CIRCUITS SYSTEMS SIGNAL PROCESS
VOL. 1, NO. 3-4, 1982

0278-081X/82/3-4471-25$6.50

SCATTERING THEORY AND MATRIX ORTHOGONAL POLYNOMIALS ON THE REAL LINE*

J. S. Geronimo[1]

Abstract. The techniques of scattering and inverse scattering theory are used to investigate the properties of matrix orthogonal polynomials. The discrete matrix analog of the Jost function is introduced and its properties investigated. The matrix distribution function with respect to which the polynomials are orthonormal is constructed. The discrete matrix analog of the Marchenko equation is derived and used to obtain further results on the matrix Jost function and the distribution function.

1. Introduction

In recent years the techniques of scattering and inverse scattering theory have been used with much success to study the properties of scalar orthogonal polynomials, Case and Kac [1], Geronimo and Case [2,3], Geronimo [4,5], Guseinov [6], Dym and Iacob [7]. These techniques emphasize the important role played by factorization theory and also the close connection between polynomials orthogonal on the real line and those orthogonal on the unit circle (see also Nevai [8]).

Here the techniques of scattering and inverse scattering theory are used to investigate the properties of matrix orthogonal polynomials on the real line (see also Serebrjakov [9]). These polynomials are known to arise in a number of physical and engineering problems, Atkinson [10], Christian et al. [11], and in numerical analysis. In analogy with matrix polynomials orthogonal on the unit circle, Delsarte et al. [12]; Geronimo [13], a new set of recurrence formulas is introduced and it is shown that the solution of one of the equations leads naturally to a spectral factorization of the matrix distribution function.

In order to help motivate the problem we begin in Section 2 with the matrix distribution function. The polynomials orthonormal with respect to this distribution function are constructed and the three term recurrence formula

*Received August 31, 1981; revised March 8, 1982. This work was supported in part by NSF Grant MCS-8002731.

[1]School of Mathematics, Georgia Institute of Technology, Atlanta, Georgia, 30332, USA.

they satisfy is derived. In this section the Wronskian theorem and Christoffel-Darboux formula are given. We then turn the problem around and take the recurrence formula as our starting place (Section 3). Conditions are placed on the coefficients on the recurrence formula that allow us to introduce a new set of recurrence formulas. The analytic properties of various solutions of these recurrence formulas are investigated and the discrete matrix analog of the Jost function is introduced, (Agranovich and Marchenko [14]). In Section 4 the properties of the Jost function are investigated, and these are used to develop the relation between the Jost function and the distribution function (Section 5). The discrete matrix analog of the Marchenko equation is introduced in Section 6 and used to investigate further the properties of the Jost function. Finally in Section 7 integral representations for many of the solutions of the recurrence formula are given.

2. The matrix-valued distribution function

The $p \times p$ symmetric function $\rho(\lambda)$, λ real, is called a matrix distribution function, Atkinson [10], Berezanski [15], Krein [16], Rosenberg [17], if it has the following properties: a) $\rho(\lambda)$ is nondecreasing in the sense that $\rho(\lambda_1) \leq \rho(\lambda_2)$ for all $\lambda_1 \leq \lambda_2$ (The notation $A \leq B$ for hermitian matrices A and B means $B-A$ is nonnegative definite.); b) $\rho(\lambda)$ has infinitely many points of increase; and c) all the $p \times p$ matrix moments

$$S_n = \int_{\infty}^{\infty} \lambda^n d\rho(\lambda) \qquad n = 0,1,2,... \tag{2.1}$$

are finite. It is known that $\rho(\lambda)$ has countably many discontinuities, all of them simple (so that the limits $\rho(\lambda^+)$ and $\rho(\lambda^-)$ exist for all λ). Therefore $\rho(\lambda)$ has a derivative almost everywhere denoted by

$$\sigma(\lambda) = \frac{d\rho}{d\lambda} \qquad \text{a.e.} \tag{2.2}$$

From the properties of $\rho(\lambda)$ it is clear that $\sigma(\lambda) \geq 0$. In general $\rho(\lambda)$ can be decomposed into $\rho = \rho_{ac} + \rho_j + \rho_s$ where

$$\rho_{ac}(\lambda) = \int_{-\infty}^{\lambda} \sigma(x)dx, \tag{2.3}$$

ρ_j is a jump function and $\rho_s = \rho - \rho_{ac} - \rho_j$.

For continuous matrix-valued functions $F(\lambda)$ and $G(\lambda)$ the integral $\int G(\lambda)d\rho(\lambda)\, H(\lambda)$ is defined in the natural way: the (i,j)-entry is the (s,t)-sum of the scalar integrals $\int g_{is}(\lambda)\, h_{tj}(\lambda)d\rho_{st}(\lambda)$. In fact the above matrix integral has meaning for much more general functions G and H, Rosenberg [17]).

Define the $p(n+1) \times p(n+1)$ matrix

$$H(n) = \begin{bmatrix} S_0 & S_1 & S_2 & \cdots & S_n \\ S_1 & S_2 & S_3 & \cdots & S_{n+1} \\ \cdot & \cdot & & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ S_n & S_{n+1} & S_{n+2} & & S_{2n} \end{bmatrix}. \tag{2.4}$$

It is easy to see that $H(n)$ is positive definite. To show this let $\mu^T = (\mu_0^T, \mu_1^T, \ldots \mu_n^T)$ be a $p(n+1)$ row vector with $\mu_i \in R^p$, $i = 0,\ldots,n$. Setting $\mu(\lambda,n) = \mu_0 + \mu_1\lambda + \ldots + \mu_n\lambda^n$, then using (2.1) and (2.4) one finds (see Delsarte et al. [12]),

$$\int_{-\infty}^{\infty} \mu^T(\lambda,n)d\rho(\lambda)\ \mu(\lambda,n) = \mu^T \mathrm{H(n)}\mu.$$

Since the integral is positive for all $\mu(\lambda,n) \neq 0$, $H(n) > 0$.

Let $R^{p\times p}(\lambda)$ be the set of $p \times p$ polynomial matrices with real coefficients and $R_n^{pxp}(\lambda)$ be the subset of $R^{pxp}(\lambda)$ containing polynomials of the form

$$Y(\lambda,n) = Y_0 + \lambda Y_1 \ldots + \lambda^n Y_n \qquad Y_i \in R^{pxp} \text{ (real } p\times p \text{ matrices).}$$

Setting $V^T(\lambda,n) = (I,\lambda I,\lambda^2 I,\ldots,\lambda^n I)$

$$V^T(n) = (0,0\ldots,I),$$

(here I is the $p\times p$ identity matrix) and $Y^T(n) = (Y_0^T, Y_1^T \ldots Y_n^T)$, one finds

$$Y(\lambda,n) = V^T(\lambda,n)Y(n). \tag{2.5}$$

With $\rho(\lambda)$ one can associate the symmetric functional $L{:}R^{pxp}(\lambda) \to \text{sym } R^{pxp}$ defined as follows[1]

$$L(Y(\lambda,n)) = \int_{-\infty}^{\infty} Y^T(\lambda,n)d\rho(\lambda)Y(\lambda,n) - 2\,\text{sym}(Y_n). \tag{2.6}$$

Here

$$\text{sym } A = \tfrac{1}{2}(A^T + A)$$

is the symmetric part of the real matrix A. Using (2.3), (2.4) and (2.5) in (2.6) gives

$$L(Y(\lambda,n)) = Y^T(n)\ H(n)Y(n) - 2\,\text{sym}(V^T\ (n)Y(n)).$$

Setting $X = H(n)^{1/2}\ Y(n) - H(n)^{-1/2}\ V(n)$ where $H(n)^{1/2}$ is the symmetric square root of $H(n)$ which exists since $H(n)$ is real and positive definite and substituting X into the above equation yields

$$L(Y(\lambda,n)) = X^T X - V^T\ (n)\ H(n)^{-1}\ V(n). \tag{2.7}$$

Thus for each degree n there exists a unique $W(\lambda,n) \in R_n^{pxp}(\lambda)$ given by the expression

$$W(\lambda,n) = V^T(\lambda,n)\ H(n)^{-1}\ V(n) \tag{2.8}$$

that minimizes (2.7) in the sense that

$$L(W(\lambda,n)) \leq L(X(\lambda,n)) \qquad X(\lambda,n) \in R_n^{pxp}(\lambda).$$

Note that $W(\lambda,n) = W_0 + W_1\lambda + \ldots + W_n\lambda^n$ and from (2.8)

$$W_n = V^T(n)\ H(n)^{-1}\ V(n) > 0 \tag{2.9}$$

thus $W(\lambda,n)$ is a matrix polynomial of degree n. Besides minimizing (2.7) $W(\lambda,n)$ is orthogonal to all $X \epsilon R_{n-1}^{pxp}(\lambda)$. To see this observe that for $X(\lambda,n)$ $\epsilon\ R_n^{pxp}\ (\lambda)$

$$\int W^T(\lambda,n)\ d\rho(\lambda)\ X(\lambda,n) = V^T(n)\ X(n) = X_n \tag{2.10}$$

where (2.5) and (2.8) have been used. Now $X(\lambda,n)\ \epsilon\ R_{n-1}^{pxp}(\lambda)$ if and only if $X_n = 0$, proving the result. In order to obtain the matrix polynomials orthonormal with respect to $d\rho(\lambda)$ set

$$P(\lambda,n) = K_n^T W^T(\lambda,n) \tag{2.11}$$

where

$$K_n K_n^T = W_n^{-1} = (V^T(n)\ H(n)^{-1}\ V(n))^{-1}. \tag{2.12}$$

Thus from (2.10)

$$\int_{-\infty}^{\infty} P(\lambda,n)\ d\rho(\lambda)\ P^T(\lambda,m) = \delta_{n,m} I. \tag{2.13}$$

Note that K_n^T and hence $P(\lambda,n)$ are defined only up to a left orthogonal factor.

As in the scalar case, matrix orthonormal polynomials on the real line satisfy the following three term recurrence formula

$$A(n+1)\ P(\lambda,n+1) + B(n)\ P(\lambda,n) + A(n)\ P(\lambda,n-1) = \lambda P(\lambda,n)$$
$$n = 0,1,2,\ldots \tag{2.14A}$$

$$P(\lambda,-1) = 0 \qquad P(\lambda,0) = I \tag{2.14B}$$

(here and throughout the rest of the paper it will be assumed without loss of generality that $\int d\rho(\lambda) = I$) where

$$B(n) = \int_{-\infty}^{\infty} \lambda P(\lambda,n)\ d\rho(\lambda)\ P^T(\lambda,n) \qquad n = 0,1,2\ldots \tag{2.14}$$

and

$$A(n+1) = \int_{-\infty}^{\infty} \lambda P(\lambda,n)\ d\rho(\lambda) P^T(\lambda n+1) \qquad n = 0,1,2\ldots \tag{2.15}$$
$$= K_n^T (K_{n+1}^T)^{-1} \quad .$$

From the above equations it is apparent that $B(n)$ is symmetric and $A(n+1)A(n+1)^T$ positive definite. Since K_n^T is defined only up to a left orthogonal factor it is convenient to choose it so that

$$A(n+1)^T = A(n+1) > 0. \tag{2.16}$$

Letting $Q(u,n)$ be a solution of (2.14A) one finds

$$Q^*(u,n)\,A(n+1)\,P(\lambda,n+1) - Q^*(u,n+1)\,A(n+1)\,P(\lambda,n) \tag{2.17}$$
$$= (\lambda-\bar{u})Q^*(u,n)P(\lambda,n) + Q^*(u,n-1)A(n)P(\lambda,n) - Q^*(u,n)A(n)\,P(\lambda,n-1)$$

where A^* means the hermitian conjugate of the matrix A. Two consequences of (2.17) are the Christoffel-Darboux formula,

$$P^*(u,n)A(n+1)\,P(\lambda,n+1) - P^*(u,n+1)\,A(n+1)P(\lambda,n)$$
$$= (\lambda - \bar{u}) \sum_{i-0}^{n} P^*(u,i)\,P(\lambda,i) \tag{2.18}$$

and the Wronskian theorem,

$$W[Q,P] = Q^*(\bar{\lambda},n)A(n+1)P(\lambda,n+1) - Q^*(\bar{\lambda},n+1)A(n+1)P(\lambda,n) \tag{2.19}$$

is independent of n. The first follows from (2.17) by setting $Q=P$ then iterating down using (2.14C). (2.19) follows from (2.17) by setting $\bar{u} = \lambda$.

3. The recurrence formula

We now turn things around and begin the recurrence formula. Given two sequences of $p \times p$ real symmetric matrices $\{A(n+1)\}_{n=0}$, and $\{B(n)\}_{n=0}$ with $A(n) > 0$ one can construct a set of matrix polynomials according to (2.14). We now assume that there exists an $A(\infty) = a(\infty)I$, $a(\infty)$ a positive scalar and $B(\infty) = b(\infty)I$, $b(\infty)$ real such that

$$\lim_{n\to\infty} |A(n) - A(\infty)| = 0$$

and

$$\lim_{n\to\infty} |B(n) - B(\infty)| = 0. \tag{3.1}$$

For the matrix norm used above it will be convenient to use the Hilbert-Schmidt norm, i.e.,

$$|B| = \{\sum_{ij} |b_{ij}|^2\}^{1/2} = |B^*|. \tag{3.2}$$

This norm has the useful property that for matrices A and B

$$|A\,B| \leq |A|\ |B|. \tag{3.3}$$

Considering (3.1) above, it is without loss of generality that one can take

$$A(\infty) = I \text{ and } B(\infty) = 0. \tag{3.4}$$

In analogy with scalar polynomials orthogonal on the real line (see [2]) we consider the following two equations.

$$P(\lambda,n) = A(n)^{-1}\{(zI - B(n-1))\, P(\lambda,n-1) + \frac{1}{z}\, \Psi(z,n-1)\}, \qquad \textbf{(3.5A)}$$

and

$$\Psi(z,n) = A(n)^{-1}\{\frac{1}{z}\, \Psi(z,n-1) + ((I - A(n)^2)z - B(n-1)P(\lambda,n-1)\}, \textbf{(3.5B)}$$

with

$$P(\lambda,0) = \Psi(z,0) = I \qquad \textbf{(3.5C)}$$

and

$$\lambda = z + \frac{1}{z}. \qquad \textbf{(3.5D)}$$

(Here we choose the branch so that $z \to 0$ as $\lambda \to \infty$.) The above equations can be recast into the more compact form

$$\Phi(z,n) = C(n)\, \Phi(z,n-1) \quad n = 1,2,\ldots \qquad \textbf{(3.6A)}$$

with

$$\Phi(z,n) = \begin{bmatrix} P(\lambda,n) \\ \Phi(z,n) \end{bmatrix}, \quad \Phi(z,0) = \begin{pmatrix} I \\ I \end{pmatrix} \qquad \textbf{(3.6B)}$$

and

$$C(n) = A(n)^{-1} \begin{bmatrix} (zI - B(n-1)) & \frac{1}{z}\, I \\ (I - A(n)^2)z - B(n-1) & \frac{1}{z}\, I \end{bmatrix}. \qquad \textbf{(3.6C)}$$

As a first application of the above recurrence formula the following is proved;

Lemma 3.1. *Let Φ_1 and Φ_2 be solutions of 3.6A and 3.6C for $m \le n \le l$, then*

$$\tilde{W}[\Phi_1^T, \Phi_2] \equiv \Phi_1^T(z,n) \begin{bmatrix} 0 & -I \\ I & 0 \end{bmatrix} \Phi_2(z,n)$$

is independent of n, $m \le n \le l$. Furthermore,

$$\tilde{W}[\Phi_1^T, \Phi_2]^T = -\tilde{W}[\Phi_2^T, \Phi_1].$$

Proof. Calculation.

To proceed further it is convenient at this point to introduce the techniques of Banach algebras. Let $v(n)$ be a real even function of n with the following properties

$$\begin{aligned} v(0) &= 1,\ v(n) \ge 1 \\ v(n) &\le v(n+1) \quad n > 0 \\ v(n) &\le v(m)v(n-m) \quad n,m \ge 0 \end{aligned}$$

and

$$\lim_{n\to\infty} (\nu(n))^{1/n} = R \geq 1. \tag{3.7}$$

Let h_ν be the space of functions such that if $g \in h_\nu$ then

$$g(z) = \sum_{n=-\infty}^{\infty} g(n)z^n \quad \frac{1}{R} \leq |z| \leq R \tag{3.8}$$

with

$$\| g \|_\nu = \sum_{n=-\infty}^{\infty} v(n) \, | g(n) | < \infty. \tag{3.9}$$

(Here the $g(n)$'s are complex numbers), let h_ν^+ and h_ν^- denote those functions in h_ν of the form

$$g^+(z) = \sum_{k=0}^{\infty} g(k)z^k \tag{3.10}$$

and

$$g^-(z) = \sum_{k=-\infty}^{0} g(k)z^k \tag{3.11}$$

respectively. If $\| g \|_\nu$ is the norm on h_ν, h_ν^+ and h_ν^- are commutative Banach algebras, h will denote the Banach algebra where $v(n) = 1$ for all n. It is obvious that

$$h_\nu \subset h. \tag{3.12}$$

Let H_ν be the class of $p \times p$ matrices such that if $G \in H_\nu$ then the entries in G are in h_ν and

$$G(z) = \sum_{k=-\infty}^{\infty} G(k)z^k \qquad \frac{1}{R} \leq |z| \leq R. \tag{3.13}$$

with

$$\| G \|_\nu = \sum_{k=-\infty}^{\infty} v(k) \| G(k) \| < \infty. \tag{3.14}$$

Here $G(k)$ is a $p \times p$ matrix. The matrix norm used here is the one introduced earlier, the Hilbert-Schmidt norm. H_ν^+ and H_ν^- will denote the collection of matrix functions in H_ν of the form

$$G^+(z) = \sum_{k=0}^{\infty} G(k)z^k \tag{3.15}$$

and

$$G^-(z) = \sum_{k=-\infty}^{0} G(k)z^k \tag{3.16}$$

respectively. Again it is clear that

$$H_v \subset H. \tag{3.17}$$

Note that if $G_+(z) \in H_v^+$, $(G_-(z) \in H_v^-)$ then $G_+(z)$ $(G_-(z))$ is analytic inside (outside) and continuous on the disk $|z| \le R$. $(|z| \ge R)$. Let H_0^+ (H_0^-) denote those functions analytic inside (outside) the unit circle and continuous on it except at the points $z = \pm 1$.

Returning to the recurrence formulas we denote the class M_v by those polynomial systems whose recurrence formula coefficients satisfy the following conditions

$$\sum_{n=1}^{\infty} n\, v(2n)\ \{ |I-A(n)^2| + |B(n-1)| \} < \infty \tag{3.18}$$

The class M_0 will be denoted by those systems whose coefficients satisfy

$$\sum_{n=1}^{\infty} \{ |I-A(n)^2| + B(n-1)| \} < \infty. \tag{3.19}$$

Lemma 3.2. *If* $P(\lambda,n)$ *is defined by* (3.5A), *then*

$$\|\alpha(n)^{-1} z^n P(\lambda,n)\|_v$$
$$\le p(n+1)\ v(2n)\ \exp\ \{p\, D \sum_{i=1}^{n} (i+1)\ v(2i)\ \gamma(i)\} \tag{3.20}$$

where

$$\alpha(i) = \overleftarrow{\prod_{j=1}^{i}} A(j) = A(i)\ A(i-1) \cdots A(1),\ \alpha(0) \equiv I \tag{3.21}$$

$$\gamma(i) = |\, I-A(i)^2\, | + |\, (B(i-1)\, |, \tag{3.22}$$

and

$$D = \max_i \{ |\, \alpha(i)^{-1}\, |\ \ |\, \alpha(i)\, | \}.$$

Furthermore

$$|\, z^n\, \alpha(n)^{-1}\, P(\lambda,n)\, | \le$$

$$\frac{pC(n+1)}{1+|\,1-z^2|\ (n+1)} \exp\ \{pCD \sum_{i=1}^{n} \frac{(i+1)\ \Delta(i,z)}{1+|1-z^2\,|\ (i+1)} \tag{3.23}$$

where

$$\Delta(i,z) = |I-A(i)^2|\ \ |z\,|^2 + |\, B(i\text{-}1)\, |\ \ |z| \tag{3.24}$$

and C is constant.

Proof. See [3].

The above lemma leads to;

Theorem 3.1. *If* $\Psi(z,n)$ *satisfies* (3.5B), *then*

$$\|\Psi^{\wedge}(z,n) - \Psi^{\wedge}(z,m)\|_v \le p\, D \sum_{i=m+1}^{n} i\, v(2i)\gamma(i)\exp\{p\, D \sum_{j=1}^{i-1} (j+1)v(2j)\gamma(j)\} \tag{3.25}$$

and

$$|\Psi^{\wedge}(z,n) - \Psi^{\wedge}(z,m))| \le pCD \sum_{i=m+1}^{n} \frac{i\Delta(i,z)}{1+|1-z^2|i} \exp\{pCd \sum_{j=1}^{i} \frac{(j+1)\Delta(j,z)}{1+|1-z^2|\ (j+1)}\} \tag{3.26}$$

where $\Psi^{\wedge}(z,n) = \alpha(n)^{-1} z^n\, \Psi(z,n)$.

Since H_v is a Banach space one has;

Corollary 3.1. *If* (3.18) *holds then there exists a function* $zf_+(z) \in H_v^+$ *such that*

$$\lim_{n\to\infty} \|\Psi^{\wedge}(z,n) - zf_+^{\wedge}(z)\|_v = 0.$$

If (3.19) *holds then there exists a function* $zf_+(z) \in H_0^+$ *such that*

$$\lim_{n\to\infty} |\Psi^{\wedge}(z,n) - zf_+^{\wedge}(z)| = 0$$

uniformly on compact subsets of the unit disk excluding the points $z = \pm 1$.

In the theory of scalar polynomials orthogonal on the real line the scalar analog of the matrix function $f_+(z) = \alpha(\infty)f_+^{\wedge}(z)$ plays an important role in such things as, the spectral factorization of the measure, asymptotic formulas, etc., Case [18], Geronimo and Case [3]. It will be shown below that $f_+(z)$ plays an analogous role in the theory of matrix orthogonal polynomials and we shall call it the matrix Jost function.

Before developing further the properties of $f_+(z)$ we introduce three other solutions of (3.6A) and (3.6B). Let

$$\Phi_1(z,n) = \begin{pmatrix} P_1(\lambda,n) \\ \Psi_1(z,n) \end{pmatrix} \quad n\ge 1,\ \Phi_1(z,1) = A(1)^{-1}\begin{pmatrix} I \\ I \end{pmatrix}. \tag{3.27}$$

In analogy with the previous solutions one has;

Theorem 3.2. *If* (3.18) *holds then there exists a function* $zf_{1+}^{\wedge}(z) \in H_v^+$ *such that*

$$\lim_{n\to\infty} \|\Psi_1^{\wedge}(z,n) - zf_{1+}^{\wedge}(z)\|_v = 0.$$

If (3.19) *holds then there exists a function* $z\hat{f}_{1+}(z) \in H_0^+$ *such that*

$$\lim_{n\to\infty} |\hat{\Psi}_1(z,n) - z\hat{f}_{1+}(z)| = 0$$

uniformly on compact subsets of the unit disk excluding the points $z = \pm 1$.

For the other two solutions, let

$$\Phi_+(z,n) = \begin{bmatrix} P_+(z,n) \\ \Psi_+(z,n) \end{bmatrix} \tag{3.28A}$$

and

$$\Phi_-(z,n) = \begin{bmatrix} P_-(z,n) \\ \Psi_-(z,n) \end{bmatrix} \tag{3.28B}$$

be solutions of (3.6A) and (3.6B) satisfying the following boundary conditions

$$\lim_{n\to\infty} P_+(z,n) = z^n I \lim_{n\to\infty} \Psi_+(z,n) = 0, \quad |z| \le 1 \tag{3.29}$$

and

$$\lim_{n\to\infty} P_-(z,n) = z^{-n}I, \; \lim_{n\to\infty} \Psi_-(z,n) = (1-z^2)z^{-n}I, \quad |z| \ge 1. \tag{3.30}$$

In what sense the limit exists is given in the following theorem.

Theorem (3.3). *If* (3.18) *holds then*

$$\|\hat{P}_+(z,n)\|_v \le p \exp\{p N \sum_{m=n+1}^{\infty} (m-n)\, v(2n-2m+2)\gamma(m+1)\} \; n \ge 0, \tag{3.31}$$

$$\|\hat{P}_+(z,n) - I\|_v \le p N \sum_{i=n+1}^{\infty} (i-n) v(2i-2n+2)\gamma(i+1)$$

$$\exp\{p N \sum_{m=i+1}^{\infty} (m-i) v(2m-2i+2)\gamma(m+1)\} \; n \ge 0, \tag{3.32}$$

and $P_+(z,n) \in H_v^+$.

If (3.19) *holds then*

$$|\hat{P}_+(z,n)| \le Cp \exp\{pN \sum_{m=n+1}^{\infty} C \frac{(m-n)\Delta(m+1,z)}{1+|1-z^2|(m-n)} \quad |z| \le 1, \; n \ge 0$$

$$|\hat{P}_+(z,n) - I| \le CpN \sum_{i=n+1}^{\infty} \frac{(i-n)\Delta(i+1,z)}{1+|1-z^2|\,(i-n)}$$

$$\exp\{cpN \sum_{m=i+1}^{\infty} \frac{(m-i)\Delta(m+1,z)}{1+|1-z^2|\,(m-i)}\}, \tag{3.33}$$

and $^{\wedge}P_+(z,n) \in H_0^+$. *Here* $^{\wedge}P_+(z,n) = z^{-n}\alpha(n+1,\infty)P_+(z,n)$, $\alpha(n,m)$, $= \overset{\rightarrow}{\prod_{n}^{m}} A(i) \equiv A(n)A(n+1)\cdots A(m)$, *and* $N = \max_{m,k} \{ |\alpha(k,\infty)|\ |\alpha(k+1,m)|\ |\alpha(m+1,\infty)^{-1}| \}$.

Proof. From the recurrence formula one has,

$$^{\wedge}P_+(z,n) = I - \sum_{i=n+1}^{\infty} {}^{\wedge}\Psi_+(z,i) \tag{3.34}$$

where $^{\wedge}\Psi_+(z,n) = z^{-n}\alpha(n+1,\infty)\Psi_+(z,n)$.

Iterating upwards the lower component of (3.28A) and (3.6C) then substituting the result into (3.34) gives

$$^{\wedge}P_+(z,n) = I + \sum_{m=n+1}^{\infty} \sum_{k=n+1}^{m} \alpha(k+1,\infty)\alpha((k+1,m)$$
$$\{I-A(m+1)^2)z-B(m)\}\alpha(m+1,\infty)^{-1}z^{2m-2k+1}\,{}^{\wedge}P_+(z,m). \tag{3.35}$$

Writing

$$^{\wedge}P_+(z,n) = \sum_{i=0}^{\infty} G(z,n)^i \tag{3.36}$$

with $G(z,n)^0 = I$ and

$$G(z,n)^{i+1} = \sum_{m=n+1}^{\infty} \sum_{k=n+1}^{m} \alpha(k+1,\infty)\alpha((k+1,m)$$
$$\{I-A(m+1)^2)z-B(m)\}\alpha(m+1,\infty)^{-1}z^{2m-2k+1}G(z,n)^i, \tag{3.37}$$

one finds $\|G(z,n)^0\|_v = p$ and by induction

$$\|G(z,n)^{i+1}\|_v \le \frac{p}{(i+1)!}\,(pN \sum_{m=n+1}^{\infty} (m-n)v(2m-2n+2)\gamma(m+1))^{i+1}. \tag{3.38}$$

(3.31) now follows by substituting the above equations into (3.35). (3.32) is obtained by substituting (3.31) into (3.35). (3.33) follows from arguments similar to those given above and from the fact that each of the $G(z,n)^i$ can be shown by induction to be matrix functions analytic inside the unit circle and continuous on it except perhaps at the points $z = \pm 1$.

To obtain the analytic properties of $P_-(z,n)$ note that if one eliminates $\Psi_+(z,n)$ and $\Psi_-(z,n)$ from the recurrence formula one finds that $\underset{\pm}{P}(z,n)$

satisfy (2.14A) with boundary conditions given by (3.29) or (3.30). Therefore

$$P_-(z,n) = P_+(1/z,n) \tag{3.39}$$

and the following corollary is a consequence of the recurrence formulas and the above equation.

Corollary (3.2). *If* (3.18) *holds then* $^\wedge\Psi_+(z,n) \in H_v^+$, $n \geq 0$, $^\wedge\Psi_-(z,n) \in H_v^-$, $n \geq 2$, *and* $^\wedge P_-(z,n) \in H_v^-$, $n \geq 0$. *If* (3.19) *holds then* $\Psi_+(z,n) \in H_0^+$, $n \geq 0$, $\Psi_-(z,n) \in H_0^-$, $n \geq 2$, *and* $P_-(z,n) \in H_0^-$, $n \geq 0$.

Utilizing the above analytic properties of Φ_+ and Φ_-, and the fact that they are linearly independent solutions of (3.6) except at the points $z = \pm 1$, one can write

$$\Phi(z,n) = \frac{\Phi_+(z,n)C(z) - \Phi_-(z,n)D(z)}{z - 1/z} \qquad z \neq \pm 1 \tag{3.40}$$

where $C(z)$ and $D(z)$ can be evaluated using Lemma 3.1. Thus

$$\begin{aligned} D(z) &= -\frac{1}{z}\tilde{W}[\Phi_+^T,\Phi] = \frac{1}{z}(P_+^T(z,n)\Psi(z,n) - \Psi_+^T(z,n)P(\lambda,n)) \\ &= \frac{1}{z}\lim_{n\to\infty} z^n\Psi(z,n) = f_+(z) \end{aligned} \tag{3.41}$$

and

$$C(z) = -\frac{1}{z}\tilde{W}[\Phi_-^T,\Phi] = z\lim_{n\to\infty} z^{-n}\Psi\{\frac{1}{z},n) = f_+(\frac{1}{z}) \equiv f_-(z). \tag{3.42}$$

Another useful representation of f_+ following from (2.19) is

$$f_+(z) = W[P_+^T,P] = P_+^T(z,-1) \tag{3.43}$$

with $A(0) \equiv I$.

Likewise the following formula for Φ_1 may be obtained

$$\Phi_1(z,n) = \frac{\Phi_+(z,n)f_{1-}(z) - \Phi_-(z,n)f_{1+}(z)}{z - 1/z} \qquad z \neq \pm 1 \tag{3.44}$$

where $f_{1+}(z) = P_+^T(z,0)$, and $f_{1-}(z) = f_{1+}(1/z)$. Using procedures similar to those above gives,

$$\Phi_+(z,n) = \Phi(z,n)f_{1+}^T(z) - \Phi_1(z,n)f_+^T(z) \tag{3.45}$$

where the fact that $\tilde{W}[\Phi^T,\Phi] = -z$ has been used. An analogous equation for Φ_- may be obtained by replacing + by − in the subscripts of f.

An immediate consequence of the above formulas is obtained by setting $n = -1$ in (3.40) yielding

$$f_+^T(z)f_-(z) = f_-^T(z)f_+(z). \tag{3.46}$$

The regions of validity of the above equations are determined by the analytic properties of their constituents.

4. The properties of the matrix $f_+(z)$

Let ℓ^2_{+p} denote the Hilbert space of vector functions $\omega = (\omega_1, \omega_2, \ldots, \omega_p)$ where $\omega_i \in \ell^2_+ (a \in \ell^2_+ \Leftrightarrow a = \{a(i)\}_{i=0}^\infty$ and $<a,a> = \sum_{i=0}^{\infty} |a(i)|^2 < \infty$.) The scalar product on ℓ^2_{+p} is the natural one $(f,g) = \sum_{l=1}^{p} (f_i, g_i)$, where (f_i, g_i) is the scalar product in the ℓ^2_+ sense. Assuming (3.1) holds the infinite dimensional matrix given by (2.14A) acting on vectors $y = (y_1, y_2, \ldots, y_p)$ with boundary condition $y(-1) = 0$, may be thought of as a bounded self-adjoint operator J on ℓ^2_{+p}. Thus the eigenvalues of J are real.

Suppose (3.19) holds and $\det f_+(z_0) = 0$, $|z| < 1$. Then there exists a $p \times 1$ vector μ such that $f_+^T(z_0)\mu = 0$, and (3.32) and (3.43) imply that the sequence $\{P_+(z_0,n)\mu\}_{n=-1}^{\infty} \in \ell^2_{+p}$. Thus the points z_i, $|z_i| < 1$ where $\det f_+(z_i) = 0$ are real and $\lambda_i = (z_i + 1/z_i)$ are eigenvalues of J.

Consider now the system solving (2.14A-C) with

$$A(n+1) = I, \; B(n) = 0 \quad n \geq N. \tag{4.1}$$

Then $zf_+(z) = z^N \Psi(z,N)$ is a matrix polynomial in z of order $2N$. Furthermore it is easy to see that $P_+(z,n)$ is a polynomial in z while $P_-(z,n)$ and $f_-(z)/z$ are polynomials in $1/z$. Since $P_+(z,n)$ and $P_-(z,n)$ are linearly independent except at $z = \pm 1$, the theory of second order difference equations (Atkinson [10]) says that any vector solution $y(\lambda,n)$ of (2.14A) and (4.1) can be written as a linear combination of $P_+(z,n)$ and $P_-(z,n)$ times appropriate constant vectors. From (3.32) and (3.39) it is clear that in order for the components of $y(z,n)$ to be square summable for $|z| < 1$, ($|\lambda_0| > 2$), it is necessary and sufficient that $y(\lambda_0,n) = P_+(z_0,n)a$, be a constant vector. Furthermore in order for λ_0 to be an eigenvalue of J^N (defined in the same manner as J except with restriction (4.1)), $y(\lambda_0,-1) = P_+(z_0,-1)a = 0$. This of course implies that $\det f_+(z_0) = 0$. Now letting $N \to \infty$ and using the fact that $z_0 f_+(z_0)$ and $P_+(z,n)$ converge uniformly for $|z| \leq 1$, one has

Theorem (4.1). *If (3.19) holds then the points z_i, $|z_i| < 1$ where $\det f_+(z_i) = 0$ are real and $\lambda_i = (z_i + \frac{1}{z_i})$ are the only eigenvalues of J for $|\lambda| > 2$. Furthermore $f_+(z)$ is nonsingular for $|z| = 1$, $z \neq \pm 1$.*

Proof. Only the last sentence remains to be proved. From (2.17) one has

$$P_+^*(z,n)A(n+1)P_+(z,n+1) - P_+^*(z,n+1)A(n+1)P_+(z,n) = (z-\tfrac{1}{z})I, \quad |z| = 1 \quad z \neq \pm 1. \tag{4.2}$$

Therefore if there exists a nonzero vector a such that $P_+(z,n)a = 0$ then $a^*P_+^*(z,n) = 0$. Multiplying the above equation on the right by a and on the left by a^* one finds that for $z \neq \pm 1$ $a^*a = 0$ which implies $a = 0$. Thus $\det P_+(z,-1) \neq 0$ $|z| = 1$, $z \neq \pm 1$.

Lemma (4.1). *If* (3.18) *holds then* $z = \pm 1$ *is not an eigenvalue of* J.

Proof. See Appendix A.

From the preceding arguments it is clear that the zeros of $\det z f_+(z)$ inside the unit circle are real with the only possible accumulation points being $z = \pm 1$. Therefore $f_+(z)^{-1}$ is analytic inside the unit circle except at the points where $\det f_+(z) = 0$, and it is continuous on the unit circle excepting the points $z = \pm 1$. The next lemma of Newton and Jost [19] will prove useful in investigating the points where $f_+(z)^{-1}$ is singular.

Lemma (4.2). *Let* $B(z)$ *be a square matrix, which is analytic in the circle* $|z| < 1$, *such that det* $B(0) = 0$ *and* $\det B(z) \neq 0$ *for* $0 < |z| < 1$. *Then the matrix* $B(z)^{-1}$ *has a simple pole at* $z = 0$ *if and only if the relations*

$$\begin{aligned} B(0)a &= 0, \\ B(0)b + B'(0)a &= 0 \end{aligned} \tag{4.3}$$

where a *and* b *are constant vectors or matrices imply that* $a = 0$.

Theorem (4.2). *If* (3.19) *holds all the singularities of* $f_+(z)^{-1}$ *for* $|z| < 1$ *are simple poles.*

Proof. This is a discrete analog of the proof given by Newton and Jost [19] for the matrix Schrödinger equation. Setting $Q^* = P_+^*$ in (2.17), iterating upwards then differentiating with respect to z using (3.5D) and setting $\bar{\mu} = \lambda$ one finds

$$\begin{aligned} &P_+^*(\bar{z},n-1)A(n)P_+'(z,n) - P_+^*(\bar{z},n)A(n)P_+'(z,n-1) \\ &= -(1-1/z^2)\sum_{i=n}^{\infty} P_+^*(\bar{z},i)P_+(z,i). \end{aligned} \tag{4.4}$$

Suppose that $\det P_+(z_0,-1) = 0$ then there exists a vector a such that $P_+(z_0,-1)a = 0$. Setting $n = 0$ in the above equation then multiplying through by a and a^* yields

$$a^*P_+^*(z_0,0)P_+'(z_0,-1)a = (1-1/z_0^2)\sum_{i=0}^{\infty} a^*P_+^*(z_0,i)P_+(z_0,i)a \neq 0. \quad \textbf{(4.5)}$$

Suppose now that a also satisfies the bottom equation in (4.3) with $B(0) = P_+(z_0,-1)$. Taking the hermitian transpose then multiplying on the right by $P_+(z_0,0)a$ gives

$$b^*P_+^*(z_0,-1)P_+(z_0,0)a + a^*P_+^{*\prime}(z_0,-1)P_+(z_0,0)a = 0. \quad \textbf{(4.6)}$$

But from (2.19)

$$P_+^*(z_0,-1)P_+(z_0,0) = P_+^*(z_0,0)P_+(z_0,-1). \quad \textbf{(4.7)}$$

Substituting this into (4.6) then comparing the result with (4.5) leads to the conclusion that $a = 0$, and implies through Lemma (4.2) that $f_+^T(z)^{-1}$, and therfore $f_+^{-1}(z)$ has a simple pole at z_0.

Having shown that the poles of $f_+^{-1}(z)$ for $|z| < 1$ are simple let us examine the residues of $f_+^{-1}(z)P_+^T(x,0)$ at the points of singularity. These residues will play an important role in the construction of the distribution function with respect to which the matrix polynomials are orthogonal. With that in mind we will use (3.5D) (and the branch mentioned below it) and consider f_+ and P_+ as functions of the variable λ. Consider again the system given by (2.14A) and (4.1). Assuming $f_+(\lambda)^{-1}$ has a pole at λ_i one finds

$$f_+(\lambda) = f_+(\lambda_i) + (\lambda-\lambda_i)f_+'(\lambda_i) + \cdots \quad \textbf{(4.8)}$$

and

$$f_+(\lambda)^{-1} = (\lambda-\lambda_i)^{-1}M_i + M_i^\circ + \cdots \quad \textbf{(4.9)}$$

Since

$$f_+(\lambda)f_+(\lambda)^{-1} = I = f_+(\lambda)^{-1}f_+(\lambda), \quad \lambda \neq \lambda_i$$

(4.8) and (4.9) imply

$$f_+M_i = M_i f_+ = 0 \quad \textbf{(4.10)}$$

and

$$f_+ M_i^\circ + f'_+M_i = M_i^\circ f_+ + M_i f_+' = I. \quad \textbf{(4.11)}$$

(The dependence on λ will be suppressed when there is no confusion. All differentiations in the rest of this section are with respect to λ.) Let E_i be a hermitian matrix which projects onto the null space of the matrix $f_+^T(\lambda_i)$ so that

$$f_+^T(\lambda_i)E_i = 0. \quad \textbf{(4.12)}$$

From (4.10) it is clear that the set of vectors M_i^Ta where a is an arbitrary vector coincides with the the null space of $f_+^T(\lambda_i)$. Consequently

$$E_i M_i^T = M_i^T. \tag{4.13}$$

From (4.4) one has

$$P_+^*(\lambda_i,-1)P_+'(\lambda_i,0)-P_+^*(\lambda_i,0)P_+'(\lambda_i,-1) = -\sum_{j=0}^{\infty} P_+^*(\lambda_i,j)P_+(\lambda_i,j) = -z_i. \tag{4.14}$$

Multiplying on the left by M_i and on the right by $M_i^* = M_i^T$ gives, using (3.43) and the fact that λ_i is real

$$M_iP_+^*(\lambda_i,0)P_+'(\lambda_i,-1)M_i^* = M_iZ_iM_i^*. \tag{4.15}$$

Taking the hermitian conjugate of (4.11) and substituting it into the above equation using (4.7) yields

$$\rho_i \equiv M_iP_+^T(\lambda_i,0)$$

$$= M_iP_+^*(\lambda_i,0) = M_iZ_iM_i^*. \tag{4.16}$$

Thus ρ_i is a nonnegative hermitian matrix. Furthermore multiplying the upper component of (3.45) on the right by M_i^* then its hermitian conjugate on the left by M_i and substituting the result into (4.16) one finds

$$\rho_i = \rho_i\sum_{j=0}^{\infty} P^*(\lambda_i,j)P(\lambda_i,j)\rho_i^*. \tag{4.17}$$

Now letting $N\to\infty$ yields;

Theorem (4.3). *If* (3.19) *holds then*

$$\rho_i = M_iP_+^T(\lambda_i,0) = M_if_{1+}(\lambda_i)$$

is the residue of $f_+^{-1}(\lambda)P_+^T(\lambda,0)$ *at* $\lambda = \lambda_i$ *where* λ_i *is a simple pole of* $f_+^{-1}(\lambda)$ *with* $|\lambda| > 2$. ρ_i *is non-negative with rank* $\leq p$, *and the orthonormal polynomials satisfy* (4.17).

Considering now the upper component of (3.40) one has

$$P(\lambda,n) = \frac{(P_+(z,n)S(z)-P_-(z,n))f_+(z)}{z-1/z} \qquad \begin{matrix} |z| = 1 \\ z \neq \pm 1 \end{matrix} \tag{4.18}$$

where

$$S(z) = f_-(z)f_+(z)^{-1} \tag{4.19}$$

is the discrete matrix analog of the scattering operator in quantum mechanics (see Agranovich and Marchenko [14], also Serebrjakov [9]). From (3.46)

$$S(z) = f_-(z)f_+(z)^{-1} = f_+^T(z)^{-1}f_-^T(z).$$

Therefore

$$S(z)S^*(z) = S^*(z)S(z) = I \qquad |z| = 1, \quad z \neq \pm 1 \tag{4.20}$$

and

$$\overline{S(z)} = S(\bar{z}) = S^*(z) \qquad |z| = 1, \quad z \neq \pm 1 . \tag{4.21}$$

$S(z)$ will play an important role in the discrete analog of the Marchenko equation (see Section 6). A consequence of (4.18) is;

Lemma (4.2). *Suppose* (4.1) *holds. If* $f_+(z)^{-1}$ *has a pole at* $z = \pm 1$, *it is simple.*

We now turn to the problem of determining the number of zeros of $\det f_+(z)$, $|z| < 1$.

Theorem (4.4). *If* (3.18) *holds then* J *has a finite number of eigenvalues.*[2]

Proof. Only the case where $v(n) = 1$ for all n will be considered as all other cases follow. Instead of considering the matrix associated with (2.14A), J, consider $J' = J - 2I$, with the boundary condition $y(-1) = 0$. J' is now a negative self-adjoint operator acting on $\ell^2_{+p} = \ell^2_p(0,\infty)$. Consider the operators J^1 and J^2 represented by

$$J^1 = \begin{pmatrix} B(O)-2I & A(1) & & & & \\ A(1) & B(1)-2I & A(2) & & & \\ \cdot & \cdot & \cdot & & & \\ & \cdot & \cdot & & & \\ & & A(n-2) & B(n-2)-2I & A(n-1) \\ & & & A(n-1) & B(n-1)-2I \end{pmatrix}$$

with $y(-1) = y(n) = 0$, and

$$J^2 = \begin{pmatrix} B(n+1)-2I & A(n+1) & & \\ A(n+1) & B(n+2)-2I & A(n+3) & \\ & A(n+3) & B(n+3)-2I & A(n+4) \\ & \cdot & \cdot & \cdot \end{pmatrix}$$

with $y(n) = 0$. J^1 and J^2 act on $\ell_p^2(0,n-1)$ and $\ell_p^2(n+1,\infty)$ respectively. The proof now follows from arguments similar to those given in [3]. (Note the squaring argument in [3] is not needed.)

5. Construction of the distribution function

Returning to the system satisfying (2.14) and (4.1) consider the following integral,

$$\Gamma = \frac{2}{\pi}\int_0^\pi P(\lambda,m)f_+(z)^{-1}f_+(z)^{-1}P^*(\lambda,n)\sin^2\theta d\theta. \qquad \lambda = z+1/z \qquad (5.1)$$

$$z = e^{i\theta}$$

This integral is well defined since, by Lemma 4.2, $(z-1/z)f_+(z)^{-1}$ is analytic $|z| = 1$. Using (3.46), the upper component of (3.40), and (3.39) one may recast the above equation into, (see [3]),

$$\Gamma = -\frac{1}{2\pi i}\oint P(\lambda,m)f_+(z)^{-1}P_+^T(z,n)(1-z^{-2})dz.$$

The residues of the above integral are at $z = 0$ and at the simple poles of $f_+(z)^{-1}$. These residues can be evaluated using the recurrence formulas and Theorem (4.3). This leads to

Theorem (5.1). *Given the system of polynomials* $\{P(\lambda,n)\}$ *satisfying* (2.14) *and* (4.1) *one has*

$$\int_{-\infty}^{\infty} P(\lambda,n)d\rho(\lambda)P^T(\lambda,m) = \delta_{n,m}$$

where

$$d\rho(\lambda) = \begin{cases} \sigma(\theta)d\lambda & \lambda = 2\cos\theta \quad 0 \le \theta \le \pi \\ \sum_{i=1}^{L} \rho_i\delta(\lambda-\lambda_i)d\lambda & \lambda \text{ not as above } L < \infty \end{cases}$$

with

$$\sigma(\theta)d\lambda = f_+(e^{i\theta})^{-1}f_+^*(e^{i\theta})^{-1}\sin\theta d\lambda$$

and

$$\rho_i = M_iP_+^T(\lambda_i,0).$$

M_i is the residue of $f_+(\lambda)^{-1}$ at λ_i, $|\lambda_i| > 2$. ρ_i is non-negative with rank $\le p$.

Since (here we introduce the superscript N to denote the system satisfying (4.1))

$$\rho^N(\lambda) = \int_{-\infty}^{\lambda} d\rho^N(x) \leq \int_{-\infty}^{\infty} d\rho^N(x) = I$$

and $\rho^N(\lambda)$ is nondecreasing, one can use the matrix Helly-Bray theorem (Atkinson [10]) to obtain a limiting distribution function. (This holds even if only (3.1) is assumed.)

Lemma (5.1). *If* (3.18) *holds, then*

$$\lim_{N\to\infty} |(z-1/z)B^N(z)f_+^N(z)^{-1} - (z-1/z)B(z)f_+^{-1}(z)| = 0$$

uniformly on compact subsets of the open unit disk. Here

$$B^N(z) = \prod_{i=1}^{L^N} \frac{z_i^N - z}{1 - zz_i^N} \frac{|z_i^N|}{z_i^N}$$

$$B(z) = \prod_{i=1}^{L} \frac{z_i - z}{1 - zz_i} \frac{|z_i|}{z_i}$$

where $z_i^N(z_i)$ *are the singular points* of $f_+^N(z)^{-1}$ $(f_+(z)^{-1})$, $|z| < 1$.

Proof. Since det $zf_+^N(z)$ and det $zf_+(z)$ are analytic functions for $|z| < 1$, the result follows from Hurwitz's theorem, and Theorem (4.4).

Theorem (5.2). *Given* (3.18) *one has* $(z-1/z)B(z)f_+(z)^{-1} \in H_2^{p\times p}$ *(the Hilbert space of* $p\times p$ *matrix functions analytic in the open unit disk).*

Proof. Using the fact that $(z-1/z)B^N(z)f_+^N(z)^{-1}$ is analytic on the unit disk (see Lemma (4.2)) one has for $0 \leq r \leq 1$ and for all N

$$Q = ½\pi\int_{-\pi}^{\pi} |(z-1/z)B^N(z)f_+^N(z)^{-1}|^2 d\theta \qquad z = re^{i\theta} \tag{5.2}$$

$$\leq \frac{1}{2\pi}\int_{-\pi}^{\pi} |(e^{i\theta} - e^{-i\theta})f_+^N(e^{i\theta})^{-1}|^2 d\theta$$

$$= \frac{4}{\pi}\int_0^{\pi} |f_+^N(e^{i\theta})^{-1}|^2 \sin^2\theta d\theta$$

$$= 4 \text{ trace} \int_0^{\pi} f_+^N(e^{i\theta})^{-1} f_+^{N*}(e^{i\theta})^{-1} \frac{\sin^2\theta}{\pi} d\theta$$

$$\leq 4 \text{ trace} \int_{-\infty}^{\infty} d\rho^N(\lambda) = 4p$$

Thus (5.2) remains the same for fixed $r < 1$ when $(z-1/z)B^N(z)f_+^N(z)^{-1}$ is replaced by its limiting function $(z-1/z)B(z)f_+(z)^{-1}$ proving the theorem.

Corollary (5.2). $(e^{i\theta}-e^{-i\theta})f_+^{-1}(e^{i\theta}) \in L_2^{p\times p}(T)$ *(the Hilbert space of* $p\times p$ *matrix functions on the unit circle).*

As a consequence of the theorem above and the previous techniques one has

Theorem (5.3). *Theorem* (5.1) *holds when the coefficients in the recurrence formula satisfy* (3.18).

6. The discrete Marchenko equation

Returning again to the upper component of (3.40) one has

$$P(\lambda,n)f_+^{-1}(z)(z-1/z) = -P_-(z,n)+P_+(z,n)S(z) \qquad |z| = 1,\ z \neq \pm 1$$

where $S(z)$ is the discrete scattering matrix.

From (4.18) and (4.19) one sees that $S(z)$ is unitary and continuous for $|z| = 1$, $z \neq \pm 1$. Assuming (3.18) holds it is a consequence of Theorem (3.3) that $P_+(z,n) \in H^+$ and can be written as

$$P_+(z,n) = \sum_{\ell=n}^{\infty} A(n,\ell)z^{\ell}. \tag{6.1}$$

Substituting this into the previous equation multiplying by z^{m-1} then integrating around the unit circle gives

$$\frac{1}{2\pi i}\oint P(\lambda,n)f_+(z)^{-1}\, z^m(z-1/z)\frac{dz}{z}$$

$$= -\frac{1}{2\pi i}\sum_{\ell=n}^{\infty} A(n,\ell)\oint z^{m-\ell}\frac{dz}{z} + \sum_{\ell=n}^{\infty} A(n,\ell)\oint S(z)z^{\ell+m}\frac{dz}{z}, \tag{6.2}$$

Here (3.39) and the fact that $(z-1/z)\, f_+(z)^{-1} \in L^{p\times p}(T)$ have been used. Note that because of the unitarity of $S(z)$ and the summability of $A(n,\ell)$ the interchange of orders of summation and integration is justified.

For $m \geq n \geq 0$ one can evaluate the L.H.S. using the residue theorem

$$\frac{1}{2\pi i}\oint P(\lambda,n)f_+^{-1}(z)(z-1/z)z^m\frac{dz}{z}$$

$$= \sum_{i=1}^{L} P(\lambda_i,n)M_i z_i^m - A^T(n,n)^{-1}\delta_{n,m} \tag{6.3}$$

where the last term on the R.H.S. comes from the residue at $z = 0$. To recast

$P(\lambda_i,n)M_i$ multiply the upper component of (3.45) on the right by projection operator E_i (see (4.12)), which gives

$$P_+(\lambda_i,n)E_i = P(\lambda_i,n)P_+(\lambda_i,0)E_i.$$

Replacing M_i in (4.16) by E_i^*, taking the hermitian transpose, then substituting the result into the above equation using (4.13) yields

$$P_+(\lambda_i,n)E_i = P(\lambda_i,n)M_iT_i \tag{6.4}$$

where

$$T_i = E_iZ_iE_i + I - E_i$$

is positive definite. Multiplying (6.4) on the right by $T_i^{-1}E_i$ gives

$$P_+(\lambda_i,n)N_i = P(\lambda_i,n)M_i$$

with

$$N_i = E_iT_i^{-1}E_i \tag{6.5}$$

a non-negative hermitian matrix. Substituting the above into (6.3) then combining the result with (6.2) gives the discrete analogs of the matrix Marchenko equations,

$$a(n,m) + \omega(n+m) + \sum_{\ell=n+1}^{\infty} a(n,\ell)\omega(\ell+m) = 0 \qquad m > n \geq 0$$

and

$$A(n,n)^{-1}A^T(n,n)^{-1} = 1 + \omega(2n) + \sum_{\ell=n+1}^{\infty} a(n,\ell)\omega(\ell+n) \qquad n \geq 0$$

where

$$a(n,m) = A(n,n)^{-1}A(n,m) \tag{6.6}$$

and

$$\omega(k) = -\frac{1}{2\pi i}\oint S(z)z^k\frac{dz}{z} + \sum_{i=1}^{L} N_iz_i^k \quad . \tag{6.7}$$

For $n = -1$ the L.H.S. of (6.2) vanishes and one has

$$a(-1,m) + \omega^1(m-1) + \sum_{\ell=0}^{\infty} a(-1,\ell)\omega^1(\ell+m) = 0 \qquad m \geq -1 \tag{6.8}$$

where

$$\omega^1(k) = -\frac{1}{2\pi i}\oint S(z)z^k \frac{dz}{z}.$$

With the above equations one can now prove

Theorem (6.1). *Given* (3.18) *the following holds*

a) $\frac{z}{1-z}E_1P_+(z,-1) \in H_v^+$ *if* $\det P_+(1,-1) = 0$

b) $\frac{z}{1+z}E_{-1}P_+(z,-1) \in H_v^+$ *if* $\det P_+(-1,-1) = 0$

c) $\frac{z((1-z)E_{-1} + (1+z)E_1)P_+(z-1)}{1-z^2} \in H_v^+$ *if* $\det P_+(\pm 1,-1) = 0$

E_1 and E_{-1} are hermitian matrices which project onto the null spaces of $P_+(1,-1)^T$ and $P_+(-1,-1)^T$ respectively, (i.e. $P_+(1,-1)^TE_1 = 0$).

Proof. See [3] p. 478 and [14] lemma 5.62.

Lemma 6.1. *Given* (3.18) *then $Df_+(e^{i\theta})$ is nonsingular where*

a) $D=I$ if $\det f_+(\pm 1) \neq 0$

b) $D = (I-E_1+\frac{E_1}{1-z})$ *if* $\det f_+(1) = 0$

c) $D = (I-E_{-1}+\frac{E_{-1}}{1+z})$ *if* $\det f_+(-1) = 0$

d) $D = (I-E_{-1}+\frac{E_{-1}}{1+z})(I-E_1+\frac{E_1}{1-z})$ *if* $\det f_+(\pm 1) = 0$

Proof. See Agranovich and Marchenko [14], Lemma 5.6.3.

Lemma 6.2. *Given* (3.18) *then* $D^{-1}\frac{\sigma}{\sin\theta}D^{*-1} \in H_v$ and $\sin\theta\, D^*\sigma^{-1}D \in H_v$.

Proof. This follows from Theorems (5.4) and (6.1), Lemma (6.1), Corollary (3.2) and the Weiner-Levy theorem.

7. Integral representations

Having the distribution function one can find explicit integral representations for the matrix function used in the earlier sections. For example

$$P_1(\lambda,n) = \int_{-\infty}^{\infty} \frac{P(\lambda,n) - P(\lambda',n)}{\lambda - \lambda'} d\rho(\lambda') \quad n \geq 0. \tag{7.1}$$

Substituting this equation into the upper component of (3.45), then multiplying on the left by $P(\lambda,n)^{-1}$ and letting $n \to \infty$ yields

$$f_{+1}^T(z) = \int_{-\infty}^{\infty} \frac{d\rho(\lambda')}{\lambda - \lambda'} f_+^T(z) \qquad \lambda = z + 1/z \quad |z| < 1 \tag{7.2}$$

$$\lambda' = z' + 1/z' \quad \lambda \notin \text{support } d\rho(\lambda).$$

Now using (7.2) in (3.45) gives

$$P_+(z,n) = \int_{-\infty}^{\infty} \frac{P(\lambda',n)}{\lambda - \lambda'} d\rho(\lambda') f_+^T(z) \qquad |z| < 1, \quad \lambda = z + 1/z \tag{7.3}$$

$$\lambda \notin \text{support } d\rho(\lambda).$$

Notes

1. For analogous results on matrix polynomials on the unit circle, see Delsarte et al. [12].
2. This proof is adopted from Agranovich and Marchenko [14]; see also Serebrjakov [9].

Appendix A

Here we prove

Lemma (4.1). *If (3.18) holds then $z = \pm 1$ is not an eigenvalue of J.*

Proof. Only $z = 1$ is considered as the other case follows in exactly the same manner. Consider the polynomials $\{P_m(\lambda,n)\}$ and $\{\Psi_m(z,m)\}$ satisfying (3.6A) for $n \geq m$ with $P_m(\lambda,0) = \Psi_m(z,0) = I$. From the recurrence one can show [3] that,

$$P_m^\wedge(z,n) = \frac{1 - z^{2n+2}}{1 - z^2} I + \sum_{i=1}^{n} \alpha_m(i-1)^{-1}\{[I - A(i+m)^2]z^2 \left(\frac{1-z^{2n-2i}}{1-z^2}\right) \tag{A.1}$$

$$-B(i+m-1)\left(\frac{1-z^{2n-2i+2}}{1-z^2}\right) z\alpha_m(i-1) P_m^\wedge(z,i-1)$$

where $P_m^\wedge(z,n) = z^n \alpha_m(n)^{-1} P_m(z,n)$ and

$$\alpha_m(n) = \overleftarrow{\sum_{J=m+1}^{n}} A(j).$$

Using successive approximations leads to,

$$|z^n \alpha_m(n)^{-1} P_m(\lambda,n)| \leq p(n+1) C \exp\{pCD_m \sum_{i=1}^{n} (i+1)\gamma(i+m)\}$$

where D_m is defined in (3.22). Thus substituting this equation into A.1 yields

$$\left| \frac{P_m(1,n)}{n+1} - I \right| \leq \sum_{i=1}^{n} D_m \, Cp\left(\frac{n-i+1}{n+1}\right)(i+1)\,|\gamma(i+m)|$$

$$x \exp\{p\, CD_m \sum_{j=1}^{i} (j+1)\gamma(j+m)|.$$

If (3.18) holds then for large enough m the R.H.S. can be made less than ½ for all $n > m$. Therefore $P_+(z,n)$ and $P_m(\lambda,n)$ are two linearly independent solutions of 2.14A for $n \geq m$ at $z = 1$. Since all solutions of 2.14A and B can be written in terms of these two for $n \geq m$ and since neither $p_+(1,n)c$ nor $P_m(z,n)d$ are summable where c and d are arbitrary nonzero vectors one can conclude that J does not have an eigenvalue at $z = 1$.

References

1. K. M. Case and M. Kac, A discrete version of the inverse scattering problem, J. Math. Phy. 14 (1973), 544-603.
2. J. S. Geronimo and K. M Case, Scattering theory and polynomials orthogonal on the unit circle, J. Math. Phys. 20 (1979), 299-310.
3. J. S. Geronimo and K. M. Case, Scattering theory and polynomials orthogonal on the real line, Trans. Amer. Math. Soc. 258 (1980), 467-494.
4. J. S. Geronimo. A relation between the coefficients in the recurrence formula and the spectral function for orthogonal polynomials, Trans. Amer. Math. Soc. 260 (1980), 65-82.
5. J. S. Geronimo, An upper bound on the number of eigenvalues of an infinite dimensional Jacobi matrix, Accepted J. Math. Phys.
6. G. V. Guseinov, The determination of an infinite Jacobi matrix from the scattering data, Soviet Math. Dokl. 17, (1976), 596-600.
7. H. Dym and A. Iacob, Applications of factorization and Toeplitz operators to inverse problems. Toeplitz memorial conference.
8. P. G. Nevai, "Orthogonal polynomials." Memoirs Amer. Math. Soc. 18, 1979.
9. V. P. Serebrjakov, The inverse problem of scattering theory for difference equations with matrix coefficients. Soviet Math. Dokl. 21, (1980), 148.
10. F. V. Atkinson, Discrete and Continuous Boundary Problems. Academic Press, New York, 1964.
11. W. G. Christian, A. G. Law, W. F. Martens, A. L. Mullikin and M. B. Sledd. Solutions of initial-value problems for some infinite chains of harmonic oscillators. J. Math. Phys. 17 (1976), 146.

12. P. Delsarte, Y. V. Genin, Y. G. Kamp. Orthogonal polynomial matrices on the unit circle. I.E.E.E. Transactions on Circuits and Systems. CAS-25, (1978), 149.
13. J. S. Geronimo, Matrix orthogonal polynomials on the unit circle. J. Math. Phys, (1981), 1359.
14. Z. S. Agranovich and V. A. Marchenko, "The Inverse Problem of Scattering Theory," Gordon and Breach, New York, 1963.
15. Ju. M. Berezanskii. "Expansions in eigenfunctions of self-adjoint operators," Trans. Math. Mono. Amer. Math. Soc. 17 (1968).
16. M. G. Krein. Infinite *J*-matrices and a matrix moment problem, Dokl. Akad. Nauk. SSSR. 69 (1949), 125.
17. M. Rosenberg, The square-integrability of matrix-valued functions with respect to a non-negative hermitian measure, Duke Math. J. 31 (1964), 291.
18. K. M. Case, Orthogonal polynomials from the viewpoint of scattering theory, J. Math. Phys. 15 (1974), 2166.
19. R. G. Newton and R. Jost, The construction of potentials from the *S*-matrix for systems of differential equations, Il. Nuovo Cimento 1, (1955) 590.