \text{for all } i \tag{2.36}$$

Given this condition it will be proved that rank $\Gamma(A,C,n) = n$.

a. If E holds then rank $\Gamma(A_i,C_i,n) = n_i$ for all i.

Proof: From assertion 1 it follows that

$$\text{rank } \Gamma(A_i, C_i, n) = \text{rank } \Gamma(A_i - \lambda_i I, C_i, n) \tag{2.37}$$

With assertion 2 there can be written :

$$\text{rank } \Gamma(A_i, C_i, n) = \text{rank } \Gamma(A_i - \lambda_i I, C_i, \overline{n_i}) \tag{2.38}$$

Now suppose that rank $\Gamma(A_i - \lambda_i I, C_i, \overline{n_i}) < n_i$. The considered matrix has dimensions $\overline{n}_i q \times n_i$ (q= number of outputs); if its rank is smaller then n_i, a linear combination of the n_i column vectors of Γ can be brought to zero.

In other words:

There exists a nonzero column vector $\underline{q}$ in such a way that

$$\Gamma(A_i - \lambda_i I, C_i, \overline{n_i})\underline{q} = \underline{0} \tag{2.39}$$

The matrix $\Gamma(A_i - \lambda_i I, C_i, \overline{n_i})$ can be written as (in the example $\overline{n_i} = n_{1k}$)

$$\left[\begin{array}{cccc:cccc:c}
\underline{c}_{1k1} & \underline{c}_{2k1} & \cdots & \underline{c}_{nk1} & \underline{c}_{1k2} & \cdots & \cdots & \underline{c}_{nk2} & \\
 & \ddots & & \vdots & & \ddots & & \vdots & \\
 & & \ddots & \vdots & & & \ddots & \underline{c}_{1k2} & \cdots \\
 & \emptyset & & \vdots & & & & 0 & \\
 & & & \underline{c}_{1k1} & 0 & \cdots & \cdots & 0 &
\end{array}\right] \tag{2.40}$$

Because $\underline{c}_{1k1}, \underline{c}_{1k2}, \ldots \underline{c}_{1kr(k)}$ is a linearly independent set, it can be seen that $\Gamma(A_i - \lambda_i I, C_i, n_i) \cdot \underline{q}$ can only be $\underline{0}$ if $\underline{q} = \underline{0}$.
This is in contradiction with the assumption, and therefore it must hold that

$$\text{rank } \Gamma(A_i - \lambda_i I, C_i, \overline{n_i}) = n_i \tag{2.41}$$

Because $n > \overline{n_i}$, and with assertions 1. and 2. it follows that

$$\text{rank } \Gamma(A_i, C_i, n) = n_i \quad \text{for all } i \tag{2.42}$$

b. If E holds then all columns of $\Gamma(A_i,C_i,n)$ and $\Gamma(A_j,C_j,n), i\neq j$, will be independent.

Proof: Suppose

$$\text{rank}\ [\ \Gamma(A_1,C_1,n) \mid \Gamma(A_2,C_2,n)\ ] < n_1 + n_2 \tag{2.43}$$

Then there exists a n_1+n_2-dimensional column vector

$$\begin{bmatrix} q_1 \\ q_2 \end{bmatrix} \neq \underline{0}\ ,\ \text{and}\ [\Gamma(A_1,C_1,n) \mid \Gamma(A_2,C_2,n)] \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} = \underline{0} \tag{2.44}$$

With assertion 3 it follows then:

$$[\ \Gamma(A_1,C_1,n) \mid \Gamma(A_2,C_2,n)\ ] \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} = \underline{0} \tag{2.45}$$

$$(A_2-\lambda_2 I)^s = \emptyset \text{ for all } s > \overline{n_2} \tag{2.46}$$

Then it follows for the $(n-\overline{n_2})\times n_1$ left under part of the matrix equation (2.45):

$$\begin{bmatrix} C_1(A_1-\lambda_2 I)^{\overline{n_2}} \\ \cdot \\ \cdot \\ C_1(A_1-\lambda_2 I)^{n-1} \end{bmatrix} \cdot q_1 = \underline{0} \tag{2.47}$$

In other words:

$$\Gamma(A_1-\lambda_2 I, C\ \ , n-\overline{n_2})(A_1-\lambda_2 I)^{\overline{n_2}} \cdot q_1 = \underline{0} \tag{2.48}$$

$$\text{Since } \lambda_1 \neq \lambda_2\ ,\ \text{rank } (A_1-\lambda_2 I)^{\overline{n_2}} = n_1 \tag{2.49}$$

$$\begin{aligned} \text{rank } \Gamma(A_1-\lambda_2 I, C_1, n-\overline{n_2}) &= \text{rank } \Gamma(A_1,C_1,n-\overline{n_2}) = \\ &= \text{rank } \Gamma(A_1,C_1,\overline{n_1}), \text{because } \overline{n_1} < n-\overline{n_2} \end{aligned} \tag{2.50}$$

With equation (2.42) it has been proved that rank

$$\Gamma(A_i,C_i,\overline{n_i}) = n_i,\ \text{so rank } \Gamma(A_1,C_1,\overline{n_1}) = n_1 \tag{2.51}$$

Equations (2.49) and (2.50) together with the Sylvester inequality now show that equation (2.48) is a vector equation with a $n_1 \times n_1$-coefficient matrix of full rank n_1.

As a result equation (2.48) can only be fulfilled if $\underline{q_1}=\underline{0}$.
Along the same lines with substitution of λ_1 in equation (2.45) it follows that $\underline{q_2}=\underline{0}$.
Then $\underline{q_1}=\underline{0}$ and $\underline{q_2}=\underline{0}$ and this is in contradiction with the hypothesis.
Therefore:

$$\text{rank}\left[\Gamma(A_i,C_i,n) \mid \Gamma(A_j,C_j,n)\right] = n_i+n_j \quad (\text{if } \lambda_i \neq \lambda_j) \qquad (2.52)$$

RESULT:

When we consider the representaion of A and A_i in equation(2.18) and (2.19), it can be seen that $n = \sum_{i=1}^{r(i)} n_i$.
Because of the fact that every value of i is associated with a different λ_i, it follows that

$$\text{rank } \Gamma(A,C,n) = n \qquad (2.53)$$

With this result the theorem as stated in section 2.2.3 has been proved.

2.2.5. REMARKS

With the given criterion for complete observability it is much more easy to analyse the observability of a system with a more physical understanding than with the definition of Kalman, at least if the Jordan canonical form of the system is known.
Because of our purely theoretical interest in the definition at this moment, this criterion is very suitable.
In the next section it will be demonstrated under which circumstances (A^{μ},B,C) is completely observable, given (A,B,C) is completely observable. The special structure of the matrix A in the Jordan canonical form will be a great help in this task.

2.3. OBSERVABILITY OF (A^{μ},B,C)

2.3.1. JORDAN STRUCTURE OF A^{μ}

Our goal is to find all possible situations where the system (A^{μ},B,C) is nonobservable, whereas the system (A,B,C) is completely observable.
In these cases, and only in these cases, the Ho-Kalman algorithm applied to the Page matrix will not lead to a minimum realization.
The criterion for complete observability, as introduced by Chen and Desoer [9] and described in the previous section, can quite easily be applied to the new system (A^{μ},B,C).
When A is assumed to be in the Jordan canonical form, as described in equation (2.18) , A^{μ} can be found by raising each Jordan cage to the power μ.
It is known (see [8], p.154) that a Jordan cage associated with eigenvalue λ_i when raising it to the power μ can be written in the next form:
if $\mu > n_{ij}$

$$A_{ij}^{\mu} = \begin{bmatrix} \lambda_i^{\mu} & \mu\lambda_i^{\mu-1} & (\mu-1)\lambda_i^{\mu-2} & \cdot & \cdot & \cdot & (\mu-n_{ij}+2)\lambda_i^{\mu-n_{ij}+1} \\ 0 & \lambda_i^{\mu} & \mu\lambda_i^{\mu-1} & \cdot & \cdot & \cdot & (\mu-n_{ij}+3)\lambda_i^{\mu-n_{ij}+2} \\ 0 & 0 & \lambda_i^{\mu} & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \lambda_i^{\mu} \end{bmatrix} \tag{2.54a}$$

and if $\mu < n_{ij}$ there appears at least one 1 or in general:

$$A_{ij}^{\mu} = \begin{bmatrix} \lambda_i^{\mu} & \mu\lambda_i^{\mu-1} & (\mu-1)\lambda_i^{\mu-2} & \cdot & 2\lambda_i & 1 & \cdot & \cdot & \cdot & 0 \\ 0 & \lambda_i^{\mu} & \mu\lambda_i^{\mu-1} & \cdot & 0 & 2\lambda_i & 1 & 0 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \lambda_i^{\mu} \end{bmatrix} \tag{2.54b}$$

In order to find out under which condition the system (A^{μ},B,C) is completely observable it has to be investigated which properties of the matrix A are used in the criterion of observability in section 2.2.3.

If the criterion is also applicable to A-matrices with a structure as in equation (2.54) then no situations of nonobservability of (A^{μ},B,C) will occur.

For assertions 1. and 3. (section 2.2.3.) no restrictions on the matrix A are made. They hold for any matrix A, and therefore they are also applicable to matrix A^{μ}.

Assertion 2. (section 2.2.3.) assumes that for all i :

$$(A_i - \lambda_i I)^{\overline{n_i}} = \emptyset \tag{2.55}$$

$$\text{with } \overline{n_i} = \max_j n_{ij}$$

This assertion not only holds for a Jordan cage ,but for any right upper matrix with equal diagonal elements λ_i and dimension less than $\overline{n_i}$.

Because A^{μ}_{ij} fulfills this condition , assertion 2. will also remain valid.

In the proof of the theorem itself (section 2.2.4.) ,apart from the three assertions, only use has been made of the fact that matrix $\Gamma(A_i-\lambda_i I,C_i,\overline{n_i})$ could be written as:

$$\left[\begin{array}{ccccc|cccc|c}
\underline{c}_{1k1} & \underline{x} & \underline{x} & \cdot & \underline{x} & \underline{c}_{1k2} & \underline{x} & \cdot & \underline{x} & \cdot \\
\underline{0} & \underline{c}_{1k1} & \underline{x} & \cdot & \underline{x} & \underline{0} & \underline{c}_{1k2} & \cdot & \underline{x} & \cdot \\
\underline{0} & \underline{0} & \underline{c}_{1k1} & \cdot & \underline{x} & \underline{0} & \underline{0} & \cdot & \underline{x} & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \underline{0} & \cdot & \cdot & \underline{c}_{1k2} & \cdot \\
\underline{0} & \cdot & \cdot & \cdot & \underline{c}_{1k1} & \underline{0} & \cdot & \cdot & \underline{0} & \cdot
\end{array}\right]$$

where $\underline{x}$ denotes any column vector, which is irrelevant for the proof. In case we deal with cages in the form of (2.54) we get:

$$\left[\begin{array}{cccccc|cccc|c}
\underline{c}_{1k1} & \underline{x} & \underline{x} & \cdot & \underline{x} & & \underline{c}_{1k2} & \underline{x} & \cdot & \underline{x} & \cdot \\
\underline{0} & 2\lambda_i\underline{c}_{1k1} & \underline{x} & \cdot & \cdot & & \underline{0} & 2\lambda_i\underline{c}_{1k2} & \cdot & \cdot & \cdot \\
\underline{0} & \underline{0} & 2\lambda_i\underline{c}_{1k1} & \cdot & \cdot & & \underline{0} & \underline{0} & 2\lambda_i\underline{c}_{1k2} & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot & & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & & \underline{0} & \cdot & \cdot & 2\lambda_i\underline{c}_{1k2} & \cdot \\
\underline{0} & \cdot & \cdot & \cdot & 2\lambda_i\underline{c}_{1k1} & & \underline{0} & \cdot & \cdot & \underline{0} & \cdot
\end{array}\right]$$

and now also $\Gamma(A_i-\lambda_i I,C_i,\overline{n_i})\cdot\underline{q} = \underline{0}$ will always give $\underline{q}=\underline{0}$ unless $\lambda_i=0$.

The situation where $\lambda_i = 0$ gives rise to an exception, which we will discuss in the next section.
We may say now that the criterion for complete observability can also be applied to the system (A^μ,B,C), that's to say, the system (A^μ,B,C) is completely observable (controllable) if the system (A,B,C) is completely observable (controllable), apart from two exceptions, which <u>may</u> disturb this:

1. The Jordan structure has been changed because $\lambda_i^\mu = \lambda_j^\mu$ for some $i \neq j$
2. $\lambda_i = 0$ for some i

In section 2.2.2. we have seen that the Jordan structure of the matrix A is completely defined by all numbers n_{ij} $\begin{array}{l} i=1,\nu \\ j=1,r(i) \end{array}$

From the remarks in this section the conclusion can be drawn that the complete observability of (A^μ,B,C) for sure is preserved if all numbers n_{ij} of A and A^μ are the same. In this situation applying the criterion of observability to A,C and A^μ,C will lead to exactly the same results.
Problems may arise when the Jordan structure of A and A^μ are not the same. This will be dealt with in the next section.

2.3.2. SITUATIONS OF NONOBSERVABILITY

The Jordan cage structure of A^μ and A will not be the same in the following situations

a) $\lambda_i^\mu = \lambda_j^\mu$ for some $i \neq j$, while $\lambda_i \neq \lambda_j$ (2.56)

b) $\lambda_i = 0$ for some i (2.57)

Ad a) In this case two originally different eigenvalues of A will be transformed into two equal eigenvalues of A^μ. This means that the number of distinct eigenvalues ν in A is decreased for A^μ by one to ν-1 , and that two Jordan blocks are linked up into one block.
For complete observability of the system the set of r(i) q-dimensional column vectors $\underline{c}_{1i1},\underline{c}_{1i2},\dots\dots,\underline{c}_{1ir(i)}$ has to be a linearly independent set. When two Jordan blocks are linked up , r(i) increases and the condition of independence of the relevant column vectors of C has to be investigated again.
The independence of the new set of r(i) vectors is the only criterion for observability of (A^μ,B,C).

Ad b) From λ_i=0 it follows that $\lambda_i^k = 0$ for all k.
In case μ is sufficiently large ($\mu > n_{ij}$) we are dealing with the situation as in equation (2.45a). This leads to a matrix A^μ that is completely filled with zero's: $A^\mu = \emptyset$.
What originally were r(i) Jordan cages ,each with dimensions n_{ij},j=1,r(i), now become $n_i = \sum_{j=1}^{r(i)} n_{ij}$ Jordan cages of length 1.
In other words all zero eigenvalues of A become noncommon zero eigenvalues of A^μ.
For observability the criterion of independence of the set of r(i) vectors of C becomes a criterion of independence of n_i column vectors in case of A^μ.

This new criterion has to be tested again. It is clear that the criterion in the latter case can only be fulfilled if $q \geq n_i$.

In case μ is smaller than n_{ij} the matrix A^{μ}_{ij} becomes:

$$A^{\mu}_{ij} = \begin{bmatrix} 0 & 0 & \cdot & \cdot & 0 & 1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & \cdot & \cdot & 0 & 0 & 1 & 0 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & 0 \\ 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & 1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \end{bmatrix} \begin{matrix} \\ \\ \\ \\ \\ \left.\begin{matrix} \\ \\ \\ \end{matrix}\right\} w \end{matrix}$$

Here the latter w rows indicate the w quasi noncommon poles we have got in the system A^{μ}. This gives rise to a different Jordan form again, where we get w cages in stead of originally one. This can be accomplished by a suitable transformation matrix T, which just interchanges some states. Once this has been done, the criterion of Chen and Desoer may be applied again, which extends the relevant set with the columns of C corresponding to the indicated w states.

Example: $n_{ij} = 4$

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \mu = 2 \quad \rightarrow \quad A^{\mu} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} w = 2$$

interchange state 2 and 3:

$$T = T^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \rightarrow \quad A^{\mu} := \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

In both situations a) and b) the complete observability of (A^{μ},B,C) is determined by the (in)dependence of some specific column vectors of C.

2.3.3. REPLACEMENT OF POLES IN THE Z-DOMAIN

In this section we want to make clear in which situations the problem of nonobservability of (A^{μ},B,C) arises.

As mentioned in the previous section two situatiions can be distinguished:

a) $\lambda_i^{\mu} = \lambda_j^{\mu}$,while $\lambda_i \neq \lambda_j$ for $i \neq j$

b) $\lambda_i = 0$ for some i

Situation b) is very clear: a single pole in $z = 0$ will cause no problems because the Jordan structure of A^{μ} remains the same as the one for A.

Two or more common poles in $z = 0$ may cause nonobservability depending on C ,because they are transformed top noncommon poles for A^{μ} ; these noncommon poles can be non-distinguishable.

Situation a) may occur e.g. when $\lambda_i = -\lambda_j$. Then for all even μ holds:

$\lambda_i^{\mu} = (-\lambda_j)^{\mu} = \lambda_j^{\mu} > 0$ as shown in Fig. 2.1.

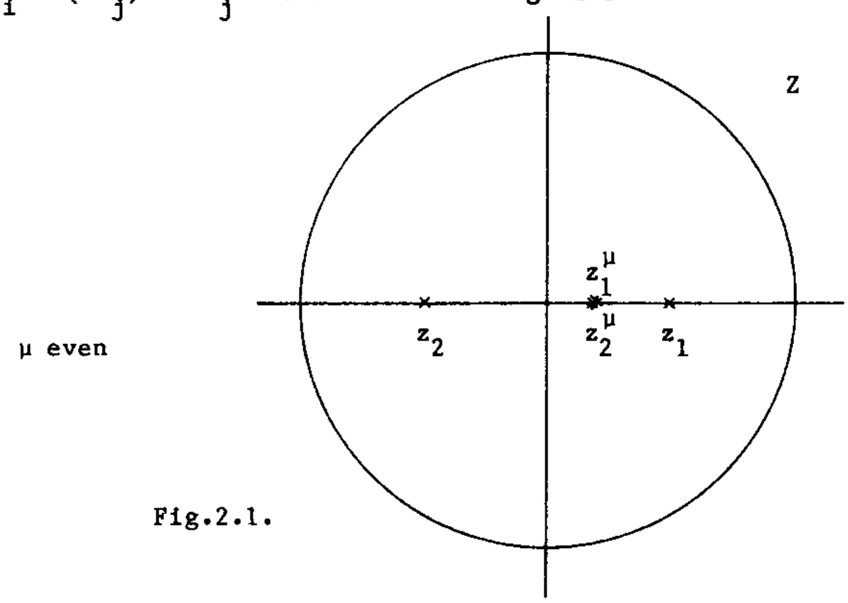

Fig.2.1.

A second possibility is that λ_i and λ_j are a complex conjugated pair $z_{1,2} = re^{\pm j\phi}$ and $\mu\phi = k\pi$,$k \in \mathbb{Z}$ (see Fig.2.2.).

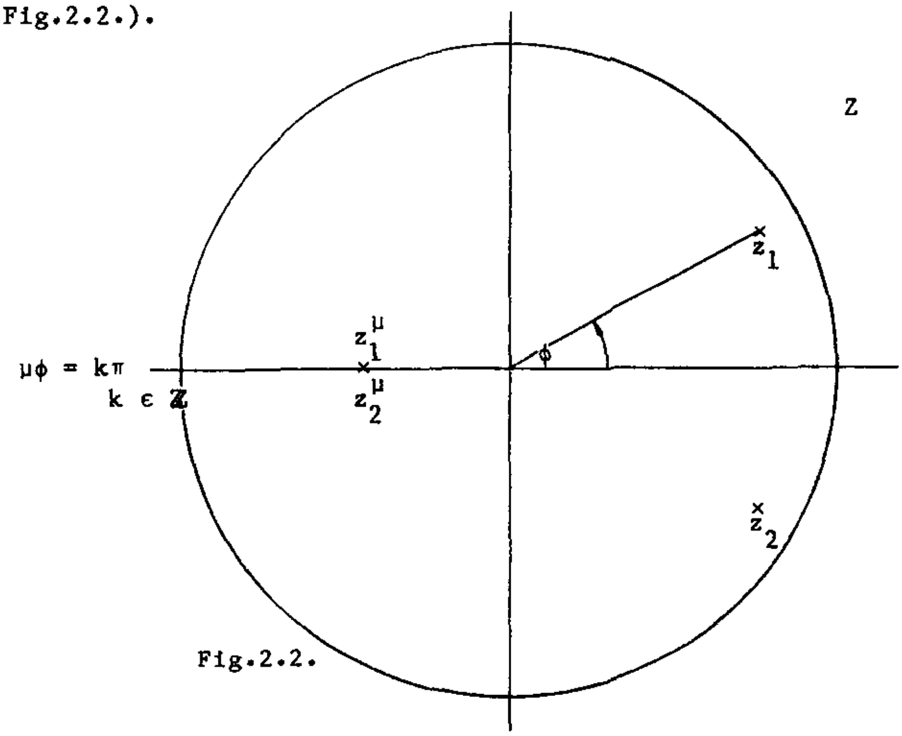

Fig.2.2.

And evidently all combinations of the two examples may occur. Notice that the final conclusion on observability of (A^μ,B,C) can only be drawn when knowing C. The described situations are the exclusive possibilities in which nonobservability might happen.

Besides, situation a) can sometimes be avoided if all poles are in the right half plane (high sampling rate) and by choosing μ small enough. Situation b) may be eliminated by excluding delays and finite responses.

2.4. REMARKS

For purpose of clearity the results of this chapter are briefly stated again:

Given an observable system (A,B,C).

The system (A^{μ},B,C) is completely observable, if condition F does <u>not</u> hold:

F 1)for some i: $\lambda_i = 0$ and $r(i) > 1$ <u>or</u>
for some $i \neq j$: $\lambda_i^{\mu} = \lambda_j^{\mu}$

<u>and</u> 2) for each i the set of r(i) q-dimensional column vectors $\underline{c}_{1i1}, \underline{c}_{1i2}, \ldots, \underline{c}_{1ir(i)}$ of the rearranged Jordan form for A^{μ} ,which are the columns of C corresponding to the <u>independent</u> states, is a linearly dependent set.

assuming that λ_i and λ_j are distinct eigenvalues of A.
If (A^{μ},B,C) is completely observable then the Ho-Kalman algorithm applied to the Page matrix will lead to a minimum realization.
It should be noted that condition F, as stated above ,will generally not hold in the noisy case, because it lays quite heavy restrictions on the positions of the poles of A.
In general one can state that the situations in which the Page matrix algorithm will not give a minimum realization appear seldom. Nevertheless in case $\lambda_i^{\mu} \approx \lambda_j^{\mu}$ and/or $\lambda_i \approx 0$, this might be a cause that the resulting solution is ill-conditioned.
A dual situation arises when the Chinese page matrix is subject of study. In stead of the observability of the system, the controllability now is the critical feature. Then in stead of the matrix C matrix B is more essential. This dual approach can be a good alternative if (A^{μ},B,C) is an nonobservable but controllable system. The Ho-Kalman algorithm applied to the Chinese page matrix then will lead to a minimum realization.

3. CONCLUSIONS

A Page matrix has been proposed as an alternative to the Hankel matrix in the realization problem.
In the deterministic case the size of the Page matrix may be chosen much smaller than the size of the Hankel matrix, which implies a significant reduction in computation.
The Page matrix is especially superior in the noisy case, due to three consecutive steps:

- The order testing: the decision concerning the dimension of the system, based on the singular values of the Page matrix, is straightforward, since all noisy data appear only once in the Page matrix. In cases of SWAYING noise, the non-relevant singular values are independent, which is not the case for the Hankel matrix (see appendix).

- The noise filtering by omitting the non-relevant singular values: there is a constant weighting factor for the Markov parameters and the total reduction equals $n/\min(h,m)$, which is optimal for a square Page matrix.

- The approximate realization: the noise filtering provides us directly with a set of unique Markov parameters in the approximated Page matrix of rank n, contrary to the situation for the Hankel matrix. This proves that we have reduced the information to the proper degree of freedom by using the Page matrix. Then the realization is straightforward, as in the deterministic case. In the approximated Hankel matrix H, however, a number of superfluous degrees of freedom is still implicitly incorporated. These are eliminated during the realization fase, be it in an uncontrolled and inadequate way.

Preliminary practical tests confirm the above theoretical expectations and we hope to present these results in a subsequent paper.

Finally, we are also optimistic about the use of the Page Matrix for the stochastic realization, where estimates of covariances are replacing the Markov parameters. Here the problem is that the uncertainties of the estimated covariances are far from independent and stationary.

ACKNOWLEDGEMENT.

The authors would like to thank prof.dr. P. Eykhoff for his comments and cooperation.

APPENDIX

The statistical behaviour of the singular values of a Hankel matrix, built up by noise corrupted Markov parameters, is a special case of the situation where all elements of a matrix have independent, additive noise, like the Page matrix. In the Hankel matrix the Markov parameters appear repeatedly and thus the noise cannot be considered as being independent. Nevertheless this independency makes it easy to study the behaviour so that we will first study the noise contaminated Page matrix.

Assume a Page matrix P with dimensions $h \cdot m$, composed of L-1 deterministic Markov parameters, all disturbed by SWAYING noise with variance σ^2 (see sect. 1.2). The deterministic Markov parameters construct a deterministic Page matrix P, so the following can be written:

$$\hat{P} = P + \Xi_p$$

where Ξ_p is the matrix containing all noise samples.

If $h < m$ we continue with $\hat{P}\,\hat{P}^T$; in the opposite case a dual version may be derived by means of $\hat{P}^T\hat{P}$.

$$\hat{P}\hat{P}^T = PP^T + \Xi_p\Xi_p^T + \Xi_p P^T + P\Xi_p^T$$

Because of the character of the noise, $E(\Xi_p)$ will be zero, and

$$E\{\Xi_p\Xi_p^T\} = m.\sigma^2 \; I_h$$

As a result:

$$E\{\hat{P}\hat{P}^T\} = PP^T + m\sigma^2 \; I_h$$

We can state that $P = U\,D\,V^T$, and also $m\sigma^2\, I_h = U(m\sigma^2 I_h)U^T$ because of the diagonal character of this matrix and the orthonormality of U.

Therefore we can write:

$$E\{\hat{P}\hat{P}^T\} = U\ (D^2 + \sigma^2 m I_h)\ U^T$$

Generally, for all possible (h,m) a description can be given for the matrix $\hat{D}$, the diagonal matrix of the singular values of $\hat{P}$:

$$E\{\hat{D}^2\} = D^2 + \sigma^2.\max(h,m).I_{\min(h,m)}$$

Although for the Hankel matrix the noise elements appear more frequently, the structure is such that it does not violate the steps used above. So the same conclusion can be made concerning the Hankel matrix. Note, however, that for the Hankel matrix the actual non-relevant singular values are highly dependent.

LITERATURE

[1] B.L. Ho and R.E. Kalman, Effective construction of linear state-variable models from input/output functions. Regelungstechnik,14(1966),545-548.

[2] H.P. Zeiger and A.J. McEwen, Approximate linear realizations of given dimension via Ho's algorithm. IEEE Trans. Autom. Control, AC-19(1974),153.

[3] J. Staar, J. Vandewalle and M. Wemans, Realization of truncated impulse response sequences with prescribed uncertainty. In: Control Science and Technology for the Progress of Society. Proc. 8th Triennial World Congress of IFAC, Kyoto,24-28 August 1981. Ed. by H.Akashi. Vol.1. Oxford: Pergamon,1982.

[4] A.A.H. Damen and A.K. Hajdasinski, Practical tests with different approximate realizations based on the singular value decomposition of the Hankel matrix. In: Identification and System Parameter Estimation. Preprints 6th IFAC Symp.,Arlington,Virginia,7-11 June 1982. Ed. by G.A.Bekey (misprinted on the cover: R.Isermann). Oxford: Pergamon, 1982. 809-814.

[5] G.A. van Zee and O.H. Bosgra, The use of realization theory in the robust identification of multivariable systems. In: Identication and System Parameter Estimation. Proc. 5th IFAC Symp.,Darmstadt, 24-28 Sept.1979. Ed. by R.Isermann. Oxford: Pergamon 1980. 477-484.

[6] A.K. Hajdasinski and A.A.H. Damen, Realization of the Markov parameter sequences using the singular value decomposition of the Hankel matrix. Department of Electrical Engineering, Eindhoven University of Technology,1979. TH-Report 79-E-95.

[7] S. Kung, A new identification and model reduction algorithm via singular value decompositions. In:Proc. 12th Asilomar Conf. on Circuits, Systems and Computers, Pacific Grove, Calif., 6-8 Nov. 1978. Ed. by Chi-Chia Hsieh. New York: IEEE,1979. 705-714.

[8] F.R.Gantmacher, The theory of matrices. Vol.1. New York: Chelsea, 1959.

[9] C.T.Chen and C.A.Desoer, A proof of controllability of Jordan form state equations. IEEE Trans. Autom. Control, AC-13, (1968),195-196.

EINDHOVEN UNIVERSITY OF TECHNOLOGY
THE NETHERLANDS
DEPARTMENT OF ELECTRICAL ENGINEERING

Reports: EUT Reports are a continuation of TH-Reports.

116) Versnel, W.
THE CIRCULAR HALL PLATE: Approximation of the geometrical correction factor for small contacts.
TH-Report 81-E-116. 1981. ISBN 90-6144-116-1

117) Fabian, K.
DESIGN AND IMPLEMENTATION OF A CENTRAL INSTRUCTION PROCESSOR WITH A MULTIMASTER BUS INTERFACE.
TH-Report 81-E-117. 1981. ISBN 90-6144-117-X

118) Wang Yen Ping
ENCODING MOVING PICTURE BY USING ADAPTIVE STRAIGHT LINE APPROXIMATION.
EUT Report 81-E-118. 1981. ISBN 90-6144-118-8

119) Heijnen, C.J.H., H.A. Jansen, J.F.G.J. Olijslagers and W. Versnel
FABRICATION OF PLANAR SEMICONDUCTOR DIODES, AN EDUCATIONAL LABORATORY EXPERIMENT.
EUT Report 81-E-119. 1981. ISBN 90-6144-119-6.

120) Piecha, J.
DESCRIPTION AND IMPLEMENTATION OF A SINGLE BOARD COMPUTER FOR INDUSTRIAL CONTROL.
EUT Report 81-E-120. 1981. ISBN 90-6144-120-X

121) Plasman, J.L.C. and C.M.M. Timmers
DIRECT MEASUREMENT OF BLOOD PRESSURE BY LIQUID-FILLED CATHETER MANOMETER SYSTEMS.
EUT Report 81-E-121. 1981. ISBN 90-6144-121-8

122) Ponomarenko, M.F.
INFORMATION THEORY AND IDENTIFICATION.
EUT Report 81-E-122. 1981. ISBN 90-6144-122-6

123) Ponomarenko, M.F.
INFORMATION MEASURES AND THEIR APPLICATIONS TO IDENTIFICATION (a bibliography).
EUT Report 81-E-123. 1981. ISBN 90-6144-123-4

124) Borghi, C.A., A. Veefkind and J.M. Wetzer
EFFECT OF RADIATION AND NON-MAXWELLIAN ELECTRON DISTRIBUTION ON RELAXATION PROCESSES IN AN ATMOSPHERIC CESIUM SEEDED ARGON PLASMA.
EUT Report 82-E-124. 1982. ISBN 90-6144-124-2

125) Saranummi, N.
DETECTION OF TRENDS IN LONG TERM RECORDINGS OF CARDIOVASCULAR SIGNALS.
EUT Report 82-E-125. 1982. ISBN 90-6144-125-0

126) Królikowski, A.
MODEL STRUCTURE SELECTION IN LINEAR SYSTEM IDENTIFICATION: Survey of methods with emphasis on the information theory approach.
EUT Report 82-E-126. 1982. ISBN 90-6144-126-9

EINDHOVEN UNIVERSITY OF TECHNOLOGY
THE NETHERLANDS
DEPARTMENT OF ELECTRICAL ENGINEERING

Eindhoven University of Technology Research Reports (ISSN 0167-9708)

(127) Damen, A.A.H., P.M.J. Van den Hof and A.K. Hajdasiński
THE PAGE MATRIX: An excellent tool for noise filtering of Markov parameters, order testing and realization.
EUT Report 82-E-127. 1982. ISBN 90-6144-127-7

(128) Nicola, V.F.
MARKOVIAN MODELS OF A TRANSACTIONAL SYSTEM SUPPORTED BY CHECKPOINTING AND RECOVERY STRATEGIES. Part 1: A model with state-dependent parameters.
EUT Report 82-E-128. 1982. ISBN 90-6144-128-5

(129) Nicola, V.F.
MARKOVIAN MODELS OF A TRANSACTIONAL SYSTEM SUPPORTED BY CHECKPOINTING AND RECOVERY STRATEGIES. Part 2: A model with a specified number of completed transactions between checkpoints.
EUT Report 82-E-129. 1982. ISBN 90-6144-129-3