Spectral Theory for Matrix Orthogonal Polynomials on the Unit Circle

by

LIYANG ZHANG

Mihai Stoiciu, Advisor

A thesis submitted in partial fulfillment
of the requirements for the
Degree of Bachelor of Arts with Honors
in Mathematics

WILLIAMS COLLEGE
Williamstown, Massachusetts

April 30, 2012

Abstract

In this thesis, we first introduce the classical theory of orthogonal polynomials on the unit circle and its corresponding matrix representations - the GGT representation and the CMV representation. We briefly discuss the Sturm oscillation theory for the CMV representation. Motivated by Schulz-Baldes' development of Sturm oscillation theory for matrix orthogonal polynomials on the real line, we study matrix orthogonal polynomials on the unit circle. We prove a connection between spectral properties of GGT representation with matrix entries, CMV representation with matrix entries with intersection of Lagrangian planes. We use this connection and Bott's theory on intersection of Lagrangian planes to develop a Sturm oscillation theory for GGT representation with matrix entries and CMV representation with matrix entries.

Liyang Zhang

Williams College

Williamstown, Massachusetts 01267, USA

Acknowledgements

I am extremely fortunate to have Professor Mihai Stoiciu as my thesis advisor. I am grateful for not only his guidance and immense patience but also the invaluable advice on my career and life. I would like to thank Professor Cesar Silva for being my second reader. I also owe great debt to Professor Steven Miller, who opened the world of mathematics to me and gave me support throughout my undergraduate career.

Contents

1 Introduction

Orthogonal polynomials on the unit circle (OPUC) were developed by Szegö around 1920. in the The theory of Matrix orthogonal polynomials on the unit circle was first developed in the study of prediction theory [**4**] [**5**]. But the foundation paper about the subject was written by Delsarte, Genin, Kamp in [**2**] and Youla, Kazanjian in [**3**].

Given a nontrivial matrix-valued measure on the unit circle $d\mu$, we can define inner products:

$$\langle\langle f, g\rangle\rangle_R = \int f(x)^\dagger d\mu g(x) \tag{1.1}$$

$$\langle\langle f, g\rangle\rangle_L = \int g(x) d\mu f(x)^\dagger. \tag{1.2}$$

This gives us two sets of matrix orthogonal polynomials on the unit circle (MOPUC). Similar to the theory of OPUC, we can define Verblunsky coefficients, GGT representation and CMV representation from it. In the study of eigenvalues of the CMV representation for OPUC, Stoiciu [**6**] described oscillation theory using the relation that the zeros of the paraorthogonal polynomials are exactly the eigenvalues of the corresponding CMV truncation. But the corresponding theory is not yet proven.

Hermann Schulz-Baldes developed a Sturm oscillation theory for matrix orthogonal polynomials on the real line in [**10**]. He used the results of symplectic geometry developed by Bott in [**7**], which connects the spectra of a matrix to the intersection of Lagrangian planes. In this note, we use the same results of Bott to develop a Sturm oscillation theory for GGT representation with matrix entries and CMV representation with matrix entries.

2 Matrix Orthogonal Polynomial on the Unit Circle

2.1 Orthogonal Functions

We first establish the orthogonality of functions. Let μ be a measure.

Definition 2.1. *An orthonormal set of functions $\{\phi_i\}_{i\in I} \subseteq L^2(S)$ where I is an arbitrary counting set is defined by the relations*

$$\langle \phi_n, \phi_m \rangle = \int_S \overline{\phi_n}\phi_m d\mu = \delta_{nm}, \qquad n, m \in I \tag{2.1}$$

One quick observation is that functions in the orthonormal set are linearly independent. Furthermore if the measure μ is supported only on a finite number of points N then I is necessarily finite and $|I| \leq N$. From this point we will call measures supported on a set of finite number of points *trivial measures*.

Theorem 2.2. *Let $\{f_i\}_{i\in I}$ be a set of linearly independent functions in $L^2(S)$ for some space S where I can be finite or infinite. Then an orthonormal set $\{\phi_i\}_{i\in I}$ exists.*

Not only we can prove the theorem but we can explicitly construct the set of orthonormal functions by the Gram-Schmidt process. We first construct a set of orthogonal functions $\{\Phi_i\}_{i\in I}$ as follows:
First we define a projection operator by

$$\mathrm{proj}_{\mathbf{u}}(\mathbf{v}) = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle}\mathbf{u}. \tag{2.2}$$

Then let

$$\Phi_1 = f_1, \tag{2.3}$$

$$\Phi_k = f_k - \sum_{j=1}^{k-1} \mathrm{proj}_{\Phi_j}(f_k). \tag{2.4}$$

Notice that this process will not stop because the set of functions are linearly independent. Finally we can let

$$\phi_i = \frac{\Phi_i}{\|\Phi_i\|}. \tag{2.5}$$

Definition 2.3. *Let $\{\phi_i\}_{i\in I}$ be a given orthonormal set in $L^2(S, d\mu)$. To an arbitrary function $f \in L^2(S, d\mu)$, it corresponds the formal Fourier expansion*

$$c_1\phi_1 + c_2\phi_2 + \cdots \tag{2.6}$$

where

$$c_i = \langle \phi_i, f \rangle = \int_S \overline{\phi_i} f d\mu \tag{2.7}$$

are the Fourier coefficients of f with respect to the orthonormal set $\{\phi_i\}$.

Now we are ready to study orthogonal polynomials.

Definition 2.4. *Let μ be a nontrivial measure on some space S so that for all n,*

$$c_n = \int_S |x|^n d\mu < \infty. \tag{2.8}$$

If we orthogonalize the set of linearly independent functions $\{1, x, x^2, \ldots\}$ by the Gram-Schmidt process we obtain a set of monic orthogonal polynomials $\{\Phi_i\}$ and a set of orthonormal polynomials $\{\phi_i\}$. Both Φ_n and ϕ_n have degree precisely n.

2.2 Orthogonal Polynomials on the Unit Circle

In this thesis, our main focus is the orthogonal polynomials with measure defined on the unit circle. Let μ be an arbitrary nontrivial measure on $\partial\mathbb{D}$. We usually normalize the measure so that $\mu(\partial\mathbb{D}) = 1$. We will assume this condition from here. By the Gram-Schmidt process we have the monic orthogonal polynomials $\Phi_n(z)$ or $\Phi_n(z; d\mu)$ for the measure μ and the orthonormal polynomials φ_n given by

$$\varphi_n(z) = \kappa_n z^n + \text{lower order terms} \tag{2.9}$$

where

$$\kappa_n = \|\Phi_n\|^{-1}. \tag{2.10}$$

One important feature of these polynomials is that they satisfy a recursion relationship called Szegö recursion. To present the recursion, we begin with investigating some properties for orthogonal polynomials on the unit circle.

Let P be a monic polynomial of degree n, then Φ_n and $\Phi_n - P$ are orthogonal. So we have

$$\langle \Phi_n, \Phi_n \rangle = \langle \Phi_n, \Phi_n \rangle + \langle \Phi_n, P - \Phi_n \rangle = \langle \Phi_n, P \rangle \tag{2.11}$$

and in particular

$$\|\Phi_n\|^2 = \langle z^n, \Phi_n \rangle = \langle z\Phi_{n-1}, \Phi_n \rangle = \kappa_n^{-2}. \tag{2.12}$$

Also notice that $z \in \partial\mathbb{D}$ implies

$$\|z\Phi_n\| = \|\Phi_n\| \quad \text{and} \quad \langle zf, g \rangle = \langle f, z^{-1}g \rangle. \tag{2.13}$$

On $L^2(\partial\mathbb{D}, d\mu)$, we define the reverse polynomial f^* for f to be

$$f^*(z) = (R_n f)(z) = z^n \overline{f(z)}. \tag{2.14}$$

This map is anti-unitary so we have

$$\langle R_n f, R_n g\rangle = \langle g, f\rangle. \tag{2.15}$$

Now we are ready to prove an important lemma.

Lemma 2.5. *Up to a constant, $\Phi_n^*(z)$ is the unique polynomial of degree (at most) n which is orthogonal to $z, z^2, \ldots, z^n$, that is, $\langle P, z^j\rangle = 0$, $j = 1, 2, \ldots, n$, and deg$(P) \le n \Rightarrow P = c\Phi_n^*$. Moreover, $\Phi_n^*(0) = 1$ and*

$$\|\Phi_n\|^2 = \|\Phi_n^*\|^2 = \int \Phi_n^*(e^{i\theta})d\mu(e^{i\theta}). \tag{2.16}$$

Proof. Since the operator R_n is anti-unitary and Φ_n is the unique polynomial orthogonal to $1, \ldots, z^{n-1}$, $\Phi_n^* = R_n\Phi_n$ is the unique polynomial orthogonal to $\{R_n z^j\}_{j=0}^{n-1} = \{z^{n-j}\}_{j=0}^{n-1} = \{z^j\}_{j=1}^{n}$. For the second statement consider

$$\begin{aligned}
\Phi_n^*(0) &= R_n(z^n + \text{lower order terms})(0) \\
&= (1 + (\text{higher order terms}))(0) \\
&= 1.
\end{aligned} \tag{2.17}$$

Finally,

$$\|\Phi_n\|^2 = \langle \Phi_n, z^n\rangle = \langle R_n z^n, R_n\Phi_n\rangle = \langle 1, \Phi_n^*\rangle = \int \Phi_n^*(e^{i\theta})d\mu(e^{i\theta}). \tag{2.18}$$

□

Theorem 2.6. *(Szegö Recursion). For any nontrivial probability measure μ on $\partial\mathbb{D}$, we can define a sequence $\{\alpha_n\}_{n=1}^{\infty}$ of numbers in $\mathbb{D}$ so that*

$$\Phi_{n+1}(z) = z\Phi_n(z) - \overline{\alpha_n}\Phi_n^*(z) \tag{2.19}$$

$$\Phi_{n+1}^*(z) = \Phi_n^*(z) - \alpha_n z\Phi_n(z). \tag{2.20}$$

Moreover,

$$\begin{aligned}
\|\Phi_{n+1}\|^2 &= (1 - |\alpha_n|^2)\|\Phi_n\|^2 \qquad (2.21)\\
&= \prod_{j=0}^{n}(1 - |\alpha_j|^2). \qquad (2.22)
\end{aligned}$$

Proof. Notice we have $\langle z\Phi_n, z^k\rangle = \langle \Phi_n, z^{k-1}\rangle = 0$ for $k = 1, \ldots, n$. So $z\Phi_n$ is orthogonal to $z, z^2, \ldots, z^n$ and Φ_{n+1} is orthogonal to $1, z, z^2, \ldots, z^n$. Combining we find $\Phi_{n+1}(z) - z\Phi_n(z)$ is a polynomial of degree n which is orthogonal to $z, z^2, \ldots, z^n$. By the previous lemma, $\Phi_{n+1}(z) - z\Phi_n(z) = -\overline{\alpha}_n \Phi_n^*(z)$ for some complex number $-\overline{\alpha}_n$. Apply R_n to both sides of this equation we get $\Phi_{n+1}^*(z) = \Phi_n^*(z) - \alpha_n z\Phi_n(z)$.

Now consider the equality we obtained:

$$\Phi_{n+1}(z) + \overline{\alpha_n}\Phi_n^*(z) = z\Phi_n(z). \tag{2.23}$$

Taking the norm on both sides gives

$$\|\Phi_{n+1}(z) + \overline{\alpha_n}\Phi_n^*(z)\|^2 = \|z\Phi_n(z)\|^2. \tag{2.24}$$

Notice that the cross term is 0 since terms on the left are orthogonal. So

$$\|\Phi_{n+1}(z)\|^2 + \|\overline{\alpha_n}\Phi_n^*(z)\|^2 = \|z\Phi_n(z)\|^2. \tag{2.25}$$

This simplifies to

$$\|\Phi_{n+1}\|^2 = (1 - |\alpha_n|^2)\|\Phi_n\|^2 \tag{2.26}$$

$$= \prod_{j=0}^{n}(1 - |\alpha_j|^2). \tag{2.27}$$

Notice both $\|\Phi_{n+1}\|^2$ and $\|\Phi_n\|^2$ are positive. So $1 - |\alpha_n|^2 > 0$ which completes the proof. □

Definition 2.7. *From now on, we will call these α_n's the Verblunsky coefficients associated to a measure.*

Theorem 2.8. *There is a bijection between nontrivial probability measures on the unit circle and sequences of complex numbers $\{\alpha_n\}_{n\geq 0}$ with $|\alpha_n| < 1$ for any n.*

Example 2.9. *One nontrivial example of orthogonal polynomials on the unit circle is the degree one Bersterin-Szegö polynomials with parameter $\zeta = re^{i\varphi} \in \mathbb{D}$. The measure on the unit circle is given by $\omega(\theta)d\mu(e^{i\theta}) = \frac{1-|\zeta|^2}{|1-\zeta e^{i\theta}|^2}d\mu(e^{i\theta})$. A simple* *calculation gives*

$$\begin{aligned}
\Phi_n(z) &= z^n - \bar{\zeta} z^{n-1}. \\
\varphi_n(z) &= \frac{1}{\sqrt{1-|\zeta|^2}}(z^n - \bar{\zeta}^{n-1}).
\end{aligned}$$

The Verblunsky coefficients are

$$\alpha_0 = \zeta \qquad \alpha_j = 0 (j \geq 1).$$

2.3 Matrix Representation

Orthogonal polynomials on the real line, similar to orthogonal polynomials on the unit circle have a recursion relationship. But it is a much simpler three term recursion:

$$xp_n(x) = a_{n+1}p_{n+1}(x) + b_{n+1}p_n(x) + a_n p_{n-1}(x). \tag{2.28}$$

Here p_n are the orthogonal polynomials on the real line. This relationship suggests a clear matrix representation, the Jacobi matrix:

$$J_\mu = \begin{pmatrix} b_1 & a_1 & 0 & 0 & \cdots \\ a_1 & b_2 & a_2 & 0 & \cdots \\ 0 & a_2 & b_3 & a_3 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix} \tag{2.29}$$

The corresponding matrix representation for orthogonal polynomials on the unit circle is less clear due to a more complicated recursion relationship. The first representation is the GGT representation. It is named after Geronimus, Gragg, and Yeplzaev. It is defined as follows:

Definition 2.10. *Given a nontrivial probability measure μ on $\partial\mathbb{D}$, we define the GGT matrix $\{\mathcal{G}_{k\ell}(d\mu)\}_{0\le k,\ell<\infty}$ by*

$$\mathcal{G}_{k\ell} = \langle \varphi_k, z\varphi_\ell \rangle \qquad 0 \le k, \ell < \infty. \tag{2.30}$$

Definition 2.11. *Given a sequence $\{\alpha_n\}_{n=0}^{\infty}$ of numbers in $\mathbb{D}$, we set $\alpha_{-1} = -1$ and define the matrix $\{\mathcal{G}_{k\ell}\{\alpha_n\}_{n=0}^{\infty}\}_{0\le k,\ell<\infty}$ by*

$$\mathcal{G} = \begin{cases} -\bar{\alpha}_\ell \alpha_{k-1} \prod_{j=k}^{\ell-1} \rho_j & 0 \le k \le \ell \\ \rho_\ell & k = \ell + 1 \\ 0 & k \ge \ell + 2 \end{cases} \tag{2.31}$$

where $\rho_\ell = (1 - |\alpha_\ell|^2)^{1/2}$.

We state without proof the fact that

Proposition 2.12. $\mathcal{G}(d\mu) = \mathcal{G}(\{\alpha_n(d\mu)\}_{n=0}^{\infty})$

By calculation, we find the GGT representation has the form

$$\mathcal{G}(\{\alpha_n\}_{n=0}^{\infty}) = \begin{pmatrix} \bar{\alpha}_0 & \bar{\alpha}_1\rho_0 & \bar{\alpha}_2\rho_0\rho_1 & \bar{\alpha}_3\rho_0\rho_1\rho_2 & \cdots \\ \rho_0 & -\bar{\alpha}_1\alpha_0 & -\bar{\alpha}_2\alpha_0\rho_1 & -\bar{\alpha}_3\alpha_0\rho_1\rho_2 & \cdots \\ 0 & \rho_1 & -\bar{\alpha}_2\alpha_1 & -\bar{\alpha}_3\alpha_1\rho_2 & \cdots \\ 0 & 0 & \rho_2 & -\bar{\alpha}_3\alpha_2 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}. \tag{2.32}$$

Notice that the GGT representation has a disadvantage of rows having infinite non-zero entries. Here we introduce the CMV representation for orthogonal polynomials on the unit circle. The name is given because of the work by Cantero, Moral and Velázquez [**1**]. It has the attractive property that both columns and rows have only finite non-zero entries. We first will define the CMV basis which is the building block for the CMV representation of the orthogonal polynomials on the unit circle. Let $\mathcal{H}_{(k,\ell)}$ be the space of Laurent polynomials spanned by $\{z^j\}_{j=k}^{\ell}$ and $P_{(k,\ell)}$ the orthogonal projection onto $\mathcal{H}_{(k,\ell)}$ in the Hilbert space $L^2(\partial\mathbb{D}, d\mu)$. Define

$$\mathcal{H}^{(n)} = \begin{cases} \mathcal{H}_{(-k,k)} & n = 2k \\ \mathcal{H}_{(-k,k+1)} & n = 2k+1 \end{cases} \tag{2.33}$$

and $P^{(n)}$ = projection onto $\mathcal{H}^{(n)}$.

Now we define $\chi_n^{(0)}$ by

$$\chi_n^{(0)} = \begin{cases} z^{-k} & n = 2k \\ z^{k+1} & n = 2k+1 \end{cases}. \tag{2.34}$$

We obtain the CMV basis by executing Gram-Schmidt to the $\chi_n^{(0)}$.

$$\chi_n = \frac{(1 - P^{(n-1)})\chi_n^{(0)}}{\|(1 - P^{(n-1)})\chi_n^{(0)}\|}. \tag{2.35}$$

Note that this process will not stop since our assumption of nontriviality of $d\mu$ ensures that $\chi_n^{(0)}$ are linearly independent. Thus $\|(1 - P^{(n-1)})\chi_n^{(0)}\| \neq 0$.

We can also use the ordered set $1, z^{-1}, z, z^{-2}, z^2, \ldots$ to define an alternate

CMV basis using the same procedures. So we define

$$\widetilde{\mathcal{H}}^{(n)} = \begin{cases} \mathcal{H}_{(-k,k)} & n = 2k \\ \mathcal{H}_{(-k-1,k)} & n = 2k+1 \end{cases} \tag{2.36}$$

and $\widetilde{P}^{(n)}$ = projection onto $\widetilde{\mathcal{H}}^{(n)}$.

$$x_n^{(0)} = \begin{cases} z^k & n = 2k \\ z^{-k-1} & n = 2k+1 \end{cases} \tag{2.37}$$

and the alternate CMV basis is obtained by Gram-Schmidt the $x_n^{(0)}$:

$$x_n = \frac{(1-\widetilde{P}^{(n-1)})x_n^{(0)}}{\|(1-\widetilde{P}^{(n-1)})x_n^{(0)}\|}. \tag{2.38}$$

Lemma 2.13. *Define $\sigma_n = \chi_{2n}, \tau_n = \chi_{2n-1}, s_n = x_{2n}, t_n = x_{2n-1}$ with σ_n, s_n labelled by $n = 0, 1, 2, \ldots$ and τ_n, t_n by $n = 1, 2, \ldots$. Then we have*

$$\begin{aligned} \tau_n &= z^{-n+1}\varphi_{2n-1} \\ \sigma_n &= z^{-n}\varphi_{2n}^* \\ t_n &= z^{-n}\varphi_{2n-1}^* \\ s_n &= z^{-n}\varphi_{2n} \end{aligned}$$

and

$$x_n(z) = \overline{\chi_n(1/\overline{z})}. \tag{2.39}$$

Proof. Using our new notation, we notice

$$\varphi_{2n-1} = \frac{(1-P_{(0,2n-2)})z^{2n-1}}{\|(1-P_{(0,2n-2)})z^{2n-1}\|}. \tag{2.40}$$

Since

$$z^\ell P_{(k,m)} z^{-\ell} = P_{(k+\ell,m+\ell)} \tag{2.41}$$

we have

$$z^{-n+1}\varphi_{2n-1} = \frac{[z^{-n+1}(1-P_{(0,2n-2)})z^{n-1}]z^n}{\|[z^{-n+1}(1-P_{(0,2n-2)})z^{n-1}]z^n\|} \tag{2.42}$$

$$= \frac{(1-P_{(-n+1,n-1)})z^n}{\|(1-P_{(-n+1,n-1)})z^n\|} \tag{2.43}$$

$$= \frac{1-P^{(2n-2)}\chi^{(0)}_{2n-1}}{\|1-P^{(2n-2)}\chi^{(0)}_{2n-1}\|} \tag{2.44}$$

$$= \chi_{2n-1} \tag{2.45}$$

$$= \tau_n. \tag{2.46}$$

The proofs of the other are similar and hence omitted. The second statement is immediate from the first statement. □

The CMV representation, $\mathcal{C}(d\mu)$ is

$$\mathcal{C}_{ij}(d\mu) = \langle \chi_i, z\chi_j \rangle \tag{2.47}$$

where $\{\chi_j\}_{j=0}^{\infty}$ is the CMV basis. The alternate CMV representation $\widetilde{\mathcal{C}}(d\mu)$ is the matrix

$$\mathcal{C}_{ij}(d\mu) = \langle x_i, zx_j \rangle \tag{2.48}$$

where $\{x_j\}_{j=0}^{\infty}$ is the alternate CMV basis. The elements of the matrix $\mathcal{C}$ can be determined by the following calculation:

$$\langle \sigma_{j-1}, z\sigma_j \rangle = \rho_{2j-1}\rho_{2j-2} \qquad \langle \sigma_j, z\sigma_j \rangle = -\bar{\alpha}_{2j}\alpha_{2j-1}$$

$$\langle \tau_j, z\sigma_j \rangle = -\alpha_{2j-2}\rho_{2j-1} \qquad \langle \tau_{j+1}, z\sigma_j \rangle = -\alpha_{2j-1}\rho_{2j}$$

$$\langle \sigma_{j-1}, z\tau_j \rangle = \bar{\alpha}_{2j-1}\rho_{2j-2} \qquad \langle \sigma_j, z\tau_j \rangle = \bar{\alpha}_{2j}\rho_{2j-1}$$

$$\langle \tau_j, z\tau_j \rangle = -\bar{\alpha}_{2j-1}\alpha_{2j-2} \qquad \langle \tau_{j+1}, z\tau_j \rangle = \rho_{2j}\rho_{2j-1}$$

where $\rho_\ell = (1 - |\alpha_\ell|^2)^{1/2}$. Thus

$$\mathcal{C} = \begin{pmatrix} \bar{\alpha}_0 & \bar{\alpha}_1\rho_0 & \rho_1\rho_0 & 0 & 0 & \dots \\ \rho_0 & -\bar{\alpha}_1\alpha_0 & -\rho_1\alpha_0 & 0 & 0 & \dots \\ 0 & \bar{\alpha}_2\rho_1 & -\bar{\alpha}_2\alpha_1 & \bar{\alpha}_3\rho_2 & \rho_3\rho_2 & \dots \\ 0 & \rho_2\rho_1 & -\rho_2\alpha_1 & -\bar{\alpha}_3\alpha_2 & -\rho_3\alpha_2 & \dots \\ 0 & 0 & 0 & \bar{\alpha}_4\rho_3 & -\bar{\alpha}_4\alpha_3 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}. \tag{2.49}$$

Similar calculations show that

$$\widetilde{\mathcal{C}} = \begin{pmatrix} \bar{\alpha}_0 & \bar{\alpha}_1\rho_0 & \rho_1\rho_0 & 0 & 0 & \dots \\ \rho_0 & -\bar{\alpha}_1\alpha_0 & -\rho_1\alpha_0 & 0 & 0 & \dots \\ 0 & \bar{\alpha}_2\rho_1 & -\bar{\alpha}_2\alpha_1 & \bar{\alpha}_3\rho_2 & \rho_3\rho_2 & \dots \\ 0 & \rho_2\rho_1 & -\rho_2\alpha_1 & -\bar{\alpha}_3\alpha_2 & -\rho_3\alpha_2 & \dots \\ 0 & 0 & 0 & \bar{\alpha}_4\rho_3 & -\bar{\alpha}_4\alpha_3 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}. \tag{2.50}$$

We can rewrite $\mathcal{C}$ as a product of matrices which is useful for computations.

Theorem 2.14. *With the pairs of bases $\{\chi_j\}_{j=0}^\infty$ and $\{x_j\}_{j=0}^\infty$ define $\mathcal{M}_{ij}(d\mu) = \langle x_i, \chi_j\rangle$ and $\mathcal{L}_{ij}(d\mu) = \langle \chi_i, zx_j\rangle$. Then we have*

$$\mathcal{C} = \mathcal{L}\mathcal{M}; \tag{2.51}$$

$$\widetilde{\mathcal{C}} = \mathcal{M}\mathcal{L}. \tag{2.52}$$

Proof.

$$\begin{aligned}
\mathcal{C}_{ij} &= \langle \chi_i, z\chi_j \rangle && (2.53)\\
&= \langle z^{-1}\chi_i, \chi_j \rangle && (2.54)\\
&= \sum_{\ell=0}^{\infty} \langle z^{-1}\chi_i, x_\ell \rangle \langle x_\ell, \chi_j \rangle && (2.55)\\
&= \sum_{\ell=0}^{\infty} \langle \chi_i, z x_\ell \rangle \langle x_\ell, \chi_j \rangle && (2.56)\\
&= \sum_{\ell=0}^{\infty} \langle \mathcal{L}_{i\ell}\mathcal{M}_{\ell j} && (2.57)\\
&= (\mathcal{L}\mathcal{M})_{ij}. && (2.58)
\end{aligned}$$

The proof for $\widetilde{\mathcal{C}}$ is identical. □

2.4 Sturm Oscillation for CMV matrix

In this section, we focus on the CMV representation. We want to study the eigenvalues of CMV matrix but since it is infinite we need to find a valid truncation. Given a measure $d\mu$ on the unit circle with the corresponding Verblunsky coefficients $\{\alpha_n\}$, orthogonal polynomials can be calculated using the Szegö recursion:

$$\Phi_{k+1} = z\Phi_k(z) - \bar{\alpha}_k \Phi_k^*(z) \qquad k \geq 0, \Phi_0 = 1. \tag{2.59}$$

Notice that if we make $|\alpha_N| = 1$, then $\rho_N = 0$. So the CMV matrix decouples. We write $\mathcal{C}^N$ as the upper half of the decoupled CMV matrix. At the same time we define the N-th paraorthogonal polynomial,

$$\Phi_N(z, d\mu, \beta) = z\Phi_{N-1}(z, d\mu) - \bar{\beta}\Phi_{N-1}^*(z, d\mu) \tag{2.60}$$

We state the following important theorem

Theorem 2.15. *The eigenvalues of the matrix $\mathcal{C}^N$ are the zeros of the N-th paraorthogonal polynomial.*

Therefore we can consider the complex function

$$B_N(z) = \frac{\beta z \Phi_{N-1}(z)}{\Phi_{N-1}^*(z)} \tag{2.61}$$

which has the property that $\Phi_N(e^{i\theta}) = 0$ if and only if $B_N(e^{i\theta}) = 1$. Since the truncated CMV matrix is a unitary matrix, all the eigenvalues are on the unit circle. So we can only consider z of the form $e^{i\theta}$. At the same time $|\beta| = 1$ and $|\Phi_n| = |\Phi_n^*|$, we have $B_N(z)$ lies on the unit circle. Thus we can rewrite the expression as follows:

$$B_N(e^{i\theta}) = e^{i\eta_N(\theta)}. \tag{2.62}$$

Here $\eta_N(\theta)$ is the Prüfer phase. The location of the eigenvalues of the truncated CMV matrix is signaled by a 2π increase in the Prüfer phase. This is the Sturm Osciallation theory for OPUC. We will develop a similar theory for Matrix Orthogonal polynomials on the unit circle.

3 Matrix Orthogonal Polynomials on the Unit Circle

3.1 Matrix Orthogonal Polynomials on the Unit Circle

In an effort to study higher dimensional systems, matrix valued orthogonal polynomials were developed by Delsarte-Genin-Kamp [**2**] and Youla-Kazanjian [**3**]. Here I will give a brief introduction. Let $\mathcal{M}_l$ be the ring of all $\ell \times \ell$ complex-valued matrices. We will use $\mathcal{P}$ to donate the set of polynomials in $z \in \mathbb{C}$ with coefficients in $\mathcal{M}_l$. A matrix-valued measure μ on $\mathbb{C}$ is the assignment of a positive semi-definite $\ell \times \ell$ matrix $\mu(X)$ to every Borel set X which is countably additive. We usually normalize it by requiring

$$\mu(\partial\mathbb{D}) = \mathbf{1}. \tag{3.1}$$

We can associate to any such measure a scalar measure

$$\mu_{\mathrm{tr}}(X) = \mathrm{Tr}(\mu(X)) \tag{3.2}$$

(the trace normalized by $\mathrm{Tr}(\mathbf{1}) = \ell$). At the same time μ_{tr} is normalized by $\mu_{\mathrm{tr}}(\partial\mathbb{D}) = \ell$.

Since the matrix value for the measure is positive semi-definite, we can apply the Radon-Nikodym theorem to the matrix elements of μ. So there is a positive semi-definite matrix function $M_{ij}(x)$ such that

$$d\mu_{ij}(x) = M_{ij}(x)d\mu_{\mathrm{tr}}(x). \tag{3.3}$$

Clearly we have

$$\mathrm{Tr}(M(x)) = 1 \tag{3.4}$$

for $d\mu_{\mathrm{tr}}$. Conversely, and scalar measure with $\mu_{\mathrm{tr}}(\partial\mathbb{D}) = l$ and positive semi-definite matrix valued function M obeying define a matrix-valued measure.

Here we will introduce an important notation. **We will use $A^\dagger$ to denote the conjugate transpose of the matrix A.** Given $\ell \times \ell$ matrix valued functions

f, g, we define the $\ell \times \ell$ matrix $\langle\langle f, g \rangle\rangle_R$ by

$$\begin{aligned} \langle\langle f, g \rangle\rangle_R &= \int f(z)^\dagger d\mu(z) g(x) \\ &= \int f(z)^\dagger M(z) g(z) d\mu_{\mathrm{tr}}(z); \end{aligned} \tag{3.5}$$

and $\langle\langle f, g \rangle\rangle_L$ by

$$\langle\langle f, g \rangle\rangle_L = \int g(z) d\mu(z) f(z)^\dagger. \tag{3.6}$$

We also define

$$\|f\|_R = \left(\mathrm{Tr}\langle\langle f, f \rangle\rangle_R\right)^{1/2}, \|f\|_L = \left(\mathrm{Tr}\langle\langle f, f \rangle\rangle_L\right)^{1/2}. \tag{3.7}$$

It satisfies these properties:

$$\begin{aligned} \langle\langle f, g\alpha \rangle\rangle_R &= \langle\langle f, g \rangle\rangle_R \alpha, \\ \langle\langle f\alpha, g \rangle\rangle_R &= \alpha^\dagger \langle\langle f, g \rangle\rangle_R, \\ \langle\langle f, g \rangle\rangle_R^\dagger &= \langle\langle g, f \rangle\rangle_R, \\ \langle\langle f, g \rangle\rangle_L &= \langle\langle g^\dagger, f^\dagger \rangle\rangle_R, \\ \langle\langle f, g \rangle\rangle_L^\dagger &= \langle\langle g, f \rangle\rangle_L, \\ \|f\|_L &= \|f^\dagger\|_R. \end{aligned}$$

Now we define monic matrix polynomials Φ_n^R, Φ_n^L by applying Gram-Schmidt orthogonalization to the set $\{\mathbf{1}, z\mathbf{1}, z^2\mathbf{1}, \ldots\}$ using the inner product $\langle\langle \cdot, \cdot \rangle\rangle_R, \langle\langle \cdot, \cdot \rangle\rangle_L$ respectively. For example, Φ_n^R is the unique matrix polynomial of order n with

$$\langle\langle z^k \mathbf{1}, \Phi_n^R \rangle\rangle_R = 0 \qquad k = 0, 1, \ldots, n-1. \tag{3.8}$$

For a matrix polynomial P_n of degree n, we define the reverse polynomial $P_n^*(z)$:

$$P_n^*(z) = z^n P_n(1/\overline{z})^\dagger. \tag{3.9}$$

It's easy to show that

$$(P_n^*)^* = P_n \tag{3.10}$$

Now we define the normalized orthogonal matrix polynomial by

$$\varphi_0^L = \varphi_0^R = \mathbf{1}, \qquad \varphi_n^L = \kappa_n^L \Phi_n^L \text{ and } \varphi_n^R = \Phi_n^R \kappa_n^R \tag{3.11}$$

where the κ's are defined according to the normalization condition

$$\langle\langle \varphi_n^R, \varphi_m^R \rangle\rangle_R = \delta_{nm}\mathbf{1} \qquad \langle\langle \varphi_n^L, \varphi_m^L \rangle\rangle_L = \delta_{nm}\mathbf{1} \tag{3.12}$$

along with

$$\kappa_{n+1}^L (\kappa_n^L)^{-1} > \mathbf{0} \quad \text{and} \quad (\kappa_n^R)^{-1} \kappa_{n+1}^R > \mathbf{0}. \tag{3.13}$$

We now define

$$\rho_n^L = \kappa_n^L (\kappa_{n+1}^L)^{-1} \quad \text{and} \quad \rho_n^R = (\kappa_{n+1}^R)^{-1} \kappa_n^R. \tag{3.14}$$

Similar to the scalar orthogonal polynomials we have the Szegö recursion:

Theorem 3.1. *(a) For suitable matrices $\alpha_n^{L,R}$, the following is true:*

$$z\varphi_n^L - \rho_n^L \varphi_{n+1}^L = (\alpha_n^L)^\dagger \varphi_n^{R,*}, \tag{3.15}$$

$$z\varphi_n^R - \varphi_{n+1}^R \rho_n^R = \varphi_n^{L,*} (\alpha_n^R)^\dagger. \tag{3.16}$$

(b) The matrices α_n^L and α_n^R are equal and will be denoted by α_n. We now call these α_n's Verblunsky coefficients.

(c) $\rho_n^L = (\mathbf{1} - \alpha_n^\dagger \alpha_n)^{1/2}$ and $\rho_n^R = (\mathbf{1} - \alpha_n \alpha_n^\dagger)^{1/2}$.

The proof for (a) is very similar to the proof in the scalar case so it will be omitted. We will give a proof for part (b) and (c):

Proof. (b)

$$\begin{aligned}
(\alpha_n^L)^\dagger &= \mathbf{0} + (\alpha_n^L)^\dagger \mathbf{1} \\
&= \langle\!\langle \varphi_n^{R,*}, \rho_n^L \varphi_{n+1}^L \rangle\!\rangle_L + (\alpha_n^L)^\dagger \langle\!\langle \varphi_n^R, \varphi_n^R \rangle\!\rangle_R \\
&= \langle\!\langle \varphi_n^{R,*}, \rho_n^L \varphi_{n+1}^L \rangle\!\rangle_L + (\alpha_n^L)^\dagger \langle\!\langle \varphi_n^{R,*}, \varphi_n^{R,*} \rangle\!\rangle_L \\
&= \langle\!\langle \varphi_n^{R,*}, \rho_n^L \varphi_{n+1}^L + (\alpha_n^L)^\dagger \varphi_n^{R,*} \rangle\!\rangle_L \\
&= \langle\!\langle \varphi_n^{R,*}, z\varphi_n^L \rangle\!\rangle_L \\
&= \langle\!\langle z\varphi_n^R, \varphi_n^{L,*} \rangle\!\rangle_R^\dagger \\
&= \langle\!\langle \varphi_{n+1}^R \rho_n^R + \varphi_n^{L,*} (\alpha_n^R)^\dagger, \varphi_n^{L,*} \rangle\!\rangle_R^\dagger \\
&= \langle\!\langle \varphi_{n+1}^R \rho_n^R, \varphi_n^{L,*} \rangle\!\rangle_R^\dagger + \langle\!\langle \varphi_n^{L,*} (\alpha_n^R)^\dagger, \varphi_n^{L,*} \rangle\!\rangle_R^\dagger \\
&= \mathbf{0} + \langle\!\langle \varphi_n^{L,*}, \varphi_n^{L,*} (\alpha_n^R)^\dagger \rangle\!\rangle_R \\
&= \langle\!\langle \varphi_n^{L,*}, \varphi_n^{L,*} \rangle\!\rangle_R (\alpha_n^R)^\dagger \\
&= \langle\!\langle \varphi_n^L, \varphi_n^L \rangle\!\rangle_L (\alpha_n^R)^\dagger \\
&= (\alpha_n^R)^\dagger.
\end{aligned}$$

(c)

$$\begin{aligned}
\mathbf{1} &= \langle\!\langle z\varphi_n^L, z\varphi_n^L \rangle\!\rangle_L \\
&= \langle\!\langle \rho_n^L \varphi_{n+1}^L + \alpha_n^\dagger \varphi_n^{R,*}, \rho_n^L \varphi_{n+1}^L + \alpha_n^\dagger \varphi_n^{R,*} \rangle\!\rangle_L \\
&= \rho_n^L \langle\!\langle \varphi_{n+1}^L, \varphi_{n+1}^L \rangle\!\rangle_L (\rho_n^L)^\dagger + \alpha_n^\dagger \langle\!\langle \varphi_n^{R,*}, \varphi_n^{R,*} \rangle\!\rangle_L \alpha_n \\
&= (\rho_n^L)^2 + \alpha_n^\dagger \langle\!\langle \varphi_n^R, \varphi_n^R \rangle\!\rangle_R \alpha_n \\
&= (\rho_n^L)^2 + \alpha_n^\dagger \alpha_n.
\end{aligned}$$

Note here that $\mathbf{1} - \alpha_n^\dagger \alpha_n$ is a positive semi-definite matrix. So it has a unique positive semi-definite square root and this is how the square root is defined here. The proof for the second equality is similar. □

Similar to the scalar orthogonal polynomials on the unit circle, there is a one-to-one correspondence between matrix-measures and matrix Verblunsky coefficients.

Theorem 3.2. *Any sequence $\{\alpha_j\}_{j=0}^\infty \in \mathbb{D}^\infty$ is the sequence of Verblunsky coefficients of a unique measure.*

3.2 Matrix Representation

In this section, we will develop the GGT and CMV representations for matrix orthogonal polynomial on the unit circle. The GGT representation is defined by

$$\mathcal{G}_{k\ell} = \langle\langle \varphi_k^R, z\varphi_\ell^R \rangle\rangle_R. \tag{3.17}$$

Of course we can define the matrix $\mathcal{G}$ by using $\langle\langle \cdot, \cdot \rangle\rangle_L$. But the theory is identical and hence it will be omitted here. A straight forward computation shows that

$$\mathcal{G}_{k\ell} = \begin{cases} -\alpha_{k-1}\rho_k^L \rho_{k+1}^L \cdots \rho_{\ell-1}\alpha_\ell^\dagger & 0 \le k \le \ell \\ \rho_\ell^R & k = \ell + 2 \\ \mathbf{0} & k \ge \ell + 2 \end{cases}. \tag{3.18}$$

In matrix form, we have

$$\mathcal{G} = \begin{pmatrix} \alpha_0^\dagger & \rho_0^L \alpha_1^\dagger & \rho_0^L \rho_1^L \alpha_2^\dagger & \rho_0^L \rho_1^L \rho_2^L \alpha_3^\dagger & \cdots \\ \rho_0^R & -\alpha_0 \alpha_1^\dagger & -\alpha_0 \rho_1^L \alpha_2^\dagger & -\alpha_0 \rho_1^L \rho_2^L \alpha_3^\dagger & \cdots \\ \mathbf{0} & \rho_1^R & -\alpha_1 \alpha_2^\dagger & -\alpha_1 \rho_2^L \alpha_3^\dagger & \cdots \\ \mathbf{0} & \mathbf{0} & \rho_2^R & -\alpha_2 \alpha_3^\dagger & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{3.19}$$

Similar to OPUC, this does not have finite non-zero row entries. From the definition of $\mathcal{G}$, we observe the following facts:

Proposition 3.3. *The matrix $\mathcal{G}$ is unitary in the following sense:*

$$\sum_{k=0}^{\infty} \mathcal{G}_{kn}^\dagger \mathcal{G}_{km} = \sum_{k=0}^{\infty} \mathcal{G}_{nk} \mathcal{G}_{mk}^\dagger = \delta_{nm} \mathbf{1}. \tag{3.20}$$

Proposition 3.4. *Let $|z| \le 1$. Then, for every $m > 0$,*

$$\sum_{n=0}^{\infty} \varphi_n(z) \mathcal{G}_{nm} = z\varphi_m(z), \qquad \sum_{n=0}^{\infty} \mathcal{G}_{mn} \varphi_n(1/\overline{z})^\dagger = z\varphi_m(1/\overline{z})^\dagger. \tag{3.21}$$

In order to have the feature of finite non-zero entries in both columns and rows, we define the following CMV basis:

$$\chi_{2k}(z) = z^{-k}\varphi_{2k}^{L,*}(z), \qquad \chi_{2k-1}(z) = z^{-k+1}\varphi_{2k-1}^{R}(z), \tag{3.22}$$

$$x_{2k}(z) = z^{-k}\varphi_{2k}^{R}(z), \qquad x_{2k-1}(z) = z^{-k}\varphi_{2k-1}^{L,*}(z). \tag{3.23}$$

Proposition 3.5. *(i) $\{\chi_\ell(z)\}_{\ell=0}^{\infty}$ and $\{x_\ell(z)\}_{\ell=0}^{\infty}$ are $\langle\langle\cdot,\cdot\rangle\rangle_R$ orthonormal, that is*

$$\langle\langle\chi_k, \chi_m\rangle\rangle_R = \langle\langle x_k, x_m\rangle\rangle_R = \delta_{km}. \tag{3.24}$$

(ii) χ_ℓ is in the module space of the first ℓ of $1, z^{-1}, z, \dots$ and x_ℓ of $1, z, z^{-1}, \dots$.

We definte the CMV matrix $\mathcal{C}_{nm} = \langle\langle\chi_n, z\chi_m\rangle\rangle_R$.

Proposition 3.6. *The matrix $\mathcal{C}$ is unitary in the following sense:*

$$\sum_{k=0}^{\infty} \mathcal{C}_{kn}^{\dagger}\mathcal{C}_{km} = \sum_{k=0}^{\infty} \mathcal{C}_{nk}\mathcal{C}_{mk}^{\dagger} = \delta_{nm}\mathbf{1}. \tag{3.25}$$

Proposition 3.7. *Let $|z| \leq 1$. Then, for every $m > 0$,*

$$\sum_{n=0}^{\infty} \chi_n(z)\mathcal{C}_{nm} = z\chi_m(z), \qquad \sum_{n=0}^{\infty} \mathcal{C}_{mn}\chi_n(1/\overline{z})^{\dagger} = z\chi_m(1/\overline{z})^{\dagger}. \tag{3.26}$$

An important feature of the matrix C is that it can represented as the product of two matrices:

$$\mathcal{C}_{nm} = \langle\langle\chi_n, z\chi_m\rangle\rangle_R = \sum_{k=0}^{\infty} \langle\langle\chi_n, zx_k\rangle\rangle_R \langle\langle x_k, \chi_m\rangle\rangle_R = \sum_{k=0}^{\infty} \mathcal{L}_{nk}\mathcal{M}_{km}. \tag{3.27}$$

Now we define the $2\ell \times 2\ell$ unitary matrix

$$\Theta_n = \Theta(\alpha_n) = \begin{pmatrix} \alpha_n^{\dagger} & \rho_n^L \\ \rho_n^R & -\alpha_n \end{pmatrix}. \tag{3.28}$$

Proposition 3.8. *We have the following relationships:*

$$\mathcal{L} = \Theta(\alpha_0) \oplus \Theta(\alpha_2) \oplus \Theta(\alpha_4) \oplus \cdots; \tag{3.29}$$

$$\mathcal{M} = \mathbf{1}_{\ell\times\ell} \oplus \Theta(\alpha_1) \oplus \Theta(\alpha_3) \oplus \cdots. \tag{3.30}$$

$$\tag{3.31}$$

And $\mathcal{C}$ can be written in matrix form:

$$\mathcal{C} = \begin{pmatrix} \alpha_0^\dagger & \rho_0^L \alpha_1^\dagger & \rho_0^L \rho_1^L & \mathbf{0} & \mathbf{0} & \cdots \\ \rho_0^R & -\alpha_0 \alpha_1^\dagger & -\alpha_0 \rho_1^L & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{0} & \alpha_2^\dagger \rho_1^R & -\alpha_2^\dagger \alpha_1 & \rho_2^L \alpha_3^\dagger & \rho_2^L \rho_3^L & \cdots \\ \mathbf{0} & \rho_2^R \rho_1^R & -\rho_2^R \alpha_1 & -\alpha_2 \alpha_3^\dagger & -\alpha_2 \rho_3^L & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \alpha_4^\dagger \rho_3^R & -\alpha_4^\dagger \alpha_3 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{3.32}$$

The proof is a straight forward computation and will be omitted here.

4 Sturm Oscillation Theory for Matrix Orthogonal Polynomial on the Unit Circle

4.1 Matrix Paraorthogonal Polynomial on the Unit Circle and Its Matrix Representation

Both GGT and CMV representation are infinite dimensional matrices. We want to study its eigenvalues so we need to find a valid truncation.

Definition 4.1. *Given any unitary $\beta \in \mathcal{M}_\ell$, we define the N-th paraorhogonal polynomial to be:*

$$\varphi_N^R(z;\beta) = z\varphi_{N-1}^R(z) - \varphi_{N-1}^{L,*}(z)\beta^\dagger. \tag{4.1}$$

With $\beta \in \mathcal{M}_\ell$ as the N-th Verblunsky coefficients both GGT and CMV matrices decouple. This is because for $\alpha_N = \beta$, $\rho_N^L = \rho_N^R = \mathbf{0}$. We will use $\mathcal{G}^N(\beta)$ to denote the truncated GGT matrix.

$$\mathcal{G}^N(\beta) = \begin{pmatrix} \alpha_0^\dagger & \rho_0^L\alpha_1^\dagger & \rho_0^L\rho_1^L\alpha_2^\dagger & \rho_0^L\rho_1^L\rho_2^L\alpha_3^\dagger & \cdots & \rho_0^L\rho_1^L\cdots\rho_{N-1}^L\beta^\dagger \\ \rho_0^R & -\alpha_0\alpha_1^\dagger & -\alpha_0\rho_1^L\alpha_2^\dagger & -\alpha_0\rho_1^L\rho_2^L\alpha_3^\dagger & \cdots & -\alpha_0\rho_1^L\cdots\rho_{N-1}^L\beta^\dagger \\ \mathbf{0} & \rho_1^R & -\alpha_1\alpha_2^\dagger & -\alpha_1\rho_2^L\alpha_3^\dagger & \cdots & -\alpha_1\rho_2^L\cdots\rho_{N-1}^L\beta^\dagger \\ \mathbf{0} & \mathbf{0} & \rho_2^R & -\alpha_2\alpha_3^\dagger & \cdots & -\alpha_2\rho_3^L\cdots\rho_{N-1}^L\beta^\dagger \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & -\alpha_{N-1}\beta^\dagger \end{pmatrix} \tag{4.2}$$

Note that the matrix $\mathcal{G}^N(\beta)$ has dimension $N+1$.

Proposition 4.2. $\mathcal{G}^N(\beta)$ *is unitary.*

For the CMV representation, we want to first study the $\mathcal{LM}$-decomposition, it is known that if we define

$$\Theta_n = \Theta(\alpha_n) = \begin{pmatrix} \alpha_n^\dagger & \rho_n^L \\ \rho_n^R & -\alpha_n \end{pmatrix} \tag{4.3}$$

where α_n are the Verblunsky coefficients then we have

$$\mathcal{L} = \begin{pmatrix} \Theta_0 & & & \\ & \Theta_2 & & \\ & & \Theta_4 & \\ & & & \ddots \end{pmatrix}$$
$$\mathcal{M} = \begin{pmatrix} \mathbf{1} & & & \\ & \Theta_1 & & \\ & & \Theta_3 & \\ & & & \ddots \end{pmatrix}$$

and

$$\mathcal{C} = \mathcal{L}\mathcal{M}. \tag{4.4}$$

Note $\|\alpha_n\| < 1$ for all n. If we replace α_N with a unitary matrix β then $\rho_N^L = \rho_N^R = \mathbf{0}$. As a result the CMV matrix $\mathcal{C}$ decouples into $\mathcal{C}^N(\beta)$ with dimension $n+1$ and a lower half matrix. Let us define the following:

$$\begin{aligned} \mathcal{L}^{2m} &= \Theta(\alpha_0) \oplus \Theta(\alpha_2) \oplus \cdots \oplus \Theta(\alpha_{2m}) \\ \mathcal{M}^{2n+1} &= \mathbf{1} \oplus \Theta(\alpha_1) \oplus \cdots \oplus \Theta(\alpha_{2n+1}). \end{aligned} \tag{4.5}$$

Then we have

$$\mathcal{C}^N(\beta) = \begin{cases} \mathcal{L}^{N-1}(\mathcal{M}^{N-2} \oplus \beta^\dagger), & \text{for } N \text{ odd} \\ (\mathcal{L}^{N-2} \oplus \beta^\dagger)(\mathcal{M}^{N-1}), & \text{for } N \text{ even.} \end{cases} \tag{4.6}$$

In matrix form we have

$$\mathcal{C}^N(\beta) = \begin{pmatrix} \alpha_0^\dagger & \rho_0^L\alpha_1^\dagger & \rho_0^L\rho_1^L & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \rho_0^R & -\alpha_0\alpha_1^\dagger & -\alpha_0\rho_1^L & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \alpha_2^\dagger\rho_1^R & -\alpha_2^\dagger\alpha_1 & \rho_2^L\alpha_3^\dagger & \rho_2^L\rho_3^L & \cdots & \mathbf{0} \\ \mathbf{0} & \rho_2^R\rho_1^R & -\rho_2^R\alpha_1 & -\alpha_2\alpha_3^\dagger & -\alpha_2\rho_3^L & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \alpha_4^\dagger\rho_3^R & -\alpha_4^\dagger\alpha_3 & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & -\alpha_{N-1}\beta^\dagger \end{pmatrix}, N \text{ odd}$$

$$\mathcal{C}^N(\beta) = \begin{pmatrix} \alpha_0^\dagger & \rho_0^L\alpha_1^\dagger & \rho_0^L\rho_1^L & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \rho_0^R & -\alpha_0\alpha_1^\dagger & -\alpha_0\rho_1^L & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \alpha_2^\dagger\rho_1^R & -\alpha_2^\dagger\alpha_1 & \rho_2^L\alpha_3^\dagger & \rho_2^L\rho_3^L & \cdots & \mathbf{0} \\ \mathbf{0} & \rho_2^R\rho_1^R & -\rho_2^R\alpha_1 & -\alpha_2\alpha_3^\dagger & -\alpha_2\rho_3^L & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \alpha_4^\dagger\rho_3^R & -\alpha_4^\dagger\alpha_3 & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & -\beta^\dagger\alpha_{N-1} \end{pmatrix}, N \text{ even}$$

Similar to $\mathcal{G}^N(\beta)$, $\mathcal{C}^N(\beta)$ has dimension $N+1$.

Proposition 4.3. $\mathcal{C}^N(\beta)$ *is unitary.*

4.2 Transition Matrices

One important feature of matrix orthogonal polynomials on the unit circle is the existence of transition matrices. They are essential in analyzing the eigenvalues for GGT and CMV matrices with matrix entries. Recall the Szegö recursion:

$$z\varphi_n^L - \rho_n^L\varphi_{n+1}^L = (\alpha_n^L)^\dagger\varphi_n^{R,*}, \tag{4.7}$$

$$z\varphi_n^R - \varphi_{n+1}^R\rho_n^R = \varphi_n^{L,*}(\alpha_n^R)^\dagger. \tag{4.8}$$

These imply a transition matrix for orthogonal polynomials:

$$\begin{pmatrix} \varphi_{n+1}^L \\ \varphi_{n+1}^{R,*} \end{pmatrix} = A^L(a_n, z)\begin{pmatrix} \varphi_n^L \\ \varphi_n^{R,*} \end{pmatrix} \tag{4.9}$$

where

$$A^L(\alpha, z)\begin{pmatrix} z(\rho^L)^{-1} & -(\rho^L)^{-1}\alpha^\dagger \\ -z(\rho^R)^{-1}\alpha & (\rho^R)^{-1} \end{pmatrix}. \tag{4.10}$$

At the same time, the Szegö recursion implies another set of relations:

$$\varphi_{n+1}^R = z\varphi_n^R(\rho_n^R)^{-1} - \varphi_n^{L,*}\alpha_n^\dagger(\rho_n^R)^{-1}, \tag{4.11}$$

$$\varphi_{n+1}^{L,*} = \varphi_n^{L,*}(\rho_n^L)^{-1} - z\varphi_n^R\alpha_n(\rho_n^L)^{-1}. \tag{4.12}$$

In other words,

$$\begin{pmatrix} \varphi_{n+1}^R & \varphi_{n+1}^{L,*} \end{pmatrix} = \begin{pmatrix} \varphi_n^R & \varphi_n^{L,*} \end{pmatrix} A^R(\alpha_n, z) \tag{4.13}$$

where

$$A^R(\alpha, z) = \begin{pmatrix} z(\rho^R)^{-1} & -z\alpha(\rho^L)^{-1} \\ -\alpha^\dagger(\rho^R)^{-1} & (\rho^L)^{-1} \end{pmatrix}. \tag{4.14}$$

For our convenience, we can reorganize the above to the following:

Proposition 4.4. *For $z \in \partial\mathbb{D}$*

$$\begin{pmatrix} \varphi_{n+1}^R \\ \varphi_{n+1}^{L,*} \end{pmatrix} = T_n^{\mathcal{G}}(\alpha_n, z)\begin{pmatrix} \varphi_{n+1}^R \\ \varphi_n^{L,*} \end{pmatrix} \tag{4.15}$$

where

$$T_n^{\mathcal{G}}(\alpha_n, z) = zA_n^{R,\dagger}. \tag{4.16}$$

Proof. First we take the conjugate transpose of equation (4.13) and we get

$$\begin{pmatrix} z^{-(n+1)}\varphi_{n+1}^R \\ \varphi_{n+1}^{L,\dagger} \end{pmatrix} = A_n^{R,\dagger}(\alpha_n, z) \begin{pmatrix} z^{-n}\varphi_{n+1}^R \\ \varphi_n^{L,\dagger} \end{pmatrix}. \tag{4.17}$$

Now multiply both sides by z^{n+1}

$$\begin{pmatrix} \varphi_{n+1}^R \\ \varphi_{n+1}^{L,*} \end{pmatrix} = zA_n^{R,\dagger}(\alpha_n, z) \begin{pmatrix} \varphi_{n+1}^R \\ \varphi_n^{L,*} \end{pmatrix}. \tag{4.18}$$

□

For the CMV representation, it is more convenient to study the transition properties of the CMV basis. We have that

$$\chi_{2k}(z) = z^{-k}\varphi_{2k}^{L,*}(z), \qquad \chi_{2k-1}(z) = z^{-k+1}\varphi_{2k-1}^R(z)$$
$$x_{2k}(z) = z^{-k}\varphi_{2k}^R(z), \qquad x_{2k-1}(z) = z^{-k}\varphi_{2k-1}^{L,*}(z).$$

By plugging these into Szegö recursion we find

$$\begin{aligned} zx_{2k} &= \chi_{2k}\alpha_{2k}^\dagger + \chi_{2k+1}\rho_{2k}^R \\ zx_{2k+1} &= \chi_{2k}\rho_{2k}^L - \chi_{2k+1}\alpha_{2k} \\ \chi_{2k-1} &= x_{2k-1}\alpha_{2k-1}^\dagger + x_{2k}\rho_{2k-1}^R \\ \chi_{2k} &= x_{2k-1}\rho_{2k-1}^L - x_{2k}\alpha_{2k-1}. \end{aligned}$$

Reorganizing these equations will give us two sets of transition functions:

$$\begin{pmatrix} \chi_{2k+1}^\dagger \\ x_{2k+1}^\dagger \end{pmatrix} = T_{2k}^{\mathcal{C}}(\alpha_{2k}, z) \begin{pmatrix} \chi_{2k}^\dagger \\ x_{2k}^\dagger \end{pmatrix} \tag{4.19}$$

$$\begin{pmatrix} \chi_{2k}^\dagger \\ x_{2k}^\dagger \end{pmatrix} = T_{2k-1}^{\mathcal{C}}(\alpha_{2k-1}, z) \begin{pmatrix} \chi_{2k-1}^\dagger \\ x_{2k-1}^\dagger \end{pmatrix}. \tag{4.20}$$

Proposition 4.5. *For* $z \in \partial\mathbb{D}$

$$T^{\mathcal{C}}_{2k}(\alpha_{2k}, z) = z \begin{pmatrix} 0 & 1 \\ z & 0 \end{pmatrix} A^L(\alpha_{2k}, z) \tag{4.21}$$

$$T^{\mathcal{C}}_{2k-1}(\alpha_{2k-1}, z) = A^L(\alpha_{2k-1}, z) \begin{pmatrix} 0 & z^{-1} \\ 1 & 0 \end{pmatrix} \tag{4.22}$$

Proof. If $z \in \partial\mathbb{D}$, from the definitions of χ's and x's we have that

$$\chi^\dagger_{2k}(z) = z^{-k}\varphi^L_{2k}(z) \qquad \chi^\dagger_{2k-1}(z) = z^{-k}\varphi^{R,*}_{2k-1}(z) \tag{4.23}$$

$$x^\dagger_{2k}(z) = z^{-k}\varphi^{R,*}_{2k}(z) \qquad x^\dagger_{2k-1}(z) = z^{-k+1}\varphi^L_{2k-1}(z). \tag{4.24}$$

This gives

$$\begin{pmatrix} \chi^\dagger_{2k} \\ x^\dagger_{2k} \end{pmatrix} = z^{-k} \begin{pmatrix} \varphi^L_{2k} \\ \varphi^{R,*}_{2k} \end{pmatrix}; \tag{4.25}$$

$$\begin{pmatrix} \chi^\dagger_{2k-1} \\ x^\dagger_{2k-1} \end{pmatrix} = z^{-k} \begin{pmatrix} 0 & 1 \\ z & 0 \end{pmatrix} \begin{pmatrix} \varphi^L_{2k-1} \\ \varphi^{R,*}_{2k-1} \end{pmatrix}. \tag{4.26}$$

Using the transition function for φ's we get

$$\begin{pmatrix} \chi^\dagger_{2k} \\ x^\dagger_{2k} \end{pmatrix} = A^L(\alpha_{2k-1}, z) \begin{pmatrix} 0 & 1 \\ z & 0 \end{pmatrix}^{-1} \begin{pmatrix} \chi^\dagger_{2k-1} \\ x^\dagger_{2k-1} \end{pmatrix} \tag{4.27}$$

and

$$\begin{pmatrix} \chi^\dagger_{2k+1} \\ x^\dagger_{2k+1} \end{pmatrix} = z \begin{pmatrix} 0 & 1 \\ z & 0 \end{pmatrix} A^L(\alpha_{2k}, z) \begin{pmatrix} \chi^\dagger_{2k} \\ x^\dagger_{2k} \end{pmatrix}. \tag{4.28}$$

This completes the proof. □

4.3 Intersection of Lagrangian Planes and Sturm Oscillation

We first consider the matrix

$$\mathcal{J} = \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{pmatrix} \tag{4.29}$$

where $\mathbf{1}$ is the $\ell \times \ell$ identity matrix. The $\mathcal{J}$-Lagrangian planes in $\mathbb{C}^{2L}$ is defined to be ℓ-dimensional planes on which the form $\mathcal{J}$ vanishes.

Definition 4.6. *Let $\Phi = (\phi(1), \dots, \phi(L))$ with the vectors $\phi(1), \dots, \phi(L) \in \mathcal{C}^{2\ell}$. Then we call Φ a $\mathcal{J}$-Lagrangian frame if*

$$\Phi^\dagger \mathcal{J} \Phi = 0. \tag{4.30}$$

Definition 4.7. *A $\mathcal{J}$-Lagrangian plane is the equivalence class $[\Phi]_\sim$ with the relation $\sim$: $\Phi \sim \Phi^\dagger$ if $\Phi = \Phi^\dagger c$ for some $c \in Gl(\ell, \mathbb{C})$. The $\mathcal{J}$-Lagrangian Grassmannian $\mathbb{L}_\ell$ is the set of all equivalence classes of $\mathcal{J}$-Lagrangian planes.*

Proposition 4.8. *The set $\mathbb{L}_\ell^* = \{\Phi = (a, b)^T | a, b \text{ are invertible}\} \subset \mathbb{L}_L$ is bijective to $U(L)$ via the map*

$$\Pi([\Phi]_\sim) = ab^{-1}. \tag{4.31}$$

Proof. Let $\Phi = \begin{pmatrix} a \\ b \end{pmatrix}$ be $\mathcal{J}$-Lagrangian. We have that $a^\dagger a = b^\dagger b$.Therefore $\Pi([\Phi]_\sim)^\dagger \Pi([\Phi]_\sim) = (b^\dagger)^{-1} a^\dagger a b^{-1} = \mathbf{1}$. So $\Pi([\Phi]_\sim)$ is unitary. Let $U = ab^{-1}$ we have $\Phi \begin{pmatrix} U \\ \mathbf{1} \end{pmatrix} b$. This establishes the bijection. □

Proposition 4.9. *Let Φ and Ψ be $\mathcal{J}$-Lagrangian planes, and Let $U = \Pi([\Phi]_\sim)$ and $V = \Pi([\Psi]_\sim)$. Then*

$$\dim(\Phi\mathbb{C}^\ell \cap \Psi\mathbb{C}^\ell) = \dim(\ker(\Phi^\dagger \mathcal{J} \Psi)) = \dim(\ker(V^\dagger U - \mathbf{1}_\ell)). \tag{4.32}$$

Proof. Let us begin with the inequality $\leq$ of the first equality. Suppose there are two $\ell \times p$ matrices v, w of rank p such that $\Phi v = \Psi w$. Then $\Phi^\dagger \mathcal{J} \Psi w = \Phi^\dagger \mathcal{J} v = 0$

so that the kernel of $\Phi^\dagger \mathcal{J}\Psi$ is at least of dimension p. Inversely, given a $\ell \times p$ matrix w of rank p such that $\Phi^\dagger \mathcal{J}\Psi w = 0$, we deduce that $(\mathcal{J}\Phi)^\dagger \Psi w = 0$. As the column vectors of Φ and $\mathcal{J}\Psi$ are orthogonal and span $\mathbb{C}^{2\ell}$, it follows that the column vectors of Ψw lie in the span of Φ, that is, there exists an $\ell \times p$ matrix v of rank p such that $\Psi w = \Phi v$. This shows that the other inequality and hence proves the first equality of the proposition. For the second, we first note that the dimension of the kernel of $\Phi^\dagger \mathcal{J}\Psi$ does not depend on the choice of the representative. We use the representative of $\Phi = \begin{pmatrix} \pi([\Phi]_\sim) \\ \mathbf{1} \end{pmatrix}$ and $\Psi = \begin{pmatrix} \pi([\Psi]_\sim) \\ \mathbf{1} \end{pmatrix}$. A short calculation shows that $\Phi^\dagger \mathcal{J}\Psi = 2iU^\dagger(U - V)$ which implies the second equality.

□

Notice that $\dim(\ker(V^\dagger U - \mathbf{1}_\ell))$ also can be rephrased as the multiplicity of 1 as eigenvalue of $V^\dagger U$.

Proposition 4.10. *$\mathcal{J}$-Lagrangian planes are mapped to the $\mathcal{J}$-Lagrangian planes by matrices in the $\mathcal{J}$-invariance group defined by*

$$I(2\ell, \mathbb{C}) = \{\mathcal{T} \in Mat(2\ell \times 2\ell, \mathbb{C}) | \mathcal{T}^\dagger \mathcal{J}\mathcal{T} = \pm\mathcal{J}\}. \tag{4.33}$$

Proof. This can be seen as:

$$\begin{aligned} & (\mathcal{T}\Phi)^\dagger \mathcal{J}(\mathcal{T}\Phi) \\ = \; & \Phi^\dagger \mathcal{T}^\dagger \mathcal{J}\mathcal{T}\Phi \\ = \; & \Phi^\dagger(\pm\mathcal{J})\Phi \\ = \; & 0 \end{aligned}$$

□

Theorem 4.11. *If $z \in \partial\mathbb{D}$ we have that both $T^{\mathcal{G}}$ are $T^{\mathcal{C}}$ are $\mathcal{J}$-invariant meaning:*

$$T^\dagger \mathcal{J} T = \pm\mathcal{J} \tag{4.34}$$

Proof. First, let's prove that if $A, B \in \mathrm{I}(2\ell, \mathbb{C})$ then $AB \in \mathrm{I}(2\ell, \mathbb{C})$. This can be seen as

$$\begin{aligned} & (AB)\dagger\mathcal{J}(AB) \\ = \ & B^\dagger(A^\dagger\mathcal{J}A)B \\ = \ & B^\dagger(\pm\mathcal{J})B \\ = \ & \pm\mathcal{J} \end{aligned}$$

A simple calculation shows that

$$\begin{pmatrix} 0 & 1 \\ z & 0 \end{pmatrix}^\dagger \mathcal{J} \begin{pmatrix} 0 & 1 \\ z & 0 \end{pmatrix} = -\mathcal{J} \tag{4.35}$$

$$\begin{pmatrix} 0 & z^{-1} \\ 1 & 0 \end{pmatrix}^\dagger \mathcal{J} \begin{pmatrix} 0 & z^{-1} \\ 1 & 0 \end{pmatrix} = -\mathcal{J} \tag{4.36}$$

At the same time

$$\begin{aligned} & A^L(\alpha,z)^\dagger\mathcal{J}A^L(\alpha,z) \\ = \ & \begin{pmatrix} \bar{z}(\rho^L)^{-1} & -\bar{z}(\rho^R)^{-1} \\ -\alpha(\rho^L)^{-1} & (\rho^R)^{-1} \end{pmatrix} \begin{pmatrix} z(\rho^L)^{-1} & -(\rho^L)^{-1}\alpha^\dagger \\ z(\rho^R)^{-1}\alpha & -(\rho^R)^{-1} \end{pmatrix} \\ = \ & \begin{pmatrix} \bar{z}z(\rho^L)^{-2} - \bar{z}z(\rho^R)^{-2}\alpha & -\bar{z}(\rho^L)^{-2}\alpha^\dagger + \bar{z}\alpha^\dagger(\rho^R)^{-2} \\ -z\alpha(\rho^L)^{-2} + z(\rho^R)^{-2}\alpha & \alpha(\rho^L)^{-2}\alpha^\dagger - (\rho^R)^{-2} \end{pmatrix} \\ = \ & \begin{pmatrix} \bar{z}z\mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{pmatrix} \\ = \ & \mathcal{J} \end{aligned}$$

Similarly,

$$A^R(\alpha,z)^\dagger\mathcal{J}A^R(\alpha,z) = \mathcal{J}.$$

This completes the proof. □

Proposition 4.12. *Let $z \in \partial\mathbb{D}$. With* $\begin{pmatrix} \varphi_0^R \\ \varphi_0^{L,*} \end{pmatrix} = \begin{pmatrix} \mathbf{1} \\ \mathbf{1} \end{pmatrix}$ *and* $\begin{pmatrix} \chi_0^\dagger \\ x_0^\dagger \end{pmatrix} = \begin{pmatrix} \mathbf{1} \\ \mathbf{1} \end{pmatrix}$, *we have that* $\begin{pmatrix} \varphi_n^R \\ \varphi_n^{L,*} \end{pmatrix}$ *and* $\begin{pmatrix} \chi_n^\dagger \\ x_n^\dagger \end{pmatrix}$ *are* $\mathcal{J}$*-Lagrangian.*

Proof. Clearly $\begin{pmatrix} \varphi_0^R \\ \varphi_0^{L,*} \end{pmatrix}$ and $\begin{pmatrix} \chi_0^\dagger \\ x_0^\dagger \end{pmatrix}$ are $\mathcal{J}$-Lagrangian. At the same time the transition matrices are $\mathcal{J}$-invriant. So Lagrangian planes will be mapped to Lagrangian planes by these transition matrices. Thus $\begin{pmatrix} \varphi_n^R \\ \varphi_n^{L,*} \end{pmatrix}$ and $\begin{pmatrix} \chi_n^\dagger \\ x_n^\dagger \end{pmatrix}$ are $\mathcal{J}$-Lagrangian. □

We are interested in study the truncated versions of both matrix representations. For the GGT representation, let's consider the N-th vector $\begin{pmatrix} \varphi_N^R \\ \varphi_N^{L,*} \end{pmatrix}$. It can calculated from applying the first $N-1$ transition matrix to $\begin{pmatrix} \varphi_0^R \\ \varphi_0^{L,*} \end{pmatrix}$. But we are imposing an extra condition at the end, which is $\alpha_N = \beta$. From the recursion formula, this requires the vector $\begin{pmatrix} \varphi_N^R \\ \varphi_N^{L,*} \end{pmatrix}$ satisfy

$$z\varphi_N^R = \varphi_N^{L,*}\beta^\dagger. \tag{4.37}$$

It is important to note that this is not necessarily satisfied by the N-th vector. We will show that the degree which the N-th vector satisfies this condition corresponds to the multiplicity of z as an eigenvalue of $\mathcal{G}^N$. We introduce the vector

$$\mathcal{B}^{\mathcal{G}} = \begin{pmatrix} \beta^\dagger \\ z \end{pmatrix}, \tag{4.38}$$

as the boundary condition.

Similar to the GGT representation, the recursion formula for CMV basis gives the following boundary conditions based on the parity of N:

$$\overline{z}x_N^\dagger = -\beta^\dagger\chi_N^\dagger \qquad \text{for } N \text{ even} \tag{4.39}$$

$$\chi_N^\dagger = -\beta^\dagger x_N^\dagger \qquad \text{for } N \text{ odd.} \tag{4.40}$$

These are extra constraints on the final vector $\begin{pmatrix} \chi_N^\dagger \\ x_N^\dagger \end{pmatrix}$. We introduce the following vectors:

$$\mathcal{B}^{\mathcal{C}} = \begin{pmatrix} \overline{z}\mathbf{1} \\ -\beta^{\dagger} \end{pmatrix} \qquad \text{when } N \text{ is even} \tag{4.41}$$

$$\mathcal{B}^{\mathcal{C}} = \begin{pmatrix} \mathbf{1} \\ -\beta^{\dagger} \end{pmatrix} \qquad \text{when } N \text{ is odd.} \tag{4.42}$$

Again, these vectors are another way to describe the boundary conditions to suit our purpose.

Theorem 4.13.

$$\begin{aligned} & \textit{The geometric multiplicity of } z \textit{ as an eigenvalue of } \mathcal{G}^N \\ = \ & \dim\left(\begin{pmatrix} \varphi_N^R \\ \varphi_N^{L,*} \end{pmatrix} \mathbb{C}^{\ell} \cap \mathcal{B}^{\mathcal{G}}\mathbb{C}^{\ell}\right). \end{aligned} \tag{4.43}$$

$$\begin{aligned} & \textit{The geometric multiplicity of } z \textit{ as an eigenvalue of } \mathcal{C}^N \\ = \ & \dim\left(\begin{pmatrix} \chi_N^{\dagger} \\ x_N^{\dagger} \end{pmatrix} \mathbb{C}^{\ell} \cap \mathcal{B}^{\mathcal{C}}\mathbb{C}^{\ell}\right). \end{aligned} \tag{4.44}$$

Remark 4.14. *Given the complexity of the CMV and GGT representation with matrix entries, the eigenvalues are often hard to find. The above theorem relates information about eigenvalues to the intersection of two Lagrangian planes. Thus we can use the tools in symplectic geometry to find an algorithm to determine eigenvalues.*

Proof. The proofs for GGT and CMV are identical, we will present the detail for the more difficult CMV case. Recall that we have

$$\sum_{n=0}^{\infty} \mathcal{C}_{mn}\chi_n(1/\overline{z})\dagger = z\chi_m(1/\overline{z})^{\dagger}. \tag{4.45}$$

Notice that since $\mathcal{C}^N$ is unitary, all the eigenvalues are on the unit circle. So we have $z = 1/\overline{z}$. For $\mathcal{C}^N$, let's introduce the vector $\Phi_N(z) = (\chi_0^{\dagger}(z), \chi_1^{\dagger}(z), \ldots, \chi_N^{\dagger}(z))^T$.

Notice if we can find a vector $\vec{v} \in \mathbb{C}^\ell$ such that

$$\mathcal{C}^N \Phi_N(z)\vec{v} = z\Phi_N(z)\vec{v} \tag{4.46}$$

then $\Phi_N(z)\vec{v}$ is an eigenvector for $\mathcal{C}^N$ with eigenvalue z. So if we choose $\vec{v} \in \begin{pmatrix} \chi_N^\dagger \\ x_N^\dagger \end{pmatrix} \mathbb{C}^\ell \cap \mathcal{B}^{\mathcal{C}}\mathbb{C}^\ell$ then $\vec{v}$ satisfy

$$\begin{aligned} \overline{z} x_N^\dagger \vec{v} &= -\beta^\dagger \chi_N^\dagger \vec{v} \qquad \text{for } N \text{ even} \\ \chi_N^\dagger \vec{v} &= -\beta^\dagger x_N^\dagger \vec{v} \qquad \text{for } N \text{ odd.} \end{aligned}$$

It satisfies the final boundary condition and thus we have $\mathcal{C}^N \Phi_N(z)\vec{v} = z\Phi_N(z)\vec{v}$. This shows that $\Phi_N(z)\vec{v}$ is an eigenvector.

Conversely, if $\vec{t} = (t_1, t_2, \ldots, t_{\ell N+\ell})^T \in \mathbb{C}^{\ell N+\ell}$ is an eigenvector with eigenvalue z. We rewrite $\vec{t}$ as

$$\vec{T} = (T_0, T_1 \ldots, T_N), \text{ where } T_n = (t_{nL_1}, \ldots, t_{(n+1)\ell}). \tag{4.47}$$

I claim that we have

$$\mathcal{C}^N \Phi_N(z) T_0 = z\Phi_N(z) T_0 \tag{4.48}$$

with $\Phi_N(z)T_0 = \vec{T}$. This shows T_0 must be in the intersection of $\begin{pmatrix} \chi_N^\dagger \\ x_N^\dagger \end{pmatrix} \mathbb{C}^\ell$ and $\mathcal{B}\mathbb{C}^\ell$. To prove the claim $\Phi_N(z)T_0 = \vec{T}$ we show the equality component-wise by inducting on T_n. The base case T_0 is trivial. Then we consider the equation $\mathcal{C}^N \Phi_N(z)T_0 = z\Phi_N(z)T_0$. The first two lines give

$$\alpha_0^\dagger \chi_0^\dagger(z) T_0 + \rho_0^L \alpha_1^\dagger \chi_1^\dagger(z) T_0 + \rho_0^L \rho_1^L \chi_2^\dagger(z) T_0 = z\chi_0^\dagger(z) T_0 \tag{4.49}$$

$$\rho_0^R \chi_0^\dagger(z) T_0 - \alpha_0 \alpha_1^\dagger \chi_1^\dagger(z) T_0 - \alpha_0 \rho_1^L \chi_2^\dagger(z) T_0 = z\chi_1^\dagger(z) T_0 \tag{4.50}$$

Let's compare these with the first two lines of the eigenvalue equation $\mathcal{C}^N \vec{T} = z\vec{T}$:

$$\alpha_0^\dagger T_0 + \rho_0^L \alpha_1^\dagger T_1 + \rho_0^L \rho_1^L T_2 = zT_0 \tag{4.51}$$

$$\rho_0^R T_0 - \alpha_0 \alpha_1^\dagger T_1 - \alpha_0 \rho_1^L T_2 = zT_1. \tag{4.52}$$

We can consider $\chi_1^\dagger(z)T_0$ and $\chi_2^\dagger(z)$ as 2ℓ unknowns. At the same time we have 2ℓ linear equations. So we have the unique solution as

$$\chi_1^\dagger(z)T_0 = T_1 \qquad \chi_2^\dagger(z)T_0 = T_2 \tag{4.53}$$

More generally,

$$[\mathcal{C}^N \Phi_N(z)T_0]_{2n-1,2n} = [z\Phi_N(z)T_0]_{2n-1,2n} \tag{4.54}$$

and

$$[\mathcal{C}^N \vec{T}]_{2n-1,2n} = [z\Phi_N(z)\vec{T}]_{2n-1,2n} \tag{4.55}$$

gives the relation $\chi_{2n-1}^\dagger(z)T_0 = T_{2n-1}, \chi_{2n}^\dagger(z)T_0 = T_{2n}$ if $\chi_{2n-3}^\dagger(z)T_0 = T_{2n-3}, \chi_{2n-2}^\dagger(z)T_0 = T_{2n-2}$. The last such equation is exactly the boundary condition. This completes the proof. □

Theorem 4.15. *(a) The geometric multiplicity of z as an eigenvalue of $\mathcal{C}^N$ is equal to the geometric multiplicity of 1 as eigenvalue of $M_N^{\mathcal{G}}(z)$, where*

$$M_N^{\mathcal{G}}(z) = z\beta^\dagger \varphi_N^R(z)\varphi_N^{L,*}(z)^{-1} \tag{4.56}$$

(b) The geometric multiplicity of z as an eigenvalue of $\mathcal{C}^N$ is equal to the geometric multiplicity of 1 as eigenvalue of $M_N^{\mathcal{C}}(z)$, where

$$M_N^{\mathcal{C}}(z) = -z\beta^{-1}\chi_N^\dagger(z)x_N^\dagger(z)^{-1} \qquad \text{for } N \text{ even} \tag{4.57}$$

$$M_N^{\mathcal{C}}(z) = -\beta^{-1}\chi_N^\dagger(z)x_N^\dagger(z)^{-1} \qquad \text{for } N \text{ odd} \tag{4.58}$$

Proof. We can directly use Proposition 4.13 and Theorem 4.9 to conclude this theorem. □

With this theorem, we can gain spectral information of the GGT and CMV matrices with matrix entires, which have $N\ell \times N\ell$ dimension by analyzing a much simpler matrix that has dimension $\ell \times \ell$. This provides an algorithm to check whether a number on the unit circle is an eigenvalue of the GGT and CMV matrix.

5 References

[1] M.J. Cantero, L. Moral, and L. Velázquez, *Measure and para-orthogonal polynomials on the unit circle*, Linear Algebra Appl. **362** (2003), 29-56.

[2] P. Delsarte, Y.V. Genin, and Y.G. Kamp, *Orthogonal polynomial matrices on the unit circle*, IEEE Trans, Circuits and Systems **CAS-25** (1978), 149-160.

[3] D.C. Youla and N.N. Kazanjian, *Bauer-type factorization of positive matrices and the theory of matrix polynomials orthogonal on the unit circle*, IEEE Trans. Circuits and Systems **CAS-25** (1978), 57-69.

[4] H. Helson and D. Lowdenslager, *Prediction theory and Fourier series in several variables*, Acta Math. **99** (1958), 165-202.

[5] H. Helson and D. Lowdenslager, *Prediction theory and Fourier series in several variables, II*, Acta Math. **106** (1961), 175-213.

[6] M. Stoiciu, *Zeros of Random Orthogonal Polynomials on the Unit Circle*, Ph.D. Thesis at Caltech (2005)

[7] R.Bott, *On the Iteration of Closed Geodesics and the Sturm Intersection Theory*, Commun. Pure Appl. Math. **9**, 171-206 (1956)

[8] D. Damanik, A. Pushnitski, B. Simon, The Analytic Theory of Matrix Orthogonal Polynomials, Surveys in Approsimation Theory, **4** (2008), 1-85

[9] B. Simon, *CMV Matrices: Five Years After*, Journal of Computational and Applied Mathematics, 208(1). pp. 120-154.

[10] Hermann Schulz-Baldes, *Sturm intersection theory for periodic Jacobi matrices and linear Hamiltonian systems*, unpublished.

[11] B. Simon, *Orthogonal Polynomials on the Unit Circle, Part 1: Classical Theory*, AMS Colloquium Series, American Mathematical Society, Providence, RI, 2005.

[12] B. Simon, *Orthogonal Polynomials on the Unit Circle, Part 2: Spectral Theory*, AMS Colloquium Series, American Mathematical Society, Providence, RI, 2005.

[13] G. Szegö. *Orthogonal Polynomials*, 4th ed. Providence, RI: Amer. Math. Soc., pp. 44-47 and 54-55, 1975.

[14] G. Teschl, *Jacobi Operators and Completely Integrable Nonlinear Lattices*, Mathematical Surveys and Monographs 72, Amer. Math. Soc., Providence, RI, 2000.

[15] G. Teschl, *Topics in Real and Functional Analysis*, unpublished, available online at `http://www.mat.univie.ac.at/~gerald/`.