May 1989 LIDS-P-1870

Pade' Approximants to Matrix Stieltjes Series: Convergence and Related Properties

S. Basu
Department of Electrical Engineering and Computer Science
Stevens Institute of Technology
Hoboken, NJ 07030

This work was done while the author was a visitor at the Laboratory for Information and Decision Systems, M.I.T., Cambridge, MA 02139.

Abstract:

Following earlier work on Pade′ approximants to matrix Stieltjes series and their network theoretic relevance it is shown that certain paradiagonal sequences of matrix Pade′ approximants to the series under consideration always converge. Interpretation of this result in terms of representation of impedances of RC distributed multiport networks are given. Matricial generalizations of the classical Hamburger and Stieltjes moment problems are discussed in this context. Matrix polynomials of the second kind orthogonal on the real line, which fall out as numerators of the matrix Pade′ approximants of certain orders are singled out and their properties are studied.

1. Introduction

Consider a formal power series as in (1.1) where

$$T(s) = \sum_{k=0}^{\infty} T_k\, s^k \tag{1.1}$$

each T_k is a real symmetric matrix of size (pxp). The rational matrix $Q_L(s)P_M^{-1}(s)$ (or $P_M^{-1}(s)Q_L(s)$), where $Q_L(s)$ and $P_M(s)$ are (pxp) polynomial matrices of respective formal degrees L and M[1] is said to be a right matrix Pade' approximant (or left matrix Pade' approximant) to T(s) if the first (L+M+1) terms of the Maclaurin's series expansion of $Q_L(s)P_M^{-1}(s)$ (or $P_M^{-1}(s)Q_L(s)$) matches with those of T(s) in (1.1). In addition, the formal power series T(s) in (1.1) is said to be a matrix Stieltjes series [1] if for each n the block Hankel matrices $H_n(T)$ and $H'_n(T)$ as given in (1.2) are positive definite and negative definite respectively.

$$H_n(T) = \begin{bmatrix} T_0 & T_1 \cdots T_n \\ T_1 & T_2 \\ \cdot \\ \cdot \\ \cdot \\ T_n & & T_{2n} \end{bmatrix} \quad ; \quad H'_n(T) = \begin{bmatrix} T_1 & T_2 \cdots T_n \\ T_2 & T_3 \\ \cdot \\ \cdot \\ \cdot \\ T_n & & T_{2n-1} \end{bmatrix} \tag{1.2a,b}$$

Note that for matrix Stieltjes series the right and left matrix Pade' approximants uniquely exist and are necessarily identical [1],[3]. Thus, the term matrix Pade' approximant (MPA) of order [L/M] will henceforth be used to denote $[L/M](s)=Q_L(s)P_M^{-1}(s)=P_M^{-1}(s)Q_L(s)$. The fact that the paradiagonal sequences of MPA's of order [m-1/m] and [m-1/m-1] for m = 1,2,...etc. to a matrix Stieltjes series can be identified as the impedances or admittances of multiport electrical networks containing two types of elements (e.g., RC or RL) has been established in [1] via utilization of recently developed tools of matrix continued

---

[1] The formal degree of a polynomial matrix is defined as the largest degree of its polynomial entries.

fraction expansion and the Cauchy index of a rational matrix. Also, the Pade' approximation problem can be cast in terms of the partial realization problem as occurring in linear system theory [13], [19]. The positive definiteness of the block Hankel matrices $H_n(T)$ for all n then imply, in particular, that every point in the partial realization data is a 'jump point' with jump size equal to 1 [13]. Thus, as has been shown via the tools of matrix continued fraction expansion [1],[15] as well as the Cauchy index of a rational matrix [1], that the positive definiteness of $H_n(T)$ and $H_n'(T)$ can be viewed as the conditions which the partial realization data needs to satisfy so that the realized transfer function matrix is an impedance or admittance of an RC multiport (similar formulations for RL or LC impedances or admittances are also possible).

The question of convergence of the sequences [m-1/m] or [m-1/m-1], m = 1,2,....etc. of MPA's, when T(s) is a matrix Stieltjes series, however, has not been addressed in the literature. Although a discussion of this issue in the scalar (i.e., p=1) case is available in [3], system theoretic interpretation of the results are not readily available. In the present paper it is shown by exploiting the network theoretic interpretations developed in [1] that the sequences of [m-1/m] and [m-1/m-1], m = 1,2,...etc. MPA's to a matrix Stieltjes series do indeed converge uniformly in an open (bounded) region of the complex s-plane excluding the negative real axis.

Furthermore, since the sequences of [m-1/m] and [m-1/m-1] MPA's can be viewed as the successive convergents of certain special types of matrix continued fractions [1], the convergence of the paradiagonal sequences of MPA's can also be interpreted as the convergence property of the related matrix continued fraction expansions. Since the continued fraction expansion just mentioned is, in fact, associated with a ladder realizable RC multiport (cf. fig. 1 for p=1),we essentially have the result that the sequence of RC multiport ladder impedances (or admittances) so derived from a matrix Stieltjes series is always convergent (even if the formal power series T(s) in (1.1) is not). Thus, this latter result can be interpreted in terms of the important fact that the matrix Stieltjes series in (1.1), which may not necessarily converge (cf. [3] for examples in p=1 case), can be meaningfully used to represent a nonrational impedance or admittance matrix associated with a multiport RC distributed transmission line [17].

It is next shown that, as in the scalar case the positive definiteness of $H_n(T)$ in (1.2a) for all n guarantees the existence of a bounded non-decreasing real symmetric matrix valued measure $\sigma(x)$, $-\infty < x < \infty$, such that each of the T_k's in (1.2) can be viewed as the k-th order moment associated with $\sigma(x)$. This result provides a direct matricial generalization of the classical Hamburger moment problem [6],[10]. If, in addition to (1.2a), the negative definiteness of $H_n'(T)$ in (1.2b) is also imposed then the support of $\sigma(x)$ is shown to be restricted to the semi-infinite interval $0 < x < \infty$, thus providing a solution to the matrix version of classical Stieltjes moment problem [6],[10].

Furthermore, if the power series T(s) in (1.1) is assumed to converge in a disc of radius R then we show that the sequences of MPA's [m-1/m](s) and [m-1/m-1](s) indeed converge to T(s). An integral representation of T(s) in terms of the matrix valued measure $\sigma(x)$, which coincides with the Cauer's representation [19] of RC (multiport) impedances, or in the scalar case with the closely related class of classical Stieltjes functions [3], is also derived in this context.

On the otherhand, the 'denominator' polynomial matrices associated with MPA's of order [m-1/m] and [m-1/m-1] have been shown in [1] to form sequences of polynomial matrices orthogonal on the real line. Various properties of these polynomials such as the three term recurrence relation followed by them, properties of location their zeros and their relationship to the matricial Gauss quadrature formula were also derived in [1]. In the present paper it is shown that while in [1] the orthogonality of the matrix polynomials was viewed in terms of certain vector space represention s, by using the measure $\sigma(x)$ the orthogonality relationship can be seen more transparently in terms of an inner product in standard form.

More importantly, this approach also establishes the interesting fact that the sequences of 'numerator' polynomial matrices of the [m-1/m](s) and [m-1/m-1](s) sequences of MPA's also satisfy certain orthogonality properties similar to those satisfied by the 'denominators' of the paradiagonal sequences of MPA's in question. From a system theoretic standpoint this result is to be expected in view of the fact that the property of RC, RL or LC impedance (or admittance) realizability remain invariant under the operation of inversion of the rational matrix concerned. The sequence of matrix orthogonal polynomials corresponding to the 'numerator' sequences of MPA's are thus also found to provide matricial generalization of the orthogonal polynomials of the second kind discussed in the classical literature [6],[9].

In the rest of this section related previous research on specific aspects of the problem considered in the present paper will be briefly reviewed and comparisions to our approach to the problem will be made. The study of matrix Pade' approximants, their relationships to continued fractions, various moment problems and issues of convergence were initiated in [23],[24]. Both convergence of sequences of Pade' type approximants to Stieltjes series [22] as well as the related moment problems [25] have been discussed in the mathematical literature in an (infinite dimensional) operator theoretic setting by assuming that the T_k's in (1.1) are not just matrices but infinite dimensional operators in Hilbert space. Convergence of Pade' approximants to a formal power series of the matrix Stieltjes type has been previously considered in [21].

The present paper deviates from those mentioned above in the following respects. First, our prooofs are simpler, more degant and makes use of elementary tools from linear algebra and complex function theory. This is so because it makes full use of the finite dimensional (i.e., matrix) nature of the problem considered. In fact, although the major results on the moment problems in [25] is known to be in error [26] in the infinite dimensional case, a correct elementary discussion for the finite dimensional problem is not known.

The convergence proof of [21] starts from a slightly different (albeit equivalent) definition of matrix Stieltjes series, where T_k's are assumed to be the moments associated with a nondecreasing symmetric matrix valued measure at the very outset. This already amounts to assuming a solution to the corresponding moment problem referred to above, which is worked out in the present framework in section 3 of our paper. Furthermore, although the final results in [21] hold only for the finite dimensional ease, their proofs hinge on powerful operator theoretic results (e.g., Naimakar's theorem linking method of moments for selfadjoint operators in Hilbert space). While our definition of matrix Stieltjes series is via the algebraic constraints imposed on the sign definiteness of $H_n(T)$ and $H'_n(T)$, our proof is more direct, elementary and does not make use of a solution to the moment problem at all.

Finally, the most contrasting aspect of the present contribution is that our discussions including the details of proofs are guided throughout by system theoretic intuition -- an approach not adopted by earlier authors in the area.

2. Convergence Proof of Sequence of MPA's to Matrix Stieltjes Series:

The major content of this section is the proof of the fact that the $[m-1/m](s)$ and $[m-1/m-1](s)$ sequences of MPA's converge uniformly to an analytic function in the domain $D(\Delta)$ where $D(\Delta)$ is any bounded domain of the complex plane at least at a distance Δ away from the negative real axis: $-D \leq \text{Res} \leq 0$, $\text{Ims} = 0$ (cf. fig 3). The strategy of our proof is to first show that the required convergence is attained for all fixed real positive values of s. This is achieved by establishing certain monotonicity and boundedness properties of the approximants which fall out as consequences of RC Realizability of $[m-1/m](s)$ and $[m-1/m-1](s)$ as shown in [1]. Uniform convergence in $D(\Delta)$ is then proved by essentially exploiting standard arguments on convergence continuation [5]. A mathematically equivalent procedure has been pursued in [3] for the scalar (p=1) case without the use of network theoretic arguments.

The following notations will be used in the rest of the paper. If A is a real symmetric positive definite matrix then we will write $A > 0$. Also, the notations $A > B$ and $A \geq B$ will be taken to mean that the real symmetric matrix $A-B$ is positive definite or non-negative definite respectively. Obvious variations of this notation with the symbols $>$ and $\geq$ replaced by $<$ and $\leq$ will also be used.

We first need the following theorem:

Theorem 2.1: The sequences of $[m-1/m](s)$ and $[m-1/m-1](s)$ approximants to a matrix Stieltjes series each respectively form an increasing and decreasing sequence of symmetric matrix fraction descriptions on the positive real axis, i.e., for all m = 1,2,3,...etc. and for all s with $\text{Res} > 0$ and $\text{Ims} = 0$ we have

$$[m/m+1](s)-[m-1/m](s)>0, \quad [m/m](s)-[m-1/m-1](s)<0 \qquad (2.1),(2.1')$$

Proof: The proof relies on the result [1, p.211] that $[m-1/m]$ and $[m-1/m-1]$ approximants to the matrix Stieltjes series (1.1) can be obtained by truncating the matrix continued fraction expansion

$$T(s) = [B_1 + [\frac{1}{s} B_2 + \ldots [\frac{1}{s^{\lambda_k}} B_k + \frac{1}{s^{\lambda_{k+1}}} T_k(s)]^{-1} \;]^{-1} \;]^{-1} \qquad (2.2)$$

where $\lambda_k = 0$ for k odd and $\lambda_k = 1$ for k even, the B_i, i = 1,2,.... are constant real symmetric positive definite matrices and $T_k(s)$ is a matrix Stieltjes series. In particular, it is shown in [1] that for m = 1,2,....etc. (2.3) and (2.3′) hold true.

$$[m-1/m](s) = [B_1 + [\frac{1}{s} B_2 + \ldots + [\frac{1}{s} B_{2m}]^{-1} \;]^{-1} \;]^{-1} \qquad (2.3)$$

$$[m-1/m-1](s) = [B_1 + [\frac{1}{s} B_2 + \ldots + [B_{2m-1}]^{-1} \;]^{-1} \;]^{-1} \qquad (2.3')$$

We shall prove (2.1) only, the proof for (2.1′) being analogous. Note first that due to (2.3), the approximant of order [m/m+1] can be written as in (2.4).

$$[m/m+1](s) = [B_1 + [\frac{1}{s} B_2 + \ldots + [\frac{1}{s} B_{2m} + [B_{2m+1} + [\frac{1}{s} B_{2m+2}]^{-1} \;]^{-1} \;]^{-1} \;]^{-1} \;]^{-1} \qquad (2.4)$$

Obviously, for Res > 0 and Ims = 0 we have $\frac{1}{s} B_i > 0$ for all i. Since the sum as well as inverse of real symmetric positive definite matrices is also real symmetric positive definite, we have:

$[B_{2m+1}+[\frac{1}{s}B_{2m+2}]^{-1}]^{-1} > 0$; consequently, $\frac{1}{s} B_{2m}+[B_{2m+1}+[\frac{1}{s}B_{2m+2}]^{-1}]^{-1} > \frac{1}{s} B_{2m}$ for Res > 0 and Ims = 0. By using the result that if A and B are two real symmetric positive definite matrices such that A > B, then $A^{-1} < B^{-1}$ ([2], p.86) it then follows that

$$[\frac{1}{s} B_{2m} + [B_{2m+1}+[\frac{1}{s}B_{2m+2}]^{-1}]^{-1}]^{-1} < [\frac{1}{s} B_{2m}]^{-1} \qquad (2.5)$$

Repeating the process of adding the matrices $\frac{1}{s^{\lambda_i}} B_i$ and subsequently considering the inverses of the resulting matrices in the left and right hand sides of (2.5) for i = 2m-1, 2m-2,...., 1, where $\lambda_i = 0$ when i is odd and $\lambda_i = 1$ when i is even, it follows from (2.3) and (2.4) that [m/m+1](s) > [m-1/m](s) for Res > 0, Ims = 0. ▒

The physical implication of the above theorem is obviously clear in electrical network theoretic terms, when the [m-1/m](s) and

$[m-1/m-1](s)$ approximants to a Stieltjes series are interpreted as being the input impedance of RC ladder network, as depicted for the scalar case $p = 1$, in fig. 1a and fig. 1b respectively. The monotonicity property of the sequences of approximants then trivially follow from the fact that for all real and positive values of s, the input impedances can be computed by replacing the capacitors by positive resistances.

The norm $||x||$ of a vector x will be defined as the well known Euclidean norm, whereas the norm $||.||$ of a matrix A will be defined as the spectral norm $||A|| \triangleq \max [||Ax||; ||x||=1]$. We recall the following properties of the spectral norm $||.||$ of a matrix A.

<u>Property 2.1[4]</u>: $||A||$ is equal to the largest singular value of A. In particular, if A is real then $||A|| = \sqrt{(\lambda_m(A^tA))}$, where A^t is the transpose of the matrix A, and $\lambda_m(A^tA)$ denotes the largest eigenvalue of A^tA.

<u>Property 2.2:</u> For any real symmetric matrix A, $||A|| = |\lambda_m(A)|$, where $\lambda_m(A)$ is the eigenvalue of A having largest absolute value. Thus, in particular, $||\alpha A|| = |\alpha|.||A||$, where α is any real number.

<u>Proof</u>: Follows from the fact that $\lambda(A^tA) = \lambda(A^2) = \lambda^2(A)$ where $\lambda(A)$ is an eigenvalue (necessarily real) of A. ▒

<u>Property 2.3:</u> If A, B, C are real symmetric positive (non-negative) definite matrices such that $A = B+C$ then $||A|| > ||B||$ ($||A|| \geq ||B||$).

<u>Proof:</u> Since C is positive (non-negative) definite from Courant-Fisher min-max theorem (e.g., [2] p. 73) it follows that $\lambda_m(A) = \lambda_m(B+C) > \lambda_m(B)$ (or $\lambda_m(B+C) \geq \lambda_m(B)$). The result then follows from Property 2.2 via the observation that eigenvalues of A and B are positive (nonnegative). ▒

<u>Property 2.4:</u> If A and B are two real symmetric positive (non-negative) definite matrices, and a and b are two real numbers such that $0 < a < 1$ and $0 < b < 1$ then $||aA + bB|| < ||A + B||$.

Proof: Since $0 < (1-a) < 1$ and $0 < (1-b) < 0$, the matrices $(1-a)A$ and $(1-b)B$, and thus $\{(1-a)A+(1-b)B\}$ are real symmetric positive (non-negative) definite. Thus, $||A + B|| = ||(aA + bB) + \{(1-a)A + (1-b)B\}|| > ||aA + bB||$ (or $\geq ||aA + bB||$ correspondingly). The last step follows by the use of Property 2.3 above. ▩

Corollary 2.1.1: The MPA's to a matrix Stieltjes series satisfy: $||[m/m+1](s)|| > ||[m-1/m](s)||$, for $m = 1,2,...$ and $||[m/m](s)|| < ||[m-1/m-1](s)||$ for $m = 1,2,3,...$ for all real positive values of s.

Proof: We first note that the MPA's to the matrix series of Stieltjes (1.1) are necessarily symmetric rational matrices. To substantiate this, note that if $Q_L(s)P_M^{-1}(s)$ is a right MPA of order [L/M] to the series T(s), then $P_M^{-t}(s)Q_L^t(s)$ [2] is also a left MPA of order [L/M] to the matrix series $T^t(s) = T(s)$. The last equality follows from the fact that in (1.1) $T_i^t = T_i$ for all $i = 0,1,2,...$etc. However, this proves that both right and left approximants of order [L/M] to the series T(s) exist, and hence they must be equal [3], i.e., $Q_L(s)P_M^{-1}(s) = P_M^{-t}(s)Q_L^t(s)$. Thus the approximant of order [L/M] is symmetric. (Alternatively, this result also follows from the representations (2.3) and (2.3') of the sequences [m-1/m](s), m = 1,2,...etc. and [m-1/m-1](s), m = 1,2,...etc. of the approximants.) Also, it follows from Theorem 2.1 that if $[m/m+1](s) - [m-1/m](s) = P_{m1}(s)$ and $[m-1/m-1](s) - [m/m](s) = P_{m2}(s)$, then for all real positive values of s and for all m = 1,2,...etc., $P_{m1}(s)$ and $P_{m2}(s)$ are real symmetric positive definite matrices. Since due to (2.3) and (2.4) the MPA's in the last two equalities are themselves real symmetric positive definite for all real and positive values of s, the required result follows from Property 2.3 of spectral norm.

▩

Next, we consider the matrix continued fraction expansions of the MPA's [m-1/m](s) and [m-1/m-1](s) to the matrix Stieltjes series T(s). Since the MPA's just mentioned are known to be the impedance matrices of electrical networks consisting of positive resistors and capacitors

[2] superscript t denotes the transpose of a real matrix

only, the continued fraction expansions, due to results discussed in [1], can be expressed as the matrix partial fraction expansion:

$$[m-1/m](s) = \sum_{\nu=1}^{r} \frac{1}{1+\gamma_\nu s} A_\nu \qquad (2.6)$$

$$[m-1/m-1](s) = A_0' + \sum_{\nu=1}^{r'} \frac{1}{1+\gamma_\nu' s} A_\nu' \qquad (2.6')$$

where in (2.6) and (2.6'), the constant matrices A_ν, A_ν' are all real symmetric non-negative definite, and the constants γ_ν and γ_ν' are real and positive. Note that in the scalar case (i.e., if p=1) (2.6) or (2.6') can thus be interpreted as the input impedance of a circuit as shown in fig. 2, where $A_i=R_i$ and $\gamma_i=R_iC_i$. Similar interpretations are possible for p>1. Expanding the right hand side of (2.6) and (2.6') in a power series around s=0, and recalling the fact that the first 2m terms of the expansion for [m-1/m](s) and the first (2m-1) terms in the expansion for [m-1/m-1](s) must be identical with the given power series T(s) in (1.1) it respectively follows that:

$$T_k = \sum_{\nu=1}^{r} (-\gamma_\nu)^k A_\nu \text{ for } k = 0,1,\ldots 2m-1 \qquad (2.7)$$

$$T_0 = \sum_{\nu=0}^{r'} A_\nu'; \; T_k = \sum_{\nu=1}^{r'} (-\gamma_\nu)^k A_\nu' \text{ for } k = 1,2,\ldots 2(m-1) \qquad (2.7'a,b)$$

We next state the following theorem, which follows from the representations (2.6) and (2.6') associated with the MPA's of respective orders [m-1/m](s) and [m-1/m-1](s) to a matrix Stieltjes series.

If only $H_n(T)$ (but not $-H_n'(T)$) is positive definite for all n then representations (2.6) and (2.6'), and thus (2.7) and (2.7'), still hold true. However, γ_ν's may then assume positive as well as negative real values.

Theorem 2.2: For all s with Res > 0 and Ims = 0, the sequences of norms $||[m-1/m](s)||$ and $||[m-1/m-1](s)||$, m = 1,2,...etc. of MPA's to a matrix Stieltjes series each posseses an uniform upper bound.

Proof: Since $\gamma_\nu > 0$ and $\gamma'_\nu > 0$, we have $|(1 + \gamma_\nu s)^{-1}| < 1$ and $|(1 + \gamma'_\nu s)^{-1}| < 1$ for all real and positive values of s. Therefore, it follows from (2.6), (2.6′), Property 2.4 of spectral norm ||.|| and the triangle inequality for the spectral norm [4] that (2.8) and (2.8′) respectively hold true for all real and positive values of s, and for each m = 1,2,...etc.

$$||[m-1/m](s)|| \leq ||\sum_{\nu=1}^{r} A_\nu|| \qquad (2.8)$$

$$||[m-1/m-1](s)|| \leq ||\sum_{\nu=0}^{r'} A'_\nu|| \qquad (2.8')$$

Furthermore, by considering (2.7) and (2.7′) with k = 0 it immediately follows from (2.8) and (2.8′) that for Res > 0, Ims = 0 and each m = 1,2,..., etc. (2.9) and (2.9′) in the following hold:

$$||[m-1/m](s)|| \leq ||T_0|| \qquad (2.9)$$

$$||[m-1/m-1](s)|| \leq ||T_0|| \qquad (2.9')$$

The fact that the sequences ||[m-1/m](s)|| and ||[m-1/m-1](s)|| for m=1,2,...etc. each has uniform upper bounds for real positive s has, therefore, been established. ▒

In the scalar case i.e., if p=1 Theorem 2.2 admits of an obvious physical intepretation when [m-1/m](s) or [m-1/m-1](s) is viewed as an impedance of the RC circuit as in fig. 2 (for p=1), or equivalently, as in (2.6) or (2.6′). The uniform upper bound on the approximants is then provided by the sum of all resistors in the network. Similar interpretations are also possible when p>1.

Theorem 2.3: For all real positive value of s the sequences [m-1/m](s) m= 1,2,...etc, as well as [m-1/m-1](s), m = 1,2,...etc., of MPA's to a matrix Stieltjes series converge pointwise.

Proof: Since a (strictly) monotone (increasing or decreasing), bounded sequence of real numbers necessarily converge, it follows from

Corollary 2.1.1 and Theorem 2.2 that the sequences $||[m-1/m](s)||$, $m = 1,2,\ldots$etc., and $||[m-1/m-1](s)||$, $m = 1,2,\ldots$etc. are convergent for all real positive values of s. The required result then follows by noting [4] that convergence of the sequence of norms $||[m-1/m](s)||$, $m = 1,2,\ldots$etc. is a sufficient condition for the matrix sequence $[m-1/m](s)$, $m = 1,2,\ldots$etc. to converge. Similar arguments hold for $[m-1/m-1](s)$, $m = 1,2,\ldots$etc. ▩

Corollary 2.3.1: For any i, j with $1 \leq i,j \leq p$ and for any real positive value of s the sequences $[m-1/m]_{ij}(s)$ and $[m-1/m-1]_{ij}(s)$ $m=1,2,\ldots$etc of ij-th entries of MPA's of respective orders [m/m-1] and [m-1/m-1] to a matrix Stieltjes series converge pointwise.

Our next objective is to enlarge the domain of convergence of the sequence of approximants under consideration to a region $D(\Delta)$ larger than the positive real axis, where $D(\Delta)$ is any bounded region of the complex plane, which is at least at a distance Δ away from the negative real axis. The region $D(\Delta)$ is shown in fig. 3.

We first need the following lemma:

Lemma 2.4: Assuming T(s) to be a matrix Stieltjes series, if $A_\nu^{(ij)}$, $A_\nu'^{(ij)}$ are the respective ij-th elements of the matrices A_ν, A'_ν and γ_ν, γ'_ν are positive numbers as appearing in (2.6) and (2.6′), then the following inequalities hold true:

$$\sum_{\nu=1}^{r} |A_\nu^{(ij)}| \leq \sqrt{(T_0^{(ii)}T_0^{(jj)})} \; ; \quad \sum_{\nu=1}^{r} |A_\nu^{(ij)}|\gamma_\nu \leq \sqrt{(T_1^{(ii)}T_1^{(jj)})} \qquad (2.10a,b)$$

$$\sum_{\nu=0}^{r'} |A_\nu'^{(ij)}| \leq \sqrt{(T_0^{(ii)}T_0^{(jj)})} \; ; \quad \sum_{\nu=0}^{r'} |A_\nu'^{(ij)}|\gamma_\nu \leq \sqrt{(T_1^{(ii)}T_1^{(jj)})} \qquad (2.10'a,b)$$

Proof: Only proofs for (2.10a) and (2.10b) will be given. Analogous proofs hold for (2.10′a) and (2.10′b). Consider the case i = j first. Since the A_ν are real symmetric non-negative definite matrices the diagonal elements $A_\nu^{(ii)}$ are necessarily non-negative. Furthermore, considering the ii-th elements of the matrices in (2.7) with k = 0, and k = 1, one obtains respectively (2.11a) and (2.11b).

$$\sum_{\nu=1}^{r} |A_\nu^{(ii)}| = \sum_{\nu=1}^{r} A_\nu^{(ii)} = T_0^{(ii)}; \quad \sum_{\nu=1}^{r} |A_\nu^{(ii)}|\gamma_\nu = \sum_{\nu=1}^{r} A_\nu^{(ii)}\gamma_\nu = |T_1^{(ii)}| \tag{2.11a,b}$$

Next, when $i \neq j$, the (2x2) principal minor of A_ν obtained by considering the i-th row and j-th column of A_ν are also non-negative definite. Thus, it follows that $|A_\nu^{(ij)}| \leq \sqrt{(A_\nu^{(ii)}A_\nu^{(jj)})}$. The last inequality, along with an application of the well known Cauchy-Schwartz inequality, yields (2.12a) and (2.12b):

$$\sum_{\nu=1}^{r} |A_\nu^{(ij)}| \leq \sum_{\nu=1}^{r} \sqrt{(A_\nu^{(ii)}A_\nu^{(jj)})} \leq \sqrt{[(\sum_{\nu=1}^{r} A_\nu^{(ii)})(\sum_{\nu=1}^{r} A_\nu^{(jj)})]} \tag{2.12a}$$

$$\sum_{\nu=1}^{r} |A_\nu^{(ij)}|\gamma_\nu \leq \sum_{\nu=1}^{r} \sqrt{(A_\nu^{(ii)}A_\nu^{(jj)}\gamma_\nu^2)} \leq \sqrt{[(\sum_{\nu=1}^{r} A_\nu^{(ii)}\gamma_\nu)(\sum_{\nu=1}^{r} A_\nu^{(jj)}\gamma_\nu)]} \tag{2.12b}$$

Using (2.7) with $k = 0$ and $k = 1$ it easily follows that the right hand sides of (2.12a) and (2.12b) are respectively equal to $\sqrt{(T_0^{(ii)}T_0^{(jj)})}$ and $\sqrt{(T_1^{(ii)}T_1^{(jj)})}$. The inequalities (2.10a) and (2.10b) are thus established. ▒

<u>Theorem 2.5:</u> Each element of the sequences of MPA's [m-1/m](s) and [m-1/m-1](s) for m = 1,2,...etc. to a matrix Stieltjes series is uniformly bounded in $D(\Delta)$.

<u>Proof:</u> We first prove the result for the region $\{s; \operatorname{Re} s \geq 0, s \in D(\Delta)\}$. Consider the ij-th element $[m-1/m]_{ij}(s)$ and $[m-1/m-1]_{ij}(s)$ of [m-1/m](s) and [m-1/m-1](s) respectively. From (2.6) and (2.6') it respectively follows that:

$$[m-1/m]_{ij}(s) = \sum_{\nu=1}^{r} \frac{1}{1+\gamma_\nu s} A_\nu^{(ij)} \tag{2.13}$$

$$[m-1/m-1]_{ij}(s) = A_0^{(ij)} + \sum_{\nu=1}^{r'} \frac{1}{1+\gamma'_\nu s} A'^{(ij)}_\nu \tag{2.13'}$$

Clearly, since γ_ν, $\gamma'_\nu > 0$, we have for Res $\geq$ 0 and arbitrary Ims that:

$|1+\gamma_\nu s| \geq 1$, $|1+\gamma'_\nu s| \geq 1$ for all ν. By making use of the last inequalities along with (2.13) and (2.13′), (2.14) and (2.14′) respectively are obtained via the use of triangle inequality.

$$|[m-1/m]_{ij}(s)| \leq \sum_{\nu=1}^{r} |\frac{1}{1+\gamma_\nu s}||A_\nu^{(ij)}| \leq \sum_{\nu=1}^{r} |A_\nu^{(ij)}| \qquad (2.14)$$

$$|[m-1/m-1]_{ij}(s)| \leq |A_0^{(ij)}| + \sum_{\nu=1}^{r'} \frac{1}{|1+\gamma'_\nu s|} |A_\nu^{(ij)}| = \sum_{\nu=0}^{r'} |A'^{(ij)}_\nu| \qquad (2.14')$$

It then follows from (2.14) and (2.14′) via the use of (2.10a) and (2.10′a) in Lemma 2.4 that for Res $\geq$ 0 and for each m = 1,2,...etc. $|[m-1/m]_{ij}(s)| \leq \sqrt{(T_0^{(ii)}T_0^{(jj)})}$ and $|[m-1/m-1]_{ij}(s)| \leq \sqrt{(T_0^{(ii)}T_0^{(jj)})}$. The theorem has been thus proved, in particular, for all s in {s; Res $\geq$ 0, $s \varepsilon D(\Delta)$}.

In the following we consider values of s in the region {s; Res < 0, $s \varepsilon D(\Delta)$}. Note first that since the identity $|s|^2|1+\gamma s|^2 - |Ims|^2 = (\gamma|s|^2+Res)^2$ holds for any real γ, we have $|1+\gamma_\nu s| \geq |Ims|/|s|<1$ and $|1+\gamma'_\nu s| \geq |Ims|/|s|<1$ for each ν. Consequently, (2.15) and (2.15′) respectively follow from (2.6) and (2.6′) for each i,j = 1,2,...p.

$$|[m-1/m]_{ij}(s)| \leq \sum_{\nu=1}^{r} \frac{1}{|1+\gamma_\nu s|} |A^{(ij)}| \leq (|s| \sum_{\nu=1}^{r} |A_\nu^{(ij)}|)/|Ims| \qquad (2.15)$$

$$|[m-1/m-1]_{ij}(s)| \leq |A_0^{(ij)}| + \sum_{\nu=1}^{r'} \frac{1}{|1+\gamma'_\nu s|} |A'^{(ij)}_\nu|$$

$$\leq (|s| \sum_{\nu=0}^{r'} |A'^{(ij)}_\nu|)/|Ims| \qquad (2.15')$$

Again invoking (2.10a) and (2.10′a) of Lemma 2.4 along with (2.15) and (2.15′) respectively it follows that for all s with Res < 0 we have that

$$|[m-1/m]_{ij}(s)| \leq |s|\sqrt{(T_0^{(ii)}T_0^{(ij)})}/|Ims| \qquad (2.16)$$

$$|[m-1/m-1]_{ij}(s)| \leq |s|\sqrt{(T_0^{(ii)}T_0^{(ij)})}/|Ims| \qquad (2.16')$$

If $R_M < \infty$ is the radius of a circle, which completely encloses $D(\Delta)$ in the complex plane then (2.16) and (2.16′) establishes an uniform upper bound of $(R_M\sqrt{(T_0^{(ii)}T_0^{(ij)})}/\Delta)$ for the sequences $|[m-1/m]_{ij}(s)|$ and $|[m-1/m-1]_{ij}(s)|$, m = 1,2,...etc. in $\{s;\ Res < 0,\ s\varepsilon D(\Delta)\}$. If $M = Max(R_M/\Delta,\ 1)$ then due to the result proved in the last paragraph, $M\sqrt{(T_0^{(ii)}T_0^{(ij)})}$ serves as an uniform upper bound on each of the sequences $|[m-1/m]_{ij}(s)|$ and $|[m-1/m-1]_{ij}(s)|$, m = 1,2,...etc. in $D(\Delta)$. The theorem is thus proved. ▒

We are now in a position to prove the convergence of the sequences of rational matrices $[m-1/m](s)$ and $[m-1/m-1](s)$, m = 1,2,...etc., by using standard techniques from complex function theory [5]. The result is summarized in the following theorem.

<u>Theorem 2.6:</u> The sequences of MPA's of order $[m-1/m](s)$, m = 1,2,...etc. and $[m-1/m-1](s)$, m = 1,2,...etc. to a matrix Stieltjes series converge uniformly in the region $D(\Delta)$ of the complex plane. Furthermore, the matrix valued functions $G(s)$ and $G'(s)$ to which the two sequences respectively converge are both real symmetric (i.e., $\bar{G}(s) = G(\bar{s})$, $\bar{G}'(s) = G'(\bar{s})$)[3] and analytic in $D(\Delta)$.

<u>Proof</u>: The following discussion will be only in terms of the sequence $[m-1/m](s)$, m = 1,2,...etc. Analogous arguments hold for the sequence $[m-1/m-1](s)$, m = 1,2,...etc. We shall establish the convergence of $[m-1/m](s)$, m = 1,2,...etc. by showing that the sequence of ij-th elements $[m-1/m]_{ij}(s)$, m = 1,2,...etc. of $[m-1/m](s)$ converge.

Let $D(\Delta')$ be a region similar to $D(\Delta)$ but slightly larger and containing the closure of $D(\Delta)$. Then since $[m-1/m]_{ij}(s)$ is uniformly bounded in the closure of $D(\Delta')$, each subsequence of $[m-1/m]_{ij}(s)$ is normal in $D(\Delta')$, and thus contains another subsequence converging

---

[3] the bar '-' denotes complex conjugate

locally uniformly to some analytic function $D(\Delta')$. Since all these limit functions are the same on positive real axis, they are identical in $D(\Delta')$ due to analytic continuation. Hence, $[m-1/m]_{ij}(s)$ converges locally uniformly in $D(\Delta')$ to a limit function analytic on $D(\Delta')$, and in particular converges uniformly on $D(\Delta)$.

Finally, since $G(s)$ is holomorphic in $D(\Delta)$, which is symmetric with respect to the real axis, and the property of realness of $G(s)$ is inherited by the property of realness of $[m-1/m]ij(s)$ for real values of s, it follows from the well known Schwartz reflection principle that $G(s)$ is real symmetric i.e., $\bar{G}(s) = G(\bar{s})$. ▒

Note that if $T(s)$ is a Hamburger series i.e., if only $H_n(T)$ (but not $-H_n'(T)$) is positive definite for arbitrary n then properties of uniform boundedness and equicontinuity, as proved in Theorem 2.5 and Lemma 2.6, still hold true when $D(\Delta)$ is replaced by the bounded, disconnected, two-component domain $D_I(\Delta) = \{s;\ |\text{Im}s|>\Delta,\ |s|<R<\infty\}$. Consequently, the first paragraph in the proof of Theorem 2.7 applies and we may assert that there exists a subsequence of the sequence of MPA's which converge uniformly everywhere in $D_I(\Delta)$ to a real symmetric function analytic in $D_I(\Delta)$. However, since in this case γ_ν's are not necessarily positive, Theorems 2.2 and 2.3 do not apply and consequently, the pointwise convergence of the sequence of MPA's for real positive values of s cannot be established, thus making Vitali's theorem inapplicable.

3. Matricial Hamburger and Stieltjes Moment Problem and related results:

In this section we undertake the solution of the matricial version of the classical Hamburger or Stieltjes moment problem. More specifically, the following result stated in Theorem 3.1 will be proved. An integral representation of the functions $G(s)$ and $G'(s)$ of Theorem 2.7, when the Stieltjes series (1.1) has a nonzero radius of convergence is also derived in this connection.

We first need the following definition.

Definition: A real symmetric matrix valued function $\sigma(x)$ of a real variable x will be said to be non-decreasing (increasing) if the matrix $\sigma(x_1)-\sigma(x_2)$ is non-negative (positive) definite, whenever $x_1 > x_2$.

Theorem 3.1: (a) If the block Hankel matrices $H_n(T)$ in (1.2a) are positive definite for all non-negative integer values of n then there exists a non-decreasing matrix measure $\sigma(x)$ such that the matricial Stieltjes integral representation (3.1) for T_k hold true.

$$(-1)^k\, T_k = \int_{-\infty}^{\infty} x^k d\sigma(x)\ , \quad k = 0,1,2,... \tag{3.1}$$

(b) Furthermore, if in addition to the conditions stated in part (a) the block Hankel matrices $H_n'(T)$ in (1.2b) are negative definite for all non-negative integer values of n i.e., if $T(s)$ as in (1.1) is a matrix Stieltjes series then the lower limit of the integral in (3.1) can be replaced by zero.

Note that the Riemann-Stieltjes integral over a matrix measure, as appearing in (3.1), was first introduced and their properties studied by Wiener and Masani in [7] in the context of multivariate stochastic process.

Before embarking on a proof of Theorem 3.1, the matrical version of Gauss quadrature formula proved in [1, Theorem 3.3] will be recalled in a notation compatible with the present discussion.

Theorem 3.2 [1]: If $H_n(T)$ is positive definite for all non-negative integer values of n then for any fixed integer m > 0, there exist real symmetric non-negative definite (pxp) matrices A_ν and real numbers γ_ν each depending on m,[4] such that

$$(-1)^k \, T_k = \sum_{\nu=1}^{r} A_\nu \gamma_\nu^k \text{ for } k = 0,1,\ldots(2m-1) \qquad (3.2)$$

where r = mp. Furthermore, if $H_n'(T)$ is negative definite for all n, then the γ_ν's are necessarily positive.

Note that when both $H_n(T)$ and $(-H_n'(T))$ are positive definite, (3.2) follows from (2.6) and then by observing that the coefficient of s^k in the power series expansion of [m-1/m](s) around s = 0 is T_k for k = 0,1,...(2m-1), thus establishing the matricial Gauss quadrature formula (3.2) via electrical network theoretic arguments (in fact, (3.2) is identical to (2.7)). However, when only $H_n(T)$ but not $(-H_n'(T))$ is positive definite the network interpretations of [m-1/m](s) in (2.6) cannot be given and a detailed proof of (3.2) as worked out in [1] is called for.

Definition [7]: A matrix valued function $\sigma(x)$ will be said to be of bounded variation in [a,b] if $\sum_{\nu=1}^{k} ||\sigma(x_\nu)-\sigma(x_{\nu-1})||$ is bounded for any partition $a = x_0 < x_1 \ldots < x_k = b$ of the interval [a,b].

Lemma 3.3: If $\sigma(x)$ is a non-decreasing real symmetric matrix valued function such that $M \geq \sigma(x) \geq 0$ for all $x \in [a,b]$ then each element of $\sigma(x)$ is bounded for all $x \in [a,b]$. Furthermore, $\sigma(x)$ as well as each of its entries are of bounded variation in [a,b].

Proof: Let $\sigma^{(ij)}(x)$ denote the ij-th entry of the matrix $\sigma(x)$. Since $M \geq \sigma(x) \geq 0$ it follows from Property 2.3 of spectral norm that $||\sigma(x)|| \leq ||M||$ for all $x \in [a,b]$. If e_j denotes the j-th column of the (pxp) identity matrix then $\sqrt{(\sum_{i=1}^{p} |\sigma^{(ij)}(x)|^2)} = ||\sigma(x)e_j|| \leq ||\sigma(x)|| \leq ||M||$.

---

[4] To avoid clutter in notation this dependence is not reflected explicitly in (3.2)

Consequently, $|\sigma^{(ij)}(x)|<||M||$ for all $i=1,2,...p$. Since j is chosen arbitrarily each element of the matrix $\sigma(x)$ is bounded by $||M||$. This result along with the non-decreasing character of $\sigma(x)$ imply that [7, lemma 4.2(b)] the functions $\sigma^{(ij)}(x)$, $i \neq j$ are each functions of bounded variation. However, for all j, $\sigma^{(jj)}(x)$ is non-decreasing since $\sigma(x)$ is so, and furthermore $0 \leq \sigma^{(jj)}(x) \leq ||M||$ (the first inequality follows from non-negativeness of $\sigma(x)$). Thus, $\sigma^{(jj)}(x)$ is also of bounded variation in [a,b]. Consequently, each entry of $\sigma(x)$ is of bounded variation in [a,b], which is a necessary and sufficient condition for $\sigma(x)$ to be of bounded variation in [a,b] (cf. [7, Lemma 4.2(a)]). ▩

<u>Remark:</u> We note that if $\sigma(x)$ is a real symmetric matrix valued function of bounded variation in [a,b] then due to Lemma 4.2(a) of [7] $\sigma^{(ij)}(x)$ is of bounded variation for all i, j. Consequently, if $f(x)$ is any continuous function in [a,b] then $\int_a^b f(x)d\sigma^{(ij)}(x)$ exists [9] and consequently, due to [7, Lemma 4.8] the matricial Riemann-Stieltjes integral $\int_a^b f(x)d\sigma(x)$ also exists.

<u>Lemma 3.4:</u> If $\sigma(x)$ is any real symmetric non-decreasing matrix valued function of bounded variation in [a,b] then

$$||\int_a^b f(x)\, d\sigma(x)|| \geq ||\int_a^b g(x)d\sigma(x)|| \tag{3.3}$$

where $f(x)$ and $g(x)$ are continuous scalar functions such that $f(x) \geq g(x) \geq 0$ for all x in the interval of integration.

<u>Proof</u>: Consider the function $h(x)=f(x)-g(x)$ defined in [a,b]. The existence of the integrals in (3.3) and of $\int_a^b h(x)d\sigma(x)$ then immediately follow from the remark preceeding the present lemma. Furthermore, we also have:

$$\int_a^b h(x)d\sigma(x) = \int_a^b f(x)d\sigma(x) - \int_a^b g(x)d\sigma(x) \tag{3.4}$$

Since $f(x) \geq g(x) \geq 0$ the functions $f(x)$, $g(x)$ and $h(x)$ are all non-negative in [a,b]. Cosequently, due to the non-decreasing character of the matrix valued measure $\sigma(x)$ it trivially follows from the definition of the matricial Riemann-Stieltjes integrals that each of the integrals in (3.4) is real symmetric non-negative definite. The present Lemma then follows from Property 2.3 of spectral norm. ▒

Our strategy for proof of Theorem 3.1 is, in fact, a matricial generalization of a technique elaborated in [6] in the scalar context.

Proof of Theorem 3.1: Let $\gamma_1 \leq \gamma_2 \leq \ldots \leq \gamma_r$ be an ordering of the γ_ν's in (3.2) and consider the real symmetric non-negative definite matrix valued function $\sigma_m(x)$ defined over $-\infty < x < +\infty$ as in (3.5).

$$
\begin{aligned}
\sigma_m(x) &= 0 \text{ for } x < \gamma_1 \\
&= \sum_{\nu=1}^{\mu} A_\nu \text{ for } \gamma_\mu \leq x < \gamma_{\mu+1} \\
&= \sum_{\nu=1}^{r} A_\nu \text{ for } x \geq \gamma_r
\end{aligned}
\tag{3.5}
$$

Then the following properties of $\sigma_m(x)$ are clear.

(P1) $\sigma_m(x)$ is real symmetric non-negative definite for all real x. Furthermore, if $x_1 > x_2$ then $\sigma_m(x_1) \geq \sigma_m(x_2)$ i.e., $\sigma_m(x)$ is a non-decreasing matrix valued function of x.

(P2) From (3.2) with k=0 and (3.5) it follows that $(T_0 - \sigma_m(x))$ is real symmetric non-negative definite for all real x and for all m > 0.

(P3) Since due to (P1), (P2) and Lemma 3.3, $_\infty\sigma(x)$ is of bounded variation, the matrical Riemann-Stieltjes integral $\int_{-\infty}^{\infty} x^k \, d\sigma_m(x)$ exists element wise [9] and is equal to $\sum_{\nu=1}^{r} \gamma_\nu^k A_\nu$, which, due to (3.2), is equal to $(-1)^k T_k$ for $k = 0,1,\ldots(2m-1)$.

It thus follows from (P1) and (P2) above and from Theorem A1 in the appendix that a subsequence $\sigma_{m_i}(x)$; i = 1,2,...etc. of the sequence $\sigma_m(x)$, m = 1,2,...etc. converges to a real symmetric non-negative definite matrix valued function $\sigma(x)$, which is also non-decreasing. Therefore, it follows from (P3) above that

$$(-1)^k T_k = \int_{-\infty}^{\infty} x^k \, d\sigma_{m_i}(x) \text{ for all } m_i \geq \tfrac{1}{2}(k+1) \qquad (3.6)$$

The following considerations then hold for any choice of finite real numbers a, b with $a < -1$, $1 < b$ and $m_i \geq k + 1$.

Thus, from (3.6) and triangle inequality for spectral norm [4]:

$$\|(-1)^k T_k - \int_a^b x^k \, d\sigma(x)\| = \| \int_{-\infty}^{+\infty} x^k d\sigma_{m_i}(x) - \int_a^b x^k d\sigma(x)\|$$

$$\leq \|\int_{-\infty}^{a} x^k d\sigma_{m_i}(x)\| + \|\int_a^b x^k d\sigma_{m_i}(x) - \int_a^b x^k d\sigma(x)\|$$

$$+ \|\int_b^{\infty} x^k d\sigma_{m_i}(x)\| \qquad (3.7)$$

However,
$$\|\int_{-\infty}^{a} x^k d\sigma_{m_i}(x)\| = \|(-1)^k \int_{-\infty}^{a} |x^k| d\sigma_{m_i}(x)\|$$

$$= \|\int_{-\infty}^{a} |x|^k d\sigma_{m_i}(x)\|$$

$$\leq \| \frac{1}{|a|^{k+2}} \int_{-\infty}^{a} x^{2k+2} d\sigma_{m_i}(x)\|$$

$$\leq \frac{1}{|a|^{k+2}} \|\int_{-\infty}^{+\infty} x^{2k+2} d\sigma_{m_i}(x)\|$$

$$= \frac{1}{|a|^{k+2}} \|T_{2k+2}\| \qquad (3.8)$$

where in (3.8) the first equality follows from the fact that for $x<0$, $x^k=(-1)^k|x|^k$; the second equality from Property 2.2 of $\|.\|$ with $\alpha=-1$; whereas the first inequality follows from the fact that if $x \leq a < -1$ then $x^{2k+2}/|a|^{k+2} \geq |x|^k > 0$ in conjunction with Lemma 3.4; the second inequality from Properties 2.2 and 2.3 of spectral norm $\|.\|$; and the last equality from (3.6) above as a consequence of the choice $m_i \geq k+1$. It can also be shown in an analogous fashion that:

$$\|\int_b^{\infty} x^k d\sigma_{m_i}(x)\| \leq \frac{1}{|b|^{k+2}} \|T_{2k+2}\| \qquad (3.9)$$

Thus, from (3.7) it is possible to assert that (3.10) in the following holds for $a < -1$, $1 < b$, and $m_i \geq (k+1)$.

$$\|(-1)^k T_k - \int_a^b x^k d\sigma(x)\| \leq$$

$$||\int_a^b x^k d\sigma_{m_i}(x)-\int_a^b x^k d\sigma(x)||+||T_{2k+2}||(|a|^{-(k+2)}+|b|^{-(k+2)}) \qquad (3.10)$$

From Theorem A2 in the appendix it follows that the first terms in the right hand side of (3.10) goes to zero as $m_i \to \infty$. Thus, for all k = 0,1,2,...etc.

$$||(-1)^k T_k - \int_a^b x^k d\sigma(x)|| \leq ||T_{2k+2}||(|a|^{-(k+2)} + |b|^{-(k+2)}) \qquad (3.11)$$

Furthermore, as $a \to -\infty$ and $b \to \infty$ (3.11) yields $||(-1)^k T_k - \int_{-\infty}^{\infty} x^k d\sigma(x)||=0$, thus [4] proving that (3.1) holds for all k = 0,1,2,...etc.

Part (b) of the theorem follows by observing that in Theorem 3.2 if $H_n'(T)$ is negative definite for all n then γ_ν's are necessarily positive, which in turn implies that $\sigma_m(x)$, as defined in (3.5), and thus $\sigma(x)$, is zero for all negative x. ▒

Note that if $U_k = (-1)^k T_k$ and $H_n(U)$ and $H_n'(U)$ are the Hankel matrices obtained by replacing the T_k's in (1.2a,b) by the corresponding U_k's then it follows via straightforward algebraic manipulation that $H_n(T)>0$ if and only if $H_n(U) > 0$, whereas $H_n'(T) < 0$ if and only if $H_n'(U)>0$. The solutions to matricial versions of Hamburger and Stieltjes moment problems then follow in a more conventional form as stated in [10] in the scalar case from this observation.

Note that the following result can be viewed as a matricial generalization of the well known scalar result [10] that the non-decreasing Stieltjes measure $\sigma(x)$ must, in fact, have infinitely many points of increase.

<u>Property 3.5:</u> For any non-zero (1xp) real constant vector v the function $v^t\sigma(x)v$ of x must have infinite number of points of increase.

<u>Proof</u>: Assume that the result is false i.e., there exists some v such that $v^t\sigma(x)v$ can be viewed as a linear combination of a finite number N of step functions, occurring at, say, $\alpha_1,\alpha_2,\ldots\alpha_N$. Consider the polynomial p(x) as in (3.12a). Then (3.12b) follows from (3.1) and (1.2a).

$$p(x)=\prod_{i=1}^{N}(x-\alpha_i)=\sum_{k=0}^{N} a_k x^k; \quad \int_{-\infty}^{\infty} p^2(x)d\sigma(x)=A_N^t H_{N+1}(T)A_N \qquad (3.12a,b)$$

where A_N^t is the block row matrix $(a_0I, a_1I,...a_NI)$ and I is the (pxp) identity matrix. Since not all a_i's are zero, A_N is of rank p, thus implying, in view of positive definiteness of $H_{N+1}(T)$, that

$$v^t(\int_{-\infty}^{\infty} p^2(x)d\sigma(x))v > 0 \qquad (3.13)$$

However, by recalling the definition of Riemann-Stieltjes integrals over a matrix measure, it follows from the fact that $v^t\sigma(x)v$ is a linear combination of step functions that the left hand side of (3.13) is exactly equal to zero, which is a contradiction. ▒

We next assume that the matrix Stieltjes series T(s) in (1.1), which was so far considered only as a formal power series, to have a radius of convergence R. Then by using the representation (3.1) we can further prove the following.

<u>Theorem 3.6</u>: If T(s) in (1.1) has a nonzero radius of convergence R and the associated Hankel matrix $H_n(T)$ satisfies $H_n(T)>0$ for all n then (i) $\sigma(x)$ = constant for $|x|>R^{-1}$ (ii) for all s in $|s|<R$ we may write:

$$T(s) = \int_{-\infty}^{\infty} \frac{1}{1+sx}d\sigma(x) = \int_{-R^{-1}}^{R^{-1}} \frac{1}{1+sx}d\sigma(x) \qquad (3.14)$$

(iii) The sequences of MPA's $[m-1/m](s)$ and $[m-1/m-1](s)$; m=1,2,...etc. converge uniformly to the expression (3.14). In particular, if T(s) is a matrix Stieltjes series then the limit functions G(s) and G'(s) of Theorem 2.7 are both given by (3.14).

(iv) If, in addition, $H_n'(T)<0$ for all n i.e., T(s) is a matrix Stieltjes series then the lower limits in the integrals in (3.11) can be replaced by zero.

Note that in the last case (3.14) coincides with the integral representation of RC impedances known as Cauer's representation in classical network theory [20].

Proof of Theorem 3.6 uses Proposition 4.1, which, however, has been included in Section 4 for an improved categorization of results of similar nature.

Proof: (i) Let $P_m(s)$ be the denominator polynomial matrix associated with the right MPA of order [m-1/m] to T(s), and $\hat{P}_m(s)$ be the corresponding 'inverse' polynomial matrix as defined in (4.1). Define r_m via $r_m^{-1} = \max(|\hat{\alpha}_m|, |\hat{\beta}_m|)$, where $\hat{\alpha}_m$ and $\hat{\beta}_m$ are as described in Proposition 4.1 from which it also follows that $r_{m+1} \leq r_m$ for all m=1,2,...etc. Then [m-1/m](s) is analytic in $|s|<r_m$ and thus, its power series expansion around s=0 converges in $|s|<r_m$. Furthermore, as $m \to \infty$ this latter expansion coincides with T(s) in (1.1), which is assumed to have a radius of convergence R. Consequently, $R<r_m$ and thus, $|\hat{\alpha}_m|<R^{-1}$, $|\hat{\beta}_m|<R^{-1}$ for m=1,2...etc. Next, since for any fixed m, the γ_ν's in (2.6) are (a subset of) the zeros of $\det\hat{P}_m(s)$, the latter.. conclusion yields that $|\gamma_\nu|<R^{-1}$ for all ν and m. Thus, it follows from (3.5) that for all m $\sigma_m(x)$=constant if $|x|>R^{-1}$, which in turn imply that $\sigma(x)$=constant if $|x|>R^{-1}$.

(ii) The following considerations hold for real valued s with $|s|<R$.

Define $\phi_n(x)= \sum_{k=0}^{n} (-sx)^k$ and $\psi(x)=(1-|sx|)^{-1}$. Clearly then for all x in $-R^{-1} \leq x \leq R^{-1}$ we have $|\phi_n(x)|<\psi(x)$ and $\phi_n(x) \to (1+sx)^{-1}$ as $n \to \infty$. Furthermore, since $\psi(x)$ is continuous and $\sigma(x)$ is of bounded variation (cf. proof of Theorem 3.1) in $-R^{-1} \leq x \leq R^{-1}$, it follows from [7] that the matricial Riemann-Stieltjes integral of $\psi(x)$ with respect to $d\sigma(x)$ over the interval $-R^{-1} \leq x \leq R^{-1}$ exists. By applying the dominated convergence theorem of the theory of functions of a real variable to the sequences formed from the respective entries of matrices it then follows that:

$$\int_{-R^{-1}}^{R^{-1}} \phi_n(x)d\sigma(x) \to \int_{-R^{-1}}^{R^{-1}} (1+sx)^{-1}d\sigma(x) \text{ as } n \to \infty.$$

Thus, the proof of (3.14) for real values of s follows from (1.1) and (3.15) in which use of (3.1) along with the fact that $\sigma(x)$=constant for $|x|>R^{-1}$ have been made.

$$\sum_{k=0}^{n} T_k s^k = \sum_{k=0}^{n} \int_{-R^{-1}}^{R^{-1}} (-sx)^k d\sigma(x) = \int_{-R^{-1}}^{R^{-1}} \phi_n(x) d\sigma(x) \qquad (3.15)$$

The validity of (3.14) for complex values of s then follows from the principle of analytic continution by noting that both T(s) in (1.1) and the extreme right hand side of (3.14) are analytic in $|s|<R$.

(iii) Follows from the fact that the limit functions to which the sequences [m-1/m](s) and [m-1/m-1](s) of MPA's converge and T(s) are each holomorphic in $|s|<R$, in which they have identical power series expansion, namely, (1.1).

(iv) Finally, if T(s) is a matrix Stieltjes series then all γ_ν's are positive; thus due to (3.5) $\sigma(x)$=constant for $x<0$. Consequently, the lower limit of the integrals in (3.14) can be replaced by zero. ▒

The following comment is in order with respect to item (iii) of the above theorem. In the scalar case it has been shown that even if R=0, the integral representation (3.14) for the limit functions remains valid if the coefficients of the power series further satisfies the so called 'Carleman criterion' [3]. An extension of this result in the matrix case is not pursued here (see e.g.,[21] and references therein).

4. Matrix Orthogonal polynomials of the second kind

The fact that the sequence of inverse polynomial matrices, constituting the 'denominators' of MPA's to a matrix Stieltjes series form a sequence of matrix orthogonal polynomials has already been pointed out in [1]. However, in [1] the orthogonality relation was viewed as an algebraic relation i.e., in terms of orthogonality of vector spaces. Presently, it will be shown that this relation can be interpreted as an orthogonality relation with respect to the matrix valued measure $\sigma(x)$ developed in the previous section. Certain other results as natural generalizations of the scalar theory such as the orthogonal polynomials of the second kind and their properties follow as consequences of this discussion.

Consider the set of 'inverse' polynomial matrix $\hat{P}_m(s)$ as in (4.1), where $P_m(s)$ is the 'denominator' polynomial matrix associated with the $[m-1/m](s)$ MPA's for $T(s)$, and $H_n(T)$ in (1.2a) for each n is positive definite. (For the purpose of present section no restriction is imposed on $H_n'(T)$).

$$\hat{P}_m(s) = s^m P_m(s^{-1}) \text{ for all } m. \tag{4.1}$$

Then the following results hold true.

Proposition 4.1: If $\hat{\alpha}_m$ and $\hat{\beta}_m$ are respectively the largest and smallest zeros of $\det\hat{P}_m(s)$ then $\hat{\alpha}_{m+1} \geq \hat{\alpha}_m$, $\hat{\beta}_{m+1} \leq \hat{\beta}_m$ for all m=1,2,...etc.

Proof: As shown in [1] the zeros of $\det\hat{P}_m(s)$ are the eigenvalues of the block tridiagonal matrix in (4.2a), where $C_k = D_k^{-1} K_k D_k$, $C_0 = D_0^{-1} T_1$, $\lambda_k = D_{k-1}^{-1} D_k$, and D_k's are real symmetric positive definite, whereas the $K_k D_k$'s are real symmetric matrices.

$$\begin{bmatrix} C_0 & \lambda_1 & & \\ I & C_0 & \ddots & \\ & \ddots & \ddots & \lambda_{m-1} \\ & & I & C_{m-1} \end{bmatrix} ; \quad Z_{m-1} = \begin{bmatrix} T_1 & D_1^{\frac{1}{2}} & & \\ D_1^{\frac{1}{2}} & K_1 D_1 & \ddots & \\ & \ddots & \ddots & D_{m-1}^{\frac{1}{2}} \\ & & D_{m-1}^{\frac{1}{2}} & K_{m-1} D_{m-1} \end{bmatrix} \tag{4.2a,b}$$

Thus, the zeros of $\det\hat{P}_m(s)$ are also eigenvalues of the real symmetric block tridiagonal matrix Z_{m-1} in (4.2b), where $D_k^{1/2}$ stands for the Hermitian square root of D_k. It then follows from the Courant-Fisher [2] theorem that

$$\hat{\alpha}_m = \max\{x^t Z_{m-1} x;\ ||x||=1\};\quad \hat{\alpha}_{m+1} = \max\{y^t Z_m y;\ ||y||=1\} \qquad (4.3a,b)$$

where x and y are column vectors of size mp and (m+1)p respectively. Since from (4.2b) we have that

$$Z_m = \left[\begin{array}{c|c} Z_{m-1} & D_m^{1/2} \\ \hline D_m^{1/2} & K_m D_m \end{array}\right] \qquad (4.4)$$

it follows from (4.3a) that $\hat{\alpha}_m$ can also be considered as the maximum value of $y^t Z_m y$ subject to the restriction that $||y||=1$ and that the last p elements of y are zero. Thus, $\hat{\alpha}_{m+1} \geq \hat{\alpha}_m$. The result $\hat{\beta}_{m+1} \leq \hat{\beta}_m$ also follows from similar arguments if $\hat{\beta}_m$ and β_{m+1} are expressed as the minimum values of the quadratic forms in (4.3). ▒

Note that in the scalar case i.e., if p=1 the above argument also leads to the interlacing property of zeros of $\hat{P}_m(s)=\det\hat{P}_m(s)$ and $\hat{P}_{m+1}(s) = \det\hat{P}_{m+1}(s)$, whereas in the matrix case interlacing properties of this type are not known to hold.

<u>Proposition 4.2</u>: If $H_n(T)$ as given in (1.2a) is positive definite for all n then the matrical Stieltjes integral

$$\int_{-\infty}^{\infty} \hat{P}_\mu^{\,t}(x)\, d\sigma(x)\, \hat{P}_\nu(x) \qquad (4.5)$$

is positive definite when $\mu = \nu$ and is a zero matrix when $\mu \neq \nu$, where $\sigma(x)$ is the real symmetric non-decreasing matrix valued function of the real variable x as appearing in Theorem 3.1.

<u>Proof:</u> Let $P_m(s) = \sum_{k=0}^{m} p_k^{(m)} s^k$ and consequently, $\hat{P}_m(s) = \sum_{k=0}^{m} p_{m-k}^{(m)} s^k$,

where $p_k^{(m)}$'s are real (pxp) matrices. Also, note that since $\sigma(x)$ is real symmetric and $P_\mu(x)$ as well as $P_\nu(x)$ are real valued matrices for real x, it is enough to prove the result for $\nu \geq \mu$. The case of $\nu < \mu$ then æollows by considering the transpose of (4.5). It follows via the use of (3.1) in a straightforward manner that

$$\int_{-\infty}^{\infty} \hat{P}_{\mu}^{t}(x)d\sigma(x)\hat{P}_{\nu}(x) = [M_{\mu}^{t}|\overbrace{0|\ldots|0}^{\nu-\mu}]H_{\nu}(T)M_{\nu}. \qquad (4.6)$$

where $H_{\nu}(T)$ is as defined in (2.1a) and M_{ν} is defined as the $px(\nu+1)p$ matrix $M_{\nu} = [p_{\nu}^{(\nu)t}|p_{\nu-1}^{(\nu)t}|\ldots|p_{1}^{(\nu)t}|I]^{t}$. However, it also follows from the normal equations (equation (3.1) in [1]) defining the right MPA's that $H_{\nu}(T)M_{\nu} = [0|0|\ldots|0|D_{\nu}^{t}]^{t}$, where D_{ν} is a real symmetric positive definite matrix of size (pxp). Therefore, due to (4.6) we have that

$\int_{-\infty}^{\infty} \hat{P}_{\mu}^{t}(x)d\sigma(x)\hat{P}_{\nu}(x)$ is equal to D_{ν} when $\nu=\mu$ and is equal to 0 when $\nu\neq\mu$. The proposition is thus proved. ▩

<u>Proposition 4.3</u>: If P(s) is any (pxp) polynomial matrix such that each of its elements are of degree strictly less than m then

$$\int_{-\infty}^{\infty} P(x)\, d\sigma(x)\, \hat{P}_{m}(x) = 0\ ;\ \int_{-\infty}^{\infty} \hat{P}_{m}^{t}(x)\, d\sigma(x)P(x) = 0 \qquad (4.7a,b)$$

<u>Proof:</u> Since implicit in the defintion of right MPA [1] is the fact that $p_{0}^{(m)} = P_{m}(0) = I$ i.e., $\hat{P}_{m}(x)$ is monic for all m, it follows that $P^{t}(s)$ can be written as:

$$P^{t}(s) = a_{m-1}\hat{P}_{m-1}(s) + a_{m-2}\hat{P}_{m-2}(s) + \ldots + a_{0}\hat{P}_{0}(s)$$

where a_i's are constant (pxp) matrices. Then (4.7a) follows from Proposition 4.2. Analogous arguments hold for (4.7b). ▩

Next, for all m=0,1,...etc. define the matrix polynomial $\hat{Q}_{m-1}(s)$ of degree (m-1) (where $\hat{P}(s)$ is the inverse polynomial matrix corresponding to $P_{m}(s)$ as given in ((4.1)) via the relation:

$$\hat{Q}_{m-1}(s) = \int_{-\infty}^{\infty} d\sigma(x)[(\hat{P}_{m}(s) - \hat{P}_{m}(x))/(s-x)] \qquad (4.8)$$

The following properties of $\hat{Q}_{m-1}(s)$ are then imminent.

Proposition 4.4: For any m=1,2,...etc. if $P_m(s)$ is the denominator polynomial matrix associated with the right MPA of order [m-1/m] to the series T(s), which satisfies the condition $H_n(T)>0$ for all n then $Q_{m-1}(s)$ defined via $Q_{m-1}(s)=s^{m-1}\hat{Q}_{m-1}(s^{-1})$ is, in fact, the numerator polynomial matrix of the right MPA of order [m-1/m] to T(s). Furthermore, the identity (4.9) holds true for all m=1,2,...etc.

$$\int_{-\infty}^{\infty} d\sigma(x)\ [s\hat{P}_m(s)-x\hat{P}_m(x))/(s-x)] = s\hat{Q}_{m-1}(s) \tag{4.9}$$

Proof: It follows from $\hat{P}_m(s) = \sum_{k=0}^{m} p_{m-k}^{(m)} s^k$ and (4.8) that

$$\hat{Q}_{m-1}(s) = \int_{-\infty}^{\infty} d\sigma(x) \sum_{k=0}^{m} [(s^k-x^k)/(s-k)]\ p_{m-k}^{(m)}$$

$$= \int_{-\infty}^{\infty} d\sigma(x) \sum_{k=1}^{m} \left(\sum_{i=0}^{k-1} x^i s^{k-1-i}\right) p_{m-k}^{(m)} = \sum_{k=1}^{m} \sum_{i=0}^{k-1} T_i\ p_{m-k}^{(m)} s^{k-1-i} \tag{4.10}$$

where the last equality follows via the use of equation (3.1). Furthermore, we then also have:

$$Q_{m-1}(s) = s^{m-1}\hat{Q}_{m-1}(s^{-1}) = \sum_{k=1}^{m} \sum_{i=0}^{k-1} T_i\ p_{m-k}^{(m)} s^{m-k+i} = \sum_{j=0}^{m-1} \left(\sum_{h=0}^{j} T_h\ p_{j-h}^{(m)} \right) s^j \tag{4.11}$$

where the last equality follows by a straightforward rearrangement of the indices of the double sum. Since from (4.11) it follows that $Q_{m-1}(s)-T(s)P_m(s)= o(s^{2m})$ i.e., the coefficients of s^k for k=0,1,...(2m-1) are all zero, the polynomial matrix $Q_{m-1}(s)$ is indeed the 'numerator' associated with the right MPA of order [m-1/m] corresponding to the formal powes series T(s).

By following a sequence of steps analogous that used in the derivation of (4.11) above it can also be shown that

$$\int_{-\infty}^{\infty} d\sigma(x)\ [s\hat{P}_m(s)-x\hat{P}_m(x))/(s-x)] = \sum_{k=0}^{m} \sum_{i=0}^{k} T_i s^{k-i} p_{m-k}^{(m)}$$

$$= \sum_{j=0}^{m} \left(\sum_{h=0}^{j} T_h p_{j-h}^{(m)} \right) s^{m-j} \qquad (4.12)$$

where the first equality follows from straightforward algebraic manipulation and a use of (3.1), whereas the second equality involves a rearrangement of indices of the double sum. The result in (4.9) then follows by noting that due to the normal equations [1] defining the right MPA's the term in (4.12) with j = m is zero i.e., $\sum_{h=0}^{m} T_h \, p_{m-h}^{(m)} = 0$ and from (4.11) $s\hat{Q}_{m-1}(s) = \sum_{j=0}^{m-1} \sum_{h=0}^{j} T_h p_{j-h}^{(m)} s^{m-j}$ ▩

Note that in view of the properties elaborated upon in the following the sequence of matrix polynomials $\hat{Q}_{m-1}(s)$, m = 1,2,...etc. can be regarded as the natural generalization of sequence of scalar polynomials of the second kind treated in the classical literature [10].

The fact that the sequence of matrix polynomials $\hat{P}_m(s)$, m = 0,1,2...etc. satisfies the recurrence relation (4.13) has been shown in [1] i.e., (4.13) holds for m = 1,2,...etc.

$$\hat{P}_{m+1}(s) = \hat{P}_m(s)\ (sI-C_m) - \hat{P}_{m-1}(s)\lambda_m \qquad (4.13)$$

where C_m and λ_m are real (pxp) matrices such that $C_m = D_m^{-1} K_m D_m$ and $\lambda_m = D_{m-1}^{-1} D_m$ with D_m for all m are real symmetric positive definite, and $K_m D_m$ for all m are real symmetric matrices.

<u>Proposition</u> <u>4.5</u>: The sequence of polynomials $\hat{Q}_m(s)$, m = 0,1,2,...etc. satisfies the same recurrence relations as $\hat{P}_m(s)$. More specifically, following three term recurrence relation holds true.

$$\hat{Q}_{m+1}(s) = \hat{Q}_m(s)(sI - C_m) - \hat{Q}_{m-1}(s)\lambda_m, \ m=0,1,...\text{etc.} \qquad (4.14)$$

with $\hat{Q}_{-1}(s)=0$, $\hat{Q}_0(s)=T_0$, and C_m and λ_m as in the context of (4.13).

<u>Proof:</u> We substract equation (4.13) with s = s from equation (4.13) with s=x. By considering the (left) Stieltjes integral of the resulting

equation with respect to the matrix measure $d\sigma(x)$ the recurrence relation (4.14) follows by observing equations (4.8) and (4.9). Finally, the facts that $\hat{Q}_{-1}(s)=0$ and $\hat{Q}_0(s) = T_0$ follow obviously from (4.8) and that $\hat{P}_0(s)$ is monic. ▒

The following result shows that the zeros of $\hat{Q}_{m-1}(s)$, enjoy properties similar to those of the zeros of $\hat{P}_m(s)$ as discussed in [1].

<u>Proposition 4.6</u>: If $H_n(T) > 0$ for all n, then (i) all zeros of $\det\hat{Q}_{m-1}(s)$ are real (ii) if β_j is a zero of det $\hat{Q}_m(s)$ of multiplicity n there exists a set of exactly n linearly independent sets of (1 x p) vectors $\{v_j^1, v_j^2, \ldots v_j^n\}$ such that $v_j^i\hat{Q}_m(s) = 0$, $i = 1,2,\ldots n$.

(iii) any zero of det $\hat{Q}_{m-1}(s)$ cannot be of multiplicity larger than p
(iv) invariant factors in the Smith cannonical form for $\hat{Q}_{m-1}(s)$ cannot have zeros of multiple order.

Since the proof of the above proposition is essentially a consequence of the recurrence relation (4.13) and follows in exactly the same way as that of the corresponding properties of the sequence of matrix polynomials $\hat{P}_m(s)$, as elaborated in [1, Theorem 3.1, Corollaries 3.1, 3.2], it will be ommitted for the sake of brevity.

The three term recurrence relation (4.14) connecting successive members of the 'denominator' sequence of matrix polynomials, when coupled with the corresponding recurrence relation for the 'numerator' sequence (4.13) discussed in [1] provides a fast recursive algorithm for computing the paradiagonal sequence of MPA's to a matrix Stieltjes series. We note that similar recursion for the problem of computing matrix Pade' approximants in general has been discussed in [14]. If, in addition to $H_n(T)>0$, we also have $H_n'(T)<0$ for all n i.e., T(s) is a matrix Stieltjes series then it follows from the impedance or admittance interpretation of [m-1/m](s) that zeros of $\det\hat{Q}_{m-1}(s)$ are also negative.

5. Conclusion:

The present work can be viewed as a continuation of [1]. While algebraic properties of the sequences of matrix Pade' approximants of certain orders to a matrix Stieltjes series were investigated in [1], the present work is concerned with the relevant analytic and convergence properties of paradiagonal sequences of MPA's. Although, our exposition has been in terms of the sequences [m-1/m] and [m/m] of MPA's it is in general possible to derive analogous results for any paradiagonal sequence [m+j/m]; $j \geq -1$. However, the network theoretic interpretations of the results are then lost.

By using network theoretic interpretations of Pade' approximants to a matrix Stieltjes of certain orders it has been shown that the sequences of these MPA's always converge uniformly in a open bounded region of the complex plane excluding the negative real axis. Thus, a formal matrix Stieltjes series can be used to meaningfully represent a class of RC-distrubuted multiports in terms of an equivalent circuit. This result, which to the best of our knowledge has not appeared anywhere, is indeed interesting in view of the fact that the criteria for realizability of non-rational positive functions in terms of interconnections of (infinite number of) conventional lumped elements is not known [17].

Solutions to the Matricial versions of classical Hamburger and Stieltjes moments problems are obtained, and as a consequence of this discussion an integral representation for the RC-distributed multiport impedance, which in fact is closely related to the Cauer's representation for RC-impedances, is obtained when the associated Stieltjes series is assumed to be convergent in a disc of finite radius. This representation is also found to be a direct matricial generalization of the well known Stieltjes function in classical scalar literature [10].

The sequence of 'numerator' polynomial matrices of MPA's of certain orders to a Stieltjes series are shown to be natural generalization of scalar orthogonal polynomials of second kind, and their properties studied by making reference to the corresponding results for 'denominator' sequences i.e., the matrix polynomials of the first kind elaborated in [1]. Thus, the present discussions along with those in [1] is believed to provide a more complete theory of orthogonal polynomial matrices on the real line, analogous to the theory of orthogonal polynomial matrices on the unit circle discussed in [11], [12]. Finally, the relevance of orthogonal polynomials of the former kind in the context of scattering theory is also noted in [18].

It must be noted that under the present framework all results of sections 2 and 3 (except Property 3.5) including their proofs remain valid of $H_n(T)$ and $-H_n'(T)$ in (1.2) are assumed nonnegative definite. This is so primarily due to the fact that the MPA's of order [m/m-1] and [m-1/m-1] can still be interpreted as impedance or admittance matrices of RC networks even under this broader assumption [15] (the "McMillan degree" of [m-1/m](s), which is the number of capacitors in a minimal realization in such a case can be less than (m-1)p, while under the restricted assumption adopted throughout this paper it is exactly (m-1)p; but this is of no consequence to our presentation.) However, the orthogonality properties of $\hat{P}_m(s)$ and $\hat{Q}_m(s)$discussed in section 4, and Proposition 4.2 in particular, are affected if the strict positive definiteness of $H_n(T)$ and $-H_n'(T)$ are relaxed.

From the standpoint of applications it may be mentioned that although the present work primarily deals with connections of Pade' approximations to matrix Stieltjes series and their intepretations in terms of distributed RC multiport networks, in view of their relationship with problems such as inverse scattering [18], AR modelling of stationary stochastic processes [16] etc. the potential for utilizing the results developed here in other areas of signal and system theory cannot be ruled out.

References

[1] S. Basu and N.K. Bose, Matrix Stieltjes Series and network models, SIAM J. on Math. Analysis, vol. 14, no.2, pp. 209-222, 1982.

[2] E. F. Beckenbach and R. Bellman, Inequalities, Springer Verlag, 1961.

[3] G. A. Baker, Essentials of Pade' Approximants, Academic Press, 1975.

[4] P. Lancaster and M. Tismenetsky, The Theory of Matrices, Academic Press, 1985.

[5] E. Hille, Analytic function theory, vol. II, Ginn and Company, 1962.

[6] T. S. Chihara, An Introduction to Orthogonal Polynomials, Gordon and Breach Science Publishers, New York, 1978.

[7] N. Wiener and P. Masani, The prediction theory of multivariate stochastic processes, part I, Acta. Math. vol. 98, pp. 111-150, 1957.

[8] V.P. Potopov, The multiplicative structure of J-contractive matrix functions, Translations of Amer. Math. Soc., Series 2, vol. 15, pp. 131-243, 1960.

[9] L.F. Natanson, Theory of functions of a real variable, Vol. II, English Translation, Frederick Ungar Publishing Co., New York, 1960.

[10] N.I. Akhiezer, The classical moment problem, Oliver and Boyd, London, 1965.

[11] D. Delsarte, Y.V. Genin, Y.V. Kamp, Orthogonal polynomial matrices on the unit circle, IEEE trans. on Circuits and Systems, vol. 25, no. 3, pp. 149-160, 1978.

[12] D. C. Youla and N. Kazanjian, Bauer type factorization of positive matrices and theory of matrix orthogonal polynomials on unit circle, IEEE Trans. on CAS, vol. 25, no. 3, pp. 57-69, 1978.

[13] R.E. Kalman, On partial realization, transfer functions and canonical forms, ACTA polytechnica Scandinavica, (Mathematics and Computer Science), vol. 29, pp. 9-32, November 1979.

[14] A. Bultheel, Recursive algorithms for matrix Pade' problems, Math. of Comp., vol. 35, pp. 875-892, 1980.

[15] Kuchiro Morimoto, Naoki Matsumoto, Shin-Ichi Takahashi, Matrix Pade' approximants and multi-port networks, Electronics and Communications in Japan, vol. 61A, pp. 28-36, 1978.

[16] G. Cybenko, Restrictions of Normal operators Pade' approximation and Autoregressive time series, SIAM J. on Appl. Math., vol. 15, No. 4, pp 753-767, July 1984,

[17] B. Belevitch, On the realizability of non-rational positive real functions, Circuit theory and applications, vol. 1, pp. 17-30, 1973.

[18] J.S. Geronimo, Scattering theory and matrix orthogonal polynomials on the real line, Circuits, Systems and Signal Processing, vol. 1, No. 3-4, pp. 471-494, 1984.

[19] W. B. Gragg and Anders Lindquist, On the partial realization problem, Linear Algebra and its Applications, vol. 50, pp. 277-319, 1983.

[20] W. Cauer, The Poisson integral for functions with positive real parts, Bull. Amer. Math. Soc., vol. 38, pp. 713-717, Oct. 1932.

[21] Björn von Sydow, Matrix-Valued Pade' Approximation and Gaussian Quadrature, pp. 117-127, in "Det Kongelige Norse Videnskabers

Selskab, Skrifter no. 1-1983, Pade' Approximants and Continued Fractions, Universitetsforlaget Trondheim, 1983.

[22] F. J. Narcowich, R-Operators II. On the Approximation of Certain Operator-Valued Functions and the Hermitian Moment Problem, Indiana U. Math J. vol. 26, pp. 483-513, 1977.

[23] J. Zinn-Justin, Strong interaction dynamics with Pade' approximants, Phys. Lett. C1, pp. 55-102, 1971.

[24] J. L. Basdevant, D. Bessis, and J. Zinn-Justin, Pade' Approximants in Strong Ineractions: Two-Body Pion and Kaon Systems, Nuovo Cim., vol. 60A, pp. 185-237, 1969.

[25] J. S. MacNerney, Hermitian Moment Sequences, Trans. Amer. Math. Soc., vol. 103, pp. 45-81, 1962.

[26] J. S. MacNerney, Hermitian Moment Sequences -- Correction, Notices Amer. Math. Soc., vol. 25, A-427, 1978.

Appendix

In this appendix we prove the matricial version of two classical scalar theorems known as Helly's theorems [9]. We note that similar results have been derived via alternate techniques in [8] in a different context.

Theorem A1: let $\sigma_m(x)$, m = 1,2,...etc. be a sequence of non-decreasing real symmetric non-negative definite matrix valued functions defined for real values of x. If there exists a constant real symmetric non-negative definite matrix M_0 such that $M_0-\sigma_m(x)$ is non-negative definite for real x and for all m = 0,1,...etc. then there is a subsequence of the sequence $\sigma_m(x)$, m = 1,2,...etc., which converges to a real symmetric non-negative definite matrix valued function $\sigma(x)$, which is non-decreasing.

Proof: Let $f_m(x) = ||\sigma_m(x)||$, where $||.||$ denotes the spectral norm of a matrix. Since $\sigma_m(x)$ is non-decreasing, if $x_1 > x_2$ then $\sigma_m(x_1)-\sigma_m(x_2)$ is non-negative definite. Thus, due to non-negative definiteness of $\sigma_m(x)$ and Property 2.3, $||\sigma_m(x_1)|| \geq ||\sigma_m(x_2)||$ i.e., $f_m(x_1) \geq f_m(x_2)$. Consequently, $f_m(x)$ is a non-decreasing scalar function of x.

Furthermore, since $M_0-\sigma_m(x)$ is nonnegative definite it follows from Property 2.3 that $f_m(x) = ||\sigma_m(x)|| \leq ||M_0||$, for real x and all m. Thus, the scalar sequence $f_m(x)$, m = 1,2,...etc. is uniformly bounded. Therefore, by invoking a weak version of (scalar) Helley's theorem (see e.g.[6]) it follows that a subsequence of the sequence $f_m(x) = ||\sigma_m(x)||$, m = 1,2,...etc. converges to a bounded non-decreasing function f(x). However, since the convergence of the sequence of norms $||.||$ of a matrix sequence implies the convergence of the matrix sequence itself [4], it follows that the corresponding subsequence of the sequence $\sigma_m(x)$, m = 1,2,...etc. converges to $\sigma(x)$ with $||\sigma(x)|| = f(x)$. The rest of the desired properties of $\sigma(x)$ follow from the corresponding properties of $\sigma_m(x)$.

Theorem A2: Let $\sigma_m(x)$, m = 1,2,...etc. be a sequence of non-decreasing real symmetric non-negative definite matrix valued functions defined for all x in the compact interval [a,b] of the real axis such that

$M_0 \geq \sigma_m(x)$ for all m where M_0 is a constant real symmetric non-negative define matrix. Let $\sigma(x)$ be the limit function to which the above sequence converges for all x in [a,b]. Then for a continuous scalar valued function g(x) defined over [a,b], (A1) holds true

$$\lim_{m\to\infty} \int_a^b g(x)d\sigma_m(x) = \int_a^b g(x)d\sigma(x) \tag{A1}$$

Furthermore, an extension of the result holds when $a \to -\infty$ and $b \to \infty$ as in the scalar case [9].

<u>Proof:</u> First, since $0 \leq \sigma_m(x) \leq M_0$ for all $x \in [a,b]$ and for all m = 1,2,...etc., and $\sigma_m(x)$ is non-decreasing, the matrix valued functions $\sigma_m(x)$ as well as the scalar functions $\sigma_m^{(ij)}(x)$, where $\sigma_m^{(ij)}(x)$ is the ij-th element of $\sigma_m(x)$, due to Lemma 3.3, are of bounded variation in [a,b]. Consequently, $\sigma(x)$ is also of bounded variation in [a,b]. Since g(x) is continuous in [a,b] it is uniformly continuous in [a,b]. Therefore, for any $\varepsilon > 0$ it is possible to consider a partition $\{x_0, x_1, \ldots x_k\}$ of [a,b] such that

$$|g(x') - g(x'')| < \varepsilon \text{ for all } x', x'' \in [x_{\nu-1}, x_\nu],\ 1 \leq \nu \leq k \tag{A2}$$

If $\zeta_\nu \in [x_{\nu-1}, x_\nu]$, then by using mean value theorem of scalar Stieltjes integrals [9]

$$\int_{x_{\nu-1}}^{x_\nu} g(x)d\sigma^{(ij)}(x) - g(\zeta_\nu)\Delta\sigma^{(ij)}(x_\nu) = [g(\zeta'_\nu) - g(\zeta_\nu)]\Delta\sigma^{(ij)}(x_\nu) \tag{A3}$$

for some $\zeta'_\nu \in [x_{\nu-1}, x_\nu]$, where $\Delta\sigma^{(ij)}(x_\nu) = \sigma^{(ij)}(x_\nu) - \sigma^{(ij)}(x_{\nu-1})$, $\sigma^{(ij)}(x)$ being the ij-th element of the matrix $\sigma(x)$. Note that the existence of the integrals in the left hand side is guaranteed since $\sigma(x)$ is a function of bounded variation (cf. remark preceeding Lemma 3.4).

Summing (A3) over ν we obtain via the use of triangle inequality:

$$\left|\int_a^b g(x)d\sigma^{(ij)}(x) - \sum_{\nu=1}^{k} g(\zeta_\nu)\Delta\sigma^{(ij)}(x_\nu)\right| \leq \sum_{\nu=1}^{k} |g(\zeta'_\nu) - g(\zeta_\nu)||\Delta\sigma^{(ij)}(x_\nu)|$$

$$< \varepsilon \sum_{\nu=1}^{k} |\Delta\sigma^{(ij)}(x_\nu)| \leq \varepsilon V \quad \text{(A4)}$$

where $V < \infty$ is the tctal variation of the function $\sigma_b^{(ij)}(x)$ over the interval $[a,b]$. Proceeding similarily as above with $\int_a^b g(x)\, d\sigma_m^{(ij)}(x)$ instead of $\int_a^b g(x)\, d\sigma^{(ij)}(x)$ it follows that

$$\left|\int_a^b g(x)d\sigma_m^{(ij)}(x) - \sum_{\nu=1}^{k} g(\zeta_\nu)\Delta\sigma_m^{(ij)}(x_\nu)\right| < \varepsilon V \quad \text{(A5)}$$

where $\Delta\sigma_m^{(ij)}(x_\nu)=\sigma_m^{(ij)}(x_\nu)-\sigma_m^{(ij)}(x_{\nu-1})$. Thus, from (A4), (A5) and triangle inequality it follows that

$$\left|\int_a^b g(x)d\sigma^{(ij)}(x)-\int_a^b g(x)d\sigma_m^{(ij)}(x)\right|\leq 2\varepsilon V+\sum_{\nu=1}^{k} |g(\zeta_\nu)||\Delta\sigma^{(ij)}(x_\nu)-\Delta\sigma_m^{(ij)}(x_\nu)|$$

(A6)

Since the second term in the right hand side of (A6) goes to zero as $m\to\infty$ we have essentially proved that:

$$\int_a^b g(x)d\sigma^{(ij)}(x) = \lim_{m\to\infty} \int_a^b g(x)d\sigma_m^{(ij)}(x)$$

Extension of the proof when $a \to -\infty$ and $b \to \infty$ is identical to the scalar case [9] and is not repeated for brevity.

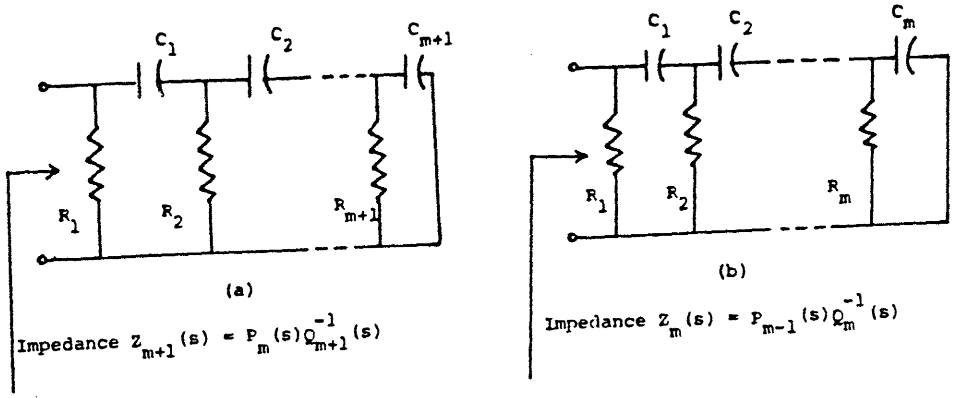

Figure 1: $Z_{m+1}(s) > Z_m(s)$ if $Ims = 0$, $Res>0$

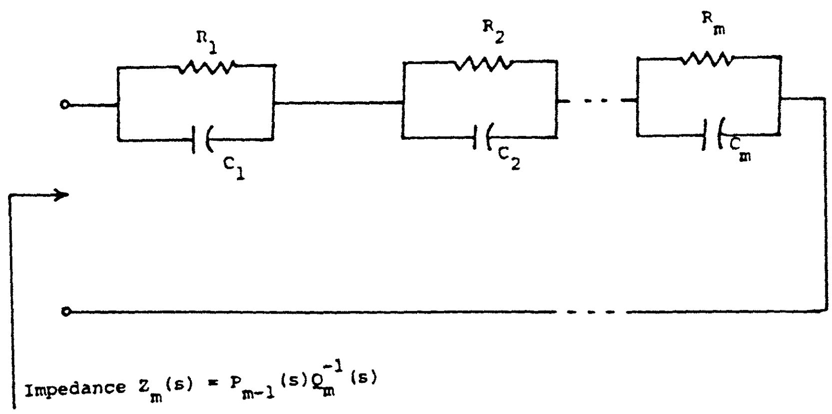

Figure 2: $Z_m(s) < R_1+R_2+\ldots+R_m$ if $Ims = 0$, $Res>0$

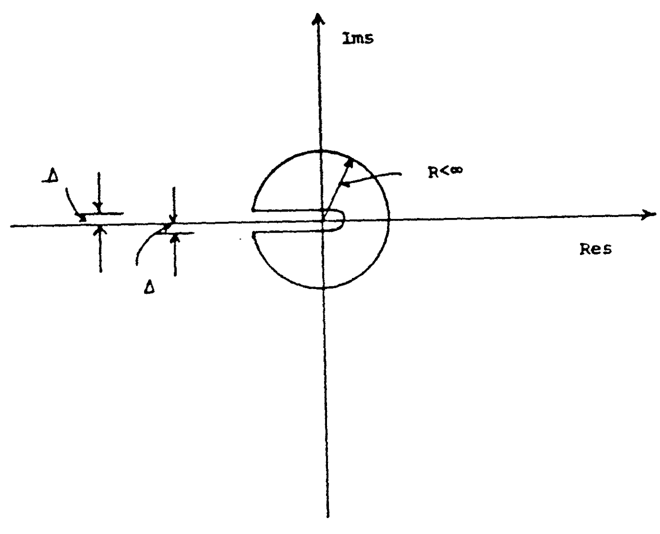

Figure 3: Region $D(\Delta)$