ON THE ADAMJAN-AROV-KREIN APPROXIMATION, IDENTIFICATION AND BALANCED REALIZATION OF A SYSTEM

by A. BULTHEEL
Katholieke Universiteit Leuven, Belgium

P. DEWILDE
Delft University of Technology, The Netherlands

Report TW 48, January 1980.
Applied Mathematics and Programming Division
KATHOLIEKE UNIVERSITEIT LEUVEN (BELGIUM)

Paper presented at the ECCTD'80 Conference,
September 2nd-5th 1980, Warsaw.

On the Adamjan-Arov-Krein approximation, identification and balanced realization of a system

A.Bultheel[1] and P. Dewilde[2]

Abstract

In a 1971 paper, Adamjan, Arov and Krein (AAK) gave an explicit formula for the approximation of an infinite Hankel matrix by a Hankel matrix of lower rank in terms of singular vectors for the Hankel matrix. This has recently received some attention in identification and realization theory where the relation between singular value decomposition (SVD) and balanced realization is now understood. Some properties of the AAK approximation are derived and it is compared with other approximations obtained from singular value decompositions of the system Hankel map.

[1]Katholieke Universiteit Leuven, Afdeling Toegepaste Wiskunde en Programmatie Celestijnenlaan 200A. B-3030 Heverlee, Belgium.

[2]Technische Hogeschool Delft, Vakgroep Netwerktheorie, Mekelweg 4, Delft 8, The Netherlands.

1. Introduction : notation and basic results

Let $\mathbf{L}_p$ and $\mathbf{H}_p$ denote the classical Banach spaces of functions defined on the unit circle in the complex plain, and let H_h denote the Hankel operator with symbol $h \in \mathbf{L}_\infty$, i.e.

$$H_h f = (I - P)\, h\, f \quad , \forall f \in \mathbf{H}_2$$

with P the orthogonal projection of $\mathbf{L}_2$ onto $\mathbf{H}_2$.

The whole class $h+\mathbf{H}_\infty$ generates the same Hankel operator H_h since H_h only depends upon the coanalytic part of h, i.e. only upon $c_k(h)$, $k \geqslant 1$, with $c_k(h)$ the Fourier coefficients

$$c_k(h) = \frac{1}{2\pi} \int_0^{2\pi} \zeta^k\, h(\zeta)\, d\theta \quad , \zeta = e^{i\theta} \ , \ k \in \mathbb{Z} \ .$$

In the basis $\zeta^k = e^{ik\theta}$ for $\mathbf{H}_2$ and $(\mathbf{H}_2)^\perp$ we can work with the matrix representation $(c_{i+j-1}(h))_{i,j \geqslant 1}$ of H_h, acting as an operator on ℓ_2.

A function $h \in \mathbf{L}_p$ will be called analytic if it is in $\mathbf{H}_p$, i.e. if $c_k(h) = 0$ for $k = 1,2,3,\ldots$ and conjugate analytic if $c_k(h) = 0$ for $k = 0, -1, -2, \ldots$.

If f is a real function, defined on the set V, then we mean by $\delta[\min_{a \in V} f(a)] = \{F,A\}$ that the minimum F of $f(a)$ for a varying over V is obtained for $a = A$. If this minimum is unique we write $\delta![\min_{a \in V} f(a)] = \{F,A\}$.

With a Hankel operator H_h we can associate the unique conjugate analytic member in the class $h + \mathbf{H}_\infty$ which is again a symbol for H_h. We denote this by

$$S(H_h) = \sum_{k=1}^{\infty} c_k(h)\, \zeta^{-k} \ .$$

Let $\mathcal{B}$ be the set of all bounded operators from $\mathbf{H}_2$ onto $(\mathbf{H}_2)^\perp$ and $\mathcal{B}^{(k)}$ the subset of $\mathcal{B}$ of bounded operators with a dimension not exceeding k (i.e. the matrix representation has rank at most k). $\mathcal{H}$ is the set of all Hankel operators in $\mathcal{B}$ and $\mathcal{H}^{(k)}$ similarly all Hankel operators in $\mathcal{B}^{(k)}$.

Concerning the symbols of these classes of Hankel operators we introduce the notation $R^{(k)}$: the set of all conjugate analytic rational functions of degree at most k, multiplicity counted and $F^{(k)} = R^{(k)} + \mathbf{H}_\infty = \{f \in \mathbf{L}_\infty \mid f=r+h$ with $r \in R^{(k)}$ and $h \in \mathbf{H}_\infty\}$.

We will use a star to indicate the hermitean transpose of a matrix and an upper bar for the complex conjugate.

Note the following properties:

property 1.1

For $H \in \mathcal{B}$ and $k \in \mathbb{N}$ fixed we have

$$\delta[\min_{K \in \mathcal{B}^{(k)}} \|H-K\|] = \{s_k(H), \sum_{j=0}^{k-1} s_j(H)(.,\xi_j)\eta_j\}$$

where (ξ_j, η_j) are singular vectors (Schmidt pair) associated with the singular value $s_j(H)$ of H.

Proof Gohberg-Krejn [1,p. 62]. □

property 1.2 (Nehari)

$H \in \mathcal{H} \Leftrightarrow \exists f \in \mathbf{L}_\infty$ s.t. $H = H_f$.

Moreover $\|H_f\| = \inf_{h \in \mathbf{H}_\infty} \|f-h\|_\infty = \text{distance}(f, \mathbf{H}_\infty)$.

proof Nehari [2]. □

property 1.3 (Kronecker [3])

For $\gamma = \{\gamma_j\}_1^\infty$ define $\gamma_-(\zeta) = \sum_{k=1}^{\infty} \gamma_k \zeta^{-k}$, $|\zeta| = 1$, then $\gamma_-(\zeta) \in R^{(k)} \Leftrightarrow H_{\gamma_-} \in \mathcal{H}^{(k)}$

(Note that $H_{\gamma_-} = (\gamma_{i+j-1})_{i,j \geq 1}$).

proof See e.g. Gantmacher II [4,p. 207] for an algebraic and Sarason [5,p.101] for an analytic approach. □

The Adamjan-Arov-Krein result is as follows

property 1.4

If $H \in \mathcal{H}$ has singualar values $s_0(H) \geq s_1(H) \geq \ldots \geq s_{k-1}(H) > s_k(H) = \ldots = s_{k+r}(H)$ and $s_k(H)$ is an eigenvalue of $(H^*H)^{1/2}$ then there exists $L_k \in \mathcal{H}^{(k)}$ s.t.

$$\delta![\min_{L \in \mathcal{H}^{(k+r)}} \|H-L\|] = \{s_k(H), L_k\}$$

with $L_k = H - H_{\phi_k} \in H^{(k)}$ and $\phi_k(\zeta) = s_k(H) \dfrac{\eta_-(\zeta)}{\xi_+(\zeta)}$ where (ξ,η) is an arbitrary pair of singular vectors of H, associated with $s_k(H)$ and $\xi_+(\zeta) = \sum_{j=1}^{\infty} \xi_j \zeta^{j-1}$ and $\eta_-(\zeta) = \sum_{j=1}^{\infty} \eta_j \zeta^{-j}$.

Moreover

(1) $\eta_-(\zeta)/\xi_+(\zeta)$ is inner (i.e. it is analytic and in modulus 1 a.e. for $|\zeta|=1$)

(2) $\xi_+(\zeta)$ and $\eta_-(\zeta)$ have the factorizations $\xi_+(\zeta) = b_+(\zeta)a(\zeta)\phi(\zeta)$ and

$\eta_-(\zeta) = \zeta\, b_-(\zeta)a(\zeta)\phi(\zeta)$, $|\zeta| = 1$, with

$\phi(\zeta)$ outer (i.e. it is analytic and $\ell n|\phi(0)| = \dfrac{1}{2\pi}\displaystyle\int_0^{2\pi} \ell n|\phi(e^{i\theta})|d\theta$)

$b_+(\zeta)$ and $b_-(\zeta)$ inner and

$a(\zeta)$ inner and independent of the choice of the pair (ξ,η).

<u>proof</u> Adamjan, Arov, Krein [6] . □

With the Nehari and Kronecker properties, this can be translated into

<u>property 1.5</u>

If $f \in \mathbb{L}_\infty$ and $s_k(H_f)$ an eigenvalue of $(H_f^* H_f)^{1/2}$, then

$$\delta![\min_{h \in F^{(k)}} \|f-h\|_\infty] = \{s_k(H_f),\ h_k\}$$

with $h_k = f - \phi_k$ and ϕ_k as in property 1.4.

<u>proof</u> Adamjan, Arov, Krein [6] □

<u>2. Realization of linear systems</u>

Consider the linear discrete system

$$x_{k+1} = A\, x_k + B\, u_k$$
$$y_k = C\, x_k \qquad\qquad k \in \mathbb{Z}$$

with

$$x_k \in \mathbb{C}^n \ , \ u_k, y_k \in \mathbb{C} \ , \ A \in \mathbb{C}^{n\times n} \ , \ B \in \mathbb{C}^{n\times 1} \ , \ C \in \mathbb{C}^{1\times n}$$

with finite dimensional state space ($n < \infty$).

Suppose now that $\{u_k\}_{k\in\mathbb{Z}} \in \ell_2$ and that all eigenvalues of A are inside the unit disc, then there exists a unique solution for the system such that $\{x_k\}_{k\in\mathbb{Z}} \in \ell_2^n$. Taking ζ-transforms, this solution is given as $x(\zeta) = (\zeta-A)^{-1}B\, u(\zeta)$. The corresponding $\{y_k\}_{k\in\mathbb{Z}} \in \ell_2$ and the transferfunction $h(\zeta) = C(\zeta-A)^{-1}B$ is conjugate analytic and uniformly bounded. The ζ-transform of the output is given by

$$y(\zeta) = C(\zeta I-A)^{-1}\, B\, u(\zeta) \triangleq S\, u(\zeta).$$

S is known as the systemoperator.

The restricted input/output map is the Hankel operator

$H = (I-P)S\big|_{H_2} = \mathcal{O}\, \mathcal{C}$ with

$$\mathcal{O} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \end{bmatrix}$$ the observability matrix and $\mathcal{C} = [B\ AB\ A^2B\ \ldots]$

the controlability matrix. We suppose the realization minimal, so that rank H = order A = rank $\mathcal{O}$ = rank $\mathcal{C}$. It is possible to chose $W \in \mathbb{C}^{n\times n}$ non-singular such that the equivalent realization $(W^{-1}A\,W, W^{-1}B, CW)$ satisfies

$$P_0 \triangleq \mathcal{O}^*\mathcal{O} = \Sigma = \mathcal{C}\mathcal{C}^* \triangleq P_c$$

where $\Sigma = \mathrm{diag}(s_0(H), s_1(H), \ldots s_{n-1}(H))$ is the diagonal appearing in the singular value decomposition of H i.e.

$$H = U\,\Sigma V^* = \sum_{j=0}^{n-1} s_j(H)\,\eta_{j+1}\,\xi_{j+1}^* \quad , \; U = [\eta_1 \ldots \eta_n\ ],\ V = [\xi_1 \ldots \xi_n\ ]$$

and $U^*U = I = V^*V$, thus $\mathcal{O} = U\Sigma^{1/2}$ and $\mathcal{C} = \Sigma^{1/2}V^*$.

If $\mathcal{O}^*\mathcal{O} = \Sigma = \mathcal{C}\mathcal{C}^*$ then the realization is called balanced [7,8].

We consider now the problem of approximating a system of finite order $n (H \in H^{(n)})$ with a system of lower order ($\hat{H} \in H^{(k)}$ with $k < n$). This is related to the problem of approximating $r_n \triangleq S(H) \in R^{(n)}$ with a lower order rational function $r_k = S(\hat{H}) \in R^{(k)}$, $k < n$. Several algorithms have been proposed to solve this problem. See [7] for some references. We mention only two possible strategies :

Suppose for simplicity that the singular values of H are all different.

a) Partial minimal realization

One method is to find an approximation $r_k \in F^{(k)}$ as a Padé approximant of r_n, which can be computed from the Markov parameters h_j, but this does not yield a guaranteed stable approximation, so that we have to take the conjugate analytic part r' of the Padé approximant r_k. Of course $r' \in R^{(j)}$ with $j \leq k$. Consequentely the associated Hankel operator $H_{r'} \in H^{(j)}$ and thus will probably not be optimal in $H^{(k)}$ and even not in $H^{(j)}$.

b) Adamjan-Arov-Krein (AAK) approximation

With the AAK approximant, the situation is reversed. I.e. the approximant $\hat{H} = L_k$, given by property 1.4 will be an optimal approximation for H, but the associated symbol $\psi_k = r_n - \phi_k$ need not be an element of $R^{(k)}$. The theorem only claims that ψ_k is in $F^{(k)}$, but ψ_k is not rational in general. Thus also here the conjugate analytic part of ψ_k must be taken to obtain $r_k \in R^{(k)}$, but then (see unicity in property 1.5), r_k is not the optimal L_∞ approximant of r_n in $F^{(k)}$ any more, but it does solve the systemtheoretical problem.

In the next section we will explore the structure of $\psi_k(\zeta)$ somewhat further.

3. Some properties concerning the AAK approximation

property 3.1

Let $H \in H^{(n)}$ and (ξ, η) be a pair of singular vectors of H, associated with $s_k(H)$, the k-th singular value of H, then

(1) $\eta_-(\zeta) \in \mathcal{R}^{(j)}$ for some $j \leq n$

(2) $\xi_+(\zeta) = \zeta^{-1}\overline{\eta_-(1/\bar{\zeta})}\ \bar{\zeta}_k$ with ζ_k a constant of modulus 1 and

$\eta_-(\zeta)$ and $\xi_+(\zeta)$ as in property 1.4.

proof

(1) Let $H = U\Sigma V^*$ with $U,V \in \mathbb{C}^{\infty\times n}$ and $\Sigma = \mathrm{diag}(s_0(H) \ldots s_{n-1}(H))$ be the SVD of H and $\mathcal{O} = U\Sigma^{1/2}$, $\mathcal{C} = \Sigma^{1/2}V^*$ the corresponding observability and controllability matrices for the balanced realization (A,B,C), then

$$\eta_-(\zeta) = \sum_1^\infty \eta_j \zeta^{-j} = Z\,\eta = Z\ \mathcal{O}\ \sigma_k$$

with $Z = [\zeta^{-1}\zeta^{-2}\ldots]$ and $\sigma_k^* = [\underbrace{0\ldots0}_{k}\ s_k(H)^{-1/2}\underbrace{0\ldots0}_{n-k-1}]$

so that $\eta_j = C\,A^{j-1}\,\sigma_k$, $j = 1,2,3,\ldots$.

This means that (A,σ_k,C) is a realization for $\eta_-(\zeta)$, thus $\eta_-(\zeta) = C(\zeta I - A)^{-1}\,\sigma_k$.

From this it follows that the poles of $\eta_-(\zeta)$ are poles of $r_n(\zeta)$ and thus lay in the unit disc.

(2) Because H is symmetric (not necessary hermitean) it follows that in the SVD of H we have $US = \bar{V}$ with $S = \mathrm{diag}(\zeta_0,\zeta_1,\ldots,\zeta_{n-1})$,a sign matrix, (i.e. $|\zeta_j| = 1, 0 \leq j \leq n-1$) so that $\xi_j = \bar{\zeta}_k\,\bar{\eta}_j$ for $j = 1,2,3,\ldots$, which implies the second part.

□

Let us denote the reciprocal of a polynomial $P(\zeta)$ of degree m by $\hat{P}(\zeta)$ i.e. $\hat{P}(\zeta) = \zeta^m\,\overline{P(1/\bar{\zeta})}$.

Suppose now $H \in \mathcal{H}^{(n)}$ has symbol $r = S(H) = \frac{R(\zeta)}{N(\zeta)} \in \mathcal{R}^{(n)}$ with $R(\zeta)$ and $N(\zeta)$ polynomials of degree at most n. If $\eta_-(\zeta)$ is as defined in the previous property, then it is clear from its proof that $\eta_-(\zeta)$ can be written as $T(\zeta)/N(\zeta)$ with $N(\zeta)$ the denominator of r and $T(\zeta)$ a polynomial of degree

at most n. Thus the AAK approximation $\psi_k(\zeta) = r(\zeta) - s_k(H)\eta_-(\zeta)/\xi_+(\zeta)$ is given by

$$\psi_k(\zeta) = \frac{R(\zeta)}{N(\zeta)} - s_k(H)\,\frac{T(\zeta)\hat{N}(\zeta)}{N(\zeta)\hat{T}(\zeta)}\,\zeta_k = \frac{R(\zeta)\hat{T}(\zeta) - s_k(H)T(\zeta)\hat{N}(\zeta)\zeta_k}{N(\zeta)\hat{T}(\zeta)} .$$

If $\alpha \in \mathbb{C}$, then we mean by its reciprocal the complex number $1/\bar{\alpha}$. We now formulate

property 3.2

Let $H \in H^{(n)}$ and $r = S(H) \in R^{(n)}$, then the AAK approximation $\psi_k \in F^{(k)}$ of r is a rational function of degree at most n-1 with all its poles among the zeros of $\xi_+(\zeta)$ (one zero at infinity excluded), which are also the reciprocals of the zeros of $\eta_-(\zeta)$ (again one zero at infinity excluded). Suppose (A,B,C) is a balanced realization of H , a_k the k-th row of A and b_k the k-th element of B $(0 \leqslant k \leqslant n-1$!), then, if $\xi_+(0) \neq 0$ we find the mentioned zeros of $\eta_-(\zeta)$ in the spectrum of $E = A - b_k^{-1}a_kB$.

proof It follows from the above expression for $\psi_k(\zeta)$ that its poles are among the zeros of $N(\zeta)\hat{T}(\zeta)$. This means they are poles of $\eta_-(\zeta)$ or zeros of $\xi_+(\zeta)$, the latter also being reciprocals of the zeros of $\eta_-(\zeta)$ (one zero at infinity always excluded). So we only have to prove that $\psi_k(\zeta)\hat{T}(\zeta) \in \mathbf{H}_2$, i.e. has no poles inside the unit disc. Now because $H\xi = s_k(H)\eta$ by definition, also $(I-P)r(\zeta)\xi_+(\zeta) = s_k(H)\eta_-(\zeta)$ where P is as before the orthogonal projection of $\mathbf{L}_2$ onto $\mathbf{H}_2$. Thus $(I-P)(r(\zeta)\xi_+(\zeta) - s_k(H)\eta_-(\zeta)) = (I-P)(\psi_k(\zeta)\xi_+(\zeta)) = 0$ because $\eta_-(\zeta) \in (\mathbf{H}_2)^{\perp}$ and thus

$$\psi_k(\zeta)\xi_+(\zeta) = \frac{\bar{\zeta}_k R(\zeta)\hat{T}(\zeta) - s_k(H)T(\zeta)\hat{N}(\zeta)}{N(\zeta)\hat{N}(\zeta)} \in \mathbf{H}_2 .$$

Because $N(\zeta)$ has only zeros in the unit disc, $1/\hat{N}$ is analytic there. Hence, multiplying with $\hat{N}(\zeta)\,\bar{\zeta}_k$ gives $\psi_k(\zeta)\hat{T}(\zeta) \in \mathbf{H}_2$ as required. Thus $\psi_k(\zeta) = Q(\zeta)/\hat{T}(\zeta)$ with $Q(\zeta)$ a polynomial of degree strict lower than the degree of $\hat{T}(\zeta)$ (because $\psi_k(\zeta)$ is conjugate analytic) and degree $\hat{T}(\zeta)$ is at most n-1. At most k of the zeros of $T(\zeta)$ can be outside the unit disc, the others must be inside.

The second part of the property is proved as follows. By definition $\overline{\xi_+(1/\bar{\zeta})} = \sum_{j=1}^{\infty} \bar{\xi}_j\,\zeta^{-j+1}$. Now if $\xi_1 \neq 0$, then also $b_k = s_k(H)^{1/2}\,\bar{\xi}_1 \neq 0$. Under this condition is $\overline{\xi_+(1/\bar{\zeta})} = \bar{\xi}_1(1 + \sum_{j=1}^{\infty} \rho_j\,\zeta^{-j})$ with $\rho_j = \bar{\xi}_{j+1}/\bar{\xi}_1 = \chi\,A^{j-1}B$, $j = 1,2,\ldots$ with

$$\chi = \overline{\xi_1^{-1}}\,\sigma_k^{\star}\,A = b_k^{-1}\,[\underbrace{0\ \ldots\ 0}_{k}\ 1\ \underbrace{0\ \ldots\ 0}_{n-k-1}]\,A = b_k^{-1}\,a_k .$$

Thus $\bar{\xi}_1^{-1}\overline{\xi_+(1/\bar{\zeta})} = 1 + \chi(\zeta I-A)^{-1}B$. Its inverse is $\bar{\xi}_1\overline{\xi_+^{-1}(1/\bar{\zeta})} = 1 + \chi(\zeta I-E)^{-1}B$ with $E = A - B\chi = A - B\, b_k^{-1}\, a_k$. Note that the (k+1)st row of E is zero. □

The condition $\xi_1 \neq 0$ is no strong restriction because if $\xi_1=\xi_2=\ldots=\xi_{\ell-1}=0$ and $\xi_\ell \neq 0$, then $\overline{\xi_+(1/\bar{\zeta})} = \bar{\xi}_\ell\, \zeta^{\ell-1}(1+ \sum_{j=1}^{\infty} \rho_j\, \zeta^{-j})$ with $\rho_j = \bar{\xi}_{j+\ell}/\bar{\xi}_\ell = \chi A^{j-1}B$, $j=1,2,\ldots$ with $\chi = \bar{\xi}_\ell^{-1}\, \sigma_k^{\star}\, A^{\ell} = b_k^{-1}(A^{\ell})_k$. and $\bar{\xi}_\ell\overline{\xi_+^{-1}(1/\bar{\zeta})} = 1 - \chi(\zeta I-E)^{-1}B$ with $E = A-B\chi$.

4. Lower order approximation via SVD

Because the AAK approximation is expressed in terms of the singular vectors and values of an infinite Hankel matrix, it is a tempting idea that some of the algorithms computing a lower order approximation via SVD should actually find the AAK approximation. We will investigate three proposed conceptual algorithms and prove that they give the same approximation, but are definitely different from the AAK approximation.

By conceptual we mean that we formulate the algorithms as if it were possible to compute the SVD of an infinite Hankel matrix $H \in \mathcal{H}^{(n)}$. We look for an approximant in $\mathcal{H}^{(k)}$. We first take an optimal approximation in $\mathcal{B}^{(k)}$, given by property 1.1, and this in turn is approximated by a Hankel matrix, characterized by a realization triple (A_k,B_k,C_k). This method is followed by Kung [7] and Zeiger-McEwen [10] while Moore [8] directly finds the triple (A_k,B_k,C_k) from a balanced realization (A,B,C). Because the balanced realization and a SVD of an infinite Hankel matrix are essentially the same there should be an obvious relation.

The Kung algorithm is as follows. Let M^+ denote the generalized (Moore-Penrose) inverse of a matrix and $A^{\uparrow} = \urcorner A$ with $\urcorner = \begin{bmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ 0 & & \ddots & \end{bmatrix}$ and $\overleftarrow{A} = A\,\urcorner^T$.

If $H = U \Sigma V^{\star}$ is the SVD of H, then $\mathcal{O} = U\Sigma^{1/2}$ and $\mathcal{C} = \Sigma^{1/2}V$ are related to the balanced realization triple (A,B,C) of H by $A = \mathcal{O}^{+}\mathcal{O}^{\uparrow} = \Sigma^{-1/2}U^{\star}U^{\uparrow}\Sigma^{1/2} = \overleftarrow{\mathcal{C}}\,\mathcal{C}^{+} = \Sigma^{1/2}\overleftarrow{V}^{\star}V\,\Sigma^{-1/2}$ while B = the first column of $\mathcal{C}$ and C is the first row of $\mathcal{O}$. Put in the SVD of H the (n-k) least significant singular values equal to zero and use the same construction then (A_k, B_k, C_k) will be found instead of (A,B,C).

This is obviously the same as the construction of Moore, dropping the last n-k rows and columns of the triple (A,B,C) as follows :

$$(A,B,C) = \left(\left[\begin{array}{c|c} A_k & \star \\ \hline \star & \star \end{array}\right], \left[\begin{array}{c} B_k \\ \hline \star \end{array}\right], [C_k \,|\, \star] \right).$$

Zeiger and McEwen [9,10] find B_k and C_k as above but A_k is found from the equation $\mathcal{O}A_k\mathcal{C} = \mathcal{O}\overleftarrow{\mathcal{C}}$ which gives the same result as Kung's.

Moore derived his algorithm for continuous systems and in that case (A_k, B_k, C_k) is again balanced. Unfortunately this is not true for discrete time systems. However $\|A_k\|$ is guaranteed to be bounded by one, so that the approximation is stable in any case.

It should also be emphasized that the equivalence of the above mentioned algorithms is only true in a conceptual context.Where the practical implementations proposed by the authors enter , there will be a difference.

In Kung's paper it is illustrated by an example that the result he obtains with his algorithm is not far from the AAK approximation. This could suggest the errorneous idea that he computes a numerically perturbed AAK approximation. That the nearness to the AAK approximation is only due to the fact that the zeroed singular values of H were very small is illustrated by the following example.

Let $h(\zeta) = \dfrac{1}{\zeta(\zeta-\varepsilon)}$ with $\varepsilon = 0.1$. The Hankel matrix with symbol h is $H_h = (h_{i+j-1})$ with $\{h_k\}_{k \in \mathbb{N}} = \{0,1,\varepsilon,\varepsilon^2,\varepsilon^3,\ldots\}$ and $h(\zeta) = \sum_{j=1}^{\infty} h_j\, \zeta^{-j}$.

The eigenvalues of H_h are $\lambda_0 = 1.0568...$ and $\lambda_1 = -0.95580...$ and all others are zero. These eigenvalues and their corresponding eigenvectors are computed in double precision by eigenvalue decomposition of a truncated 20×20 Hankel matrix by means of the EISPACK library routines [11].
The Markov parameters h_k converge rapidly to zero because ε is small, so that a 20×20 matrix is a very good approximation of H_k.

To find the AAK approximation of degree zero, take λ_0 and the corresponding eigenvector to find $\eta_-(\zeta)$ which is

$$\eta_-(\zeta) = \frac{\alpha(\zeta-\zeta_0)}{(\zeta-\varepsilon)\zeta}\ ,\ \zeta_0 = -0.946... = -\lambda_0^{-1} \text{ and } \alpha = -0.689... \ .$$

$\xi_+(\zeta)$ then follows immediately so that we find

$$r_k(\zeta) = h(\zeta) - \lambda_0 \frac{\eta_-(\zeta)}{\xi_+(\zeta)} = \frac{\lambda_0\varepsilon}{(1-\zeta_0\zeta)} = \frac{\lambda_0^2\varepsilon}{\zeta+\lambda_0}$$

with a pole at $-\lambda_0$, i.e. outside the unit disc. This gives $\hat{H} = H_{r_k} = 0$, and this is in accordance with property 3.2 which says that $r_k(\zeta)$ is a rational function of degree n-1 = 1 and has k = 0 poles inside the unit disc.

Similarly for a degree one AAK approximation we find that $\eta_-(\zeta)$ is of the same form as above but now $\zeta_0 = 1.0462... = -\lambda_1^{-1}$ and $\alpha = 0.72463...$ and so

$$h(\zeta) - \lambda_1 \frac{\eta_-(\zeta)}{\xi_+(\zeta)} = -\frac{\lambda_1^2\varepsilon}{\zeta+\lambda_1} \quad \text{with a pole } -\lambda_1 \text{ inside the unit disc.}$$

Using the Kung algorithm to find the balanced realization we get

$$\underset{\sim}{A} \begin{bmatrix} 0.549 & 0.497 \\ -0.497 & -0.449 \end{bmatrix}\ ,\ \underset{\sim}{B} \begin{bmatrix} -0.689 \\ -0.724 \end{bmatrix}\ ,\ C = B^T \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\ .$$

Thus the degree-one approximation given by Kung's algorithm is $c_1(\zeta-a_{11})^{-1} b_1 \approx \frac{0.475}{\zeta-0.549}$ with a pole which is definitely different from the AAK pole.

Remark that the Markov parameters of the AAK or Kung approximation of degree one are not close to the given Markov parameters. This is due to the fact that the neglected singular value λ_1 is about as large in magnitude as the retained one λ_0. The closeness of the approximating Markov parameters to the given ones is exactly the criterium manipulated by Kung and bounds therefore are given in his paper.

5. Algorithms for the SVD of an infinite Hankel matrix

It will be clear from the foregoing that computing the SVD of an infinite Hankel matrix $H \in H^{(n)}$ plays an important role and this is almost the same as the computation of a balanced realization.

Kung [7] proposes to take the N×N leading submatrix of H, such that the Markov parameters h_j for $j > N$ are smaller in magnitude than a preset value and compute the SVD of this submatrix. This could lead to situations where the rank n of H can be very small, while the Markov parameters h_j are slowly decreasing, e.g. $H = H_h$ with $h(\zeta) = (\zeta-\varepsilon)^{-1}$ with $|\varepsilon| \approx 1$. Thus we need the SVD of a large N×N matrix while the first Markov parameters already contain all necessary information.

Moore's technique [8] is to transform an arbitrary realization triple (A,B,C) into a balanced one. For more details, consult [8]. This requires more work (two Lyapunov equations and three SVD's) but on matrices of order n×n, where usually $n \ll N$, a saving of computer effort after all.

It is known that the singular values are rather stable with respect to perturbation of H, but the set of singular vectors can be very sensitive, thus the computation of a balanced realization is an ill conditioned problem, and this can be more critical for the Moore method than for the Kung method.

6. Conclusion

We have introduced the Adamjan, Arov and Krein approximation result and applied it to the realization problem of a system impuls response. We have compared the AAK result with some heuristic algorithms based on balanced system realization and SVD decomposition.

It has appeared that the AAK realization is not a perturbation of the heuristic realizations. Also, a possible but not very practicable way to compute the AAK realization has been deduced. The work presented has appeared first in Dutch [12]. While working on the English version attention of the authors has been drawn to a recent report of L. Silverman and Maâmar Bettayeb [13] which contains some of the results of the paper.

References

[1] I.C. Gohberg, M.G. Krejn, "Introduction à la théorie des opérateurs linéaires non auto-adjoints dans un espace Hilbertien", Dunod, Paris, 1971.

[2] Z. Nehari, "On bounded bilinear forms", Ann. of Math. vol. 65, pp. 153-162, 1957.

[3] L. Kronecker, "Algebraïsche reduction der Schaaren bilinearer Formen", S-B Akad. Berlin, pp. 763-776, 1890.

[4] F.R. Gantmacher, "Matrix theory II", Chelsea publishing comp., New York, 1959.

[5] D. Sarason, "Function theory on the unit circle". Notes lectures at conf. at Virginia Polytechnic Inst. and State Univ., Blacksburg, Virg., 1978.

[6] V.M. Adamjan, D.Z. Arov, M.C. Krein, "Analytic properties of Schmidt pairs for a Hankel operator and the generalized Schur-Takagi problem", Mat. USSR Sbornik, vol. 15, pp. 31-73, 1971.

[7] S. Kung, "A new low-order approximation algorithm via singular value decomposition", Dept. Electr. Engng., Univ. South. Calif., Los Angeles, 1979, subm. IEEE Trans. on Autom. Control.

[8] B.C. Moore, "Singular value analysis of linear systems. Part I : external variable. Part II : controlability, observability and model reduction., System control report n° 7801-7802, April-June 1978.

[9] H.P. Zeiger, A.J. McEwen, "Approximate linear realization of given dimension via Ho's algorithm", IEEE Trans. on Autom. Contr., vol. AC-19, pp. 153, 1974.

[10] U. Jaakso, U. Nurges, "Comments on 'Approximate linear realizations of given dimensions via Ho's algorithm'", IEEE Trans. on Autom. Contr., vol. AC-20, pp. 302-303, 1975.

[11] B.T. Smith et al., "Matrix eigensystem routines - EISPACK guide", 2nd ed., Lecture notes in Comp. Sci. 6, Springer Verlag, Berlin, 1976.

[12] A.Bultheel, "Numerieke aspekten van de Adamjan-Arov-Krein benadering", Bijgevoegde stelling voor het doctoraat in de wetenschappen, K.U.Leuven, Sept. 1979.

[13] L.M.Silverman and Maâmar Bettayeb, "Optimal approximation of linear systems", report dept. of electr. engng, University of Southern California, november 1979.