Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Linear Algebra and its Applications 408 (2005) 40–65
LINEAR ALGEBRA AND ITS APPLICATIONS
www.elsevier.com/locate/laa

Minimal representations of unitary operators and orthogonal polynomials on the unit circle ☆

M.J. Cantero, L. Moral *, L. Velázquez

Departamento de Matemática Aplicada, Universidad de Zaragoza, Pza. San Francisco s/n, 50009 Zaragoza, Spain

Received 9 April 2004; accepted 21 April 2005
Available online 22 August 2005
Submitted by R.A. Brualdi

Abstract

In this paper we prove that the simplest band representations of unitary operators on a Hilbert space are five-diagonal. Orthogonal polynomials on the unit circle play an essential role in the development of this result, and also provide a parameterization of such five-diagonal representations which shows specially simple and interesting decomposition and factorization properties. As an application we get the reduction of the spectral problem of any unitary Hessenberg matrix to the spectral problem of a five-diagonal one. Two applications of these results to the study of orthogonal polynomials on the unit circle are presented: the first one concerns Krein's Theorem; the second one deals with the movement of mass points of the orthogonality measure under mono-parametric perturbations of the Schur parameters.

AMS classification: 42C05; 47B36

Keywords: Orthogonal polynomials on the unit circle; Unitary band matrices; Isometric Hessenberg matrices

☆ This work was supported by Project E-12/25 of DGA (Diputación General de Aragón) and by Ibercaja under grant IBE2002-CIEN-07.
* Corresponding author. Tel.: +34 976 76 11 41; fax: +34 976 76 11 25.
E-mail address: lmoral@unizar.es (L. Moral).

0024-3795/$ - see front matter © 2005 Published by Elsevier Inc.
doi:10.1016/j.laa.2005.04.025

1. Introduction

Matrix representations are an important tool for the study of linear operators on a Hilbert space. They allow, for instance, the use of perturbation techniques for the comparison of operators defined on different Hilbert spaces. Besides, the freedom in the choice of the representation can be used to get a simple one that can make the analysis of the operator easier. Usually a band representation with minimum size band is desirable. A band matrix $(c_{i,j})$ is (p,q)-diagonal if $c_{i,j} = 0$ for $i - j > p$ and $j - i > q$. A matrix that is $([\frac{n}{2}], [\frac{n-1}{2}])$-diagonal or $([\frac{n-1}{2}], [\frac{n}{2}])$-diagonal, is called a n-diagonal matrix. If every operator of a certain class has a n-diagonal representation but not all of them have a $n-1$-diagonal one, we say that n-diagonal representations are the minimal representations of the class.

Concerning the class of self-adjoint operators, any two-diagonal representation must be diagonal due to its symmetry, but a diagonal representation is only possible in the case of pure point spectrum. Therefore, the minimal representations are at least tri-diagonal. In fact, they are tri-diagonal since, as a consequence of the spectral theorem, every self-adjoint operator is unitarily equivalent to an orthogonal sum of self-adjoint multiplication operators [28] and, hence, the use of basis of orthogonal polynomials on the real line gives a tri-diagonal representation [32].

Unitary operators, together with self-adjoint ones, are the most important examples of normal operators. However, in spite of their importance, the minimal representations for unitary operators are an open problem. Analogously to the self-adjoint case, the study can be reduced to unitary multiplication operators, but the use of basis of orthogonal polynomials on the unit circle then leads to Hessenberg instead of band representations [9,2,17,27,34]. As for the possibility of band representations, it has been recently proved in [3] that any unitary tri-diagonal matrix decomposes as a sum of 1×1 and 2×2 diagonal blocks and, therefore, it has a pure point spectrum. This shows that the minimal representations of unitary operators are at least four-diagonal. Simon has conjectured in a preliminary version of [31] that a similar decomposition should happen for any unitary four-diagonal matrix, which would imply that the minimal representations for unitary operators are at least five-diagonal.

A step to get the minimal representations of unitary operators was taken by the authors in [6]. The results presented there imply that any unitary operator has a five-diagonal representation. In the next section we introduce these five-diagonal representations and their connection with orthogonal polynomials on the unit circle. Section 3 is devoted to the study of such representations and their properties. Although Hessenberg matrices have been extensively studied, all this analysis will be done jointly for five-diagonal and Hessenberg representations for several reasons:

- It is convenient to understand the improvements given by the five-diagonal representations, if compared with the known Hessenberg ones. Some concrete examples of the advantages of the five-diagonal representations will be clearly shown in the applications discussed in Sections 4 and 5.

- The connections between Hessenberg and five-diagonal representations provide an algorithm that reduces the spectral problem of any unitary Hessenberg matrix to the spectral problem of a five-diagonal one ("five-diagonal reduction" of the spectral problem of a unitary Hessenberg matrix). The importance of this result is due to the increasing interest in the study of unitary Hessenberg matrices in numerical linear algebra [17,18,19] and digital signal processing applications (see [7] and references therein).
- The analysis of unitary Hessenberg matrices is the main tool to prove that the minimal representations of unitary operators are indeed five-diagonal. This result is a consequence of a more general one that closes Section 3: $(1, q)$-diagonal or $(p, 1)$-diagonal representations of unitary operators are possible only in the case of pure point spectrum.

In Sections 4 and 5 we consider some applications of the minimal representations of unitary operators to the study of orthogonal polynomials on the unit circle. Both applications concern the relation between the support of the measure of orthogonality and the corresponding Schur parameters. Section 4 shows the advantages of the five-diagonal representation for the analysis of the limit points of the support of the measure, while Section 5 is devoted to the study of the isolated mass points. We finish this last section giving several explicit examples of perturbations of the Schur parameters that keep an arbitrary mass point invariant.

Now we proceed with the conventions for the notation. For any subset $\mathscr{A}$ of a Hilbert space, $\overline{\mathscr{A}}$ is its closure and span $\mathscr{A}$ the set of all finite linear combinations of $\mathscr{A}$. Also, if $\mathscr{S}$ is a subspace of the Hilbert space, $\mathscr{A}^{\perp \mathscr{S}}$ means the subspace of $\mathscr{S}$ orthogonal to $\mathscr{A}$.

Given a linear operator T on a Hilbert space, T^* denotes its adjoint and $\sigma(T)$ its spectrum, while for every complex matrix M, M^{T} is its transpose and $M^* = \overline{M}^{\mathrm{T}}$. I and I_N represent the unit matrix of order infinite and N, respectively. Any matrix of order N is considered as an operator in $\mathbb{C}^N$, and any infinite bounded matrix is identified with the continuous operator that it defines in ℓ^2, the Hilbert space of square-sumable sequences in $\mathbb{C}$. The inner products in $\mathbb{C}^N$ and ℓ^2 are denoted by $(\cdot, \cdot)$, and the corresponding canonical basis by $\{e_n\}$. No misunderstanding will arise from this common notation.

The term measure always means non-negative finite Borel measure, and, without loss of generality, we will consider only probability measures. If μ is a measure on a subset of $\mathbb{C}$, suppμ is its support and L^2_μ the Hilbert space of μ-square-integrable complex functions with inner product

$$\langle f, g\rangle_\mu := \int f(z)\overline{g(z)}\, \mathrm{d}\mu(z), \quad f, g \in L^2_\mu.$$

$\mathbb{T} := \{z \in \mathbb{C} : |z| = 1\}$ is the unit circle and $\mathbb{D} := \{z \in \mathbb{C} : |z| < 1\}$ the open unit disk in the complex plane. A multiplication operator on $\mathbb{T}$ has the form

$$\begin{aligned} U_\mu :\ & L^2_\mu \to L^2_\mu \\ & f(z) \to zf(z) \end{aligned}$$

where μ is a measure on $\mathbb{T}$.

2. Representations of unitary operators and orthogonal polynomials on $\mathbb{T}$

Given a unitary operator U on a separable Hilbert space $\mathscr{H}$, the equivalence between the following assertions is known [32]:

- The spectrum of U is simple.
- U has a cyclic vector $v \in \mathscr{H}$, in the sense of $\overline{\operatorname{span}\{U^n v\}}_{n\in\mathbb{Z}} = \mathscr{H}$.
- U is unitarily equivalent to a multiplication operator on $\mathbb{T}$.

A standard application of Zorn's lemma shows that any unitary operator can be expressed as a (finite or infinite) orthogonal sum of unitary operators with cyclic vectors. Therefore, the study of unitary operators becomes the study of multiplication operators on $\mathbb{T}$. As for the spectral properties of such multiplication operators, it is known that $\sigma(U_\mu) = \operatorname{supp}\mu$, the mass points of μ being the eigenvalues of U_μ. The eigenvectors of U_μ associated with an eigenvalue λ are spanned by the characteristic function $\mathscr{X}_\lambda$ of the set $\{\lambda\}$.

For a long time, the usual attempts to get matrix representations of U_μ have dealt with basis constituted by orthogonal polynomials (OP) with respect to μ, that is, polynomials satisfying

$$\deg \varphi_n = n, \quad \langle \varphi_n, \varphi_m \rangle_\mu = \delta_{n,m}, \qquad n, m \geqslant 0. \tag{1}$$

When $\operatorname{supp}\mu$ has a finite number N of elements, $\dim(L^2_\mu) = N$ and such a basis $\mathbf{\Phi}_N := (\varphi_n)_{n=0}^{N-1}$ comes from the orthogonalization of $\{z^n\}_{n=0}^{N-1}$. $\mathbf{\Phi}_N$ is called a finite segment of OP associated with μ. If $\operatorname{supp}\mu$ is infinite, $\dim(L^2_\mu) = \aleph_0$ and the orthogonalization of the infinite set $\{z^n\}_{n\geqslant 0}$ gives a sequence $\mathbf{\Phi} := (\varphi_n)_{n\geqslant 0}$ satisfying (1) that is called a sequence of OP with respect to μ. However, such a sequence is not always a basis of L^2_μ since the polynomials are not always dense in L^2_μ.

In what follows, φ_n denotes the unique n-th OP with respect to μ with positive leading coefficient κ_n. It is known that these polynomials satisfy the recurrence relation

$$\begin{aligned} &\varphi_0(z) = 1, \\ &\rho_n \varphi_n(z) = z\varphi_{n-1}(z) + a_n \varphi^*_{n-1}(z), \qquad n \geqslant 1, \end{aligned} \tag{2}$$

where $p^*(z) := z^n \overline{p}(z^{-1})$ for a polynomial p of degree n, $\rho_n := \sqrt{1 - |a_n|^2}$ and $a_n \in \mathbb{D}$ are known as the Schur parameters associated with μ.

Besides, when supp $\mu = \{z_1, z_2, \dots, z_N\}$, the same arguments that give (2) show that the polynomial $\psi(z) = (z - z_1)(z - z_2) \cdots (z - z_N)$ satisfies

$$\kappa_{N-1}\psi(z) = z\varphi_{N-1}(z) + a_N\varphi^*_{N-1}(z), \quad a_N \in \mathbb{T}. \tag{3}$$

It is known that the preceding results establish a one to one correspondence between:

- Probability measures supported on N points of the unit circle and vectors $\mathbf{a}_N := (a_1, a_2, \dots, a_N) \in \mathbb{D}^{N-1} \times \mathbb{T}$.
- Probability measures supported on an infinite subset of the unit circle and sequences $\mathbf{a} := (a_n)_{n\in\mathbb{N}} \in \mathbb{D}^{\aleph_0}$.

If $\mathbf{\Phi}$ is the sequence of OP related to a measure μ with infinite support, from (2) we find that the matrix of U_μ with respect to $\mathbf{\Phi}$ is a Hessenberg one given by [9,2,17,27,34]

$$H(\mathbf{a}) := \begin{pmatrix} -a_1 & -\rho_1 a_2 & -\rho_1\rho_2 a_3 & -\rho_1\rho_2\rho_3 a_4 & \cdots \\ \rho_1 & -\overline{a}_1 a_2 & -\overline{a}_1\rho_2 a_3 & -\overline{a}_1\rho_2\rho_3 a_4 & \cdots \\ & \rho_2 & -\overline{a}_2 a_3 & -\overline{a}_2\rho_3 a_4 & \cdots \\ & & \rho_3 & -\overline{a}_3 a_4 & \cdots \\ & & & \rho_4 & \cdots \\ & & & & \cdots \end{pmatrix},$$

where $\mathbf{a}$ is the corresponding sequence of Schur parameters.

The principal matrix of order N of $H(\mathbf{a})$ only depends on the vector $\mathbf{a}_N$ and will be denoted by $H(\mathbf{a}_N)$. If $\mathbf{a}_N \in \mathbb{D}^{N-1} \times \mathbb{T}$ and μ is the related finitely supported measure we get from (2) and (3) that $z\mathbf{\Phi}_N(z) = H(\mathbf{a}_N)^{\mathrm{T}}\mathbf{\Phi}_N(z) + \kappa_{N-1}\psi(z)e_N$, $\mathbf{\Phi}_N$ being the corresponding finite segment of OP [5]. Since $\psi(z) = 0$ μ-a.e., we find that $H(\mathbf{a}_N)$ is the matrix of U_μ with respect to $\mathbf{\Phi}_N$.

Apart from its complexity, the infinite matrix $H(\mathbf{a})$ represents the full operator U_μ only when the polynomials are dense in L^2_μ, that is, when $\mathbf{a} \notin \ell^2$ [10,33]. In general, $H(\mathbf{a})$ represents the restriction of U_μ to the closure of $\mathbb{P} := \operatorname{span}\{z^n\}_{n\geqslant 0}$. Hence, although $H(\mathbf{a})$ is always isometric, it is unitary iff $z\overline{\mathbb{P}} = \overline{\mathbb{P}}$. Since this condition is equivalent to $\overline{\mathbb{P}} = L^2_\mu$, we see that $H(\mathbf{a})$ is unitary iff $\mathbf{a} \notin \ell^2$.

The measures corresponding to sequences $\mathbf{a} \in \ell^2$ constitute the so-called Szegö class. A possibility of getting a matrix representation for U_μ in this case is to enlarge the OP basis to get an orthonormal basis of L^2_μ. This possibility is exploited in [31], obtaining a doubly infinite unitary matrix in which $H(\mathbf{a})$ is embedded. Anyway, the complexity of the matrix representation remains.

If we want to simplify the matrix representation of U_μ solving at the same time the problem for the Szegö class, we have to change completely the choice of the basis for L^2_μ. Since the space of Laurent polynomials is always dense in L^2_μ, a more natural choice for a basis is the orthogonal Laurent polynomials (OLP) with respect to μ, related to the corresponding OP by [6,35]

$$\chi_{2k}(z) = z^{-k}\varphi_{2k}^*(z), \quad \chi_{2k+1} = z^{-k}\varphi_{2k+1}(z), \quad k \geqslant 0. \tag{4}$$

The above relation gives a finite segment of OLP $\mathbf{X}_N := (\chi_n)_{n=0}^{N-1}$ in the case of a measure supported on N points, or a sequence $\mathbf{X} := (\chi_n)_{n\geqslant 0}$ of OLP for an infinitely supported measure. $\mathbf{X}_N$ and $\mathbf{X}$ always constitute an orthonormal basis of the corresponding space L^2_μ.

If the measure μ has infinite support, we get from (2) the following matrix representation for the operator U_μ with respect to the related sequence $\mathbf{X}$ of OLP [6]

$$C(\mathbf{a}) := \begin{pmatrix} -a_1 & -\rho_1 a_2 & \rho_1\rho_2 & & & & & \\ \rho_1 & -\overline{a}_1 a_2 & \overline{a}_1\rho_2 & 0 & & & & \\ 0 & -\rho_2 a_3 & -\overline{a}_2 a_3 & -\rho_3 a_4 & \rho_3\rho_4 & & & \\ & \rho_2\rho_3 & \overline{a}_2\rho_3 & -\overline{a}_3 a_4 & \overline{a}_3\rho_4 & 0 & & \\ & & 0 & -\rho_4 a_5 & -\overline{a}_4 a_5 & -\rho_5 a_6 & \rho_5\rho_6 & \\ & & & \rho_4\rho_5 & \overline{a}_4\rho_5 & -\overline{a}_5 a_6 & \overline{a}_5\rho_6 & 0 \\ & & & & \ddots & \ddots & \ddots & \ddots & \ddots \end{pmatrix},$$

$\mathbf{a}$ being the corresponding sequence of Schur parameters.

Now we deal with a five-diagonal matrix that, contrary to the Hessenberg one, always represents the full operator U_μ and, hence, is unitary for any $\mathbf{a} \in \mathbb{D}^{\aleph_0}$. Besides, it has a much simpler dependence on the Schur parameters.

The principal matrix of order N of $C(\mathbf{a})$, that only depends on $\mathbf{a}_N$, will be denoted by $C(\mathbf{a}_N)$. Analogously to the case of the Hessenberg representation, if $\mathbf{a}_N \in \mathbb{D}^{N-1} \times \mathbb{T}$ and μ is the related measure, $C(\mathbf{a}_N)$ is the matrix of U_μ with respect to the corresponding finite segment of OLP $\mathbf{X}_N$: using (2)–(4) we find that $z\mathbf{X}_N(z) = C(\mathbf{a}_N)^{\mathrm{T}}\mathbf{X}_N(z) + \mathbf{b}_N z^{-[\frac{N-1}{2}]}\psi(z)$, where

$$\mathbf{b}_N = \begin{cases} \kappa_{N-1}e_N & \text{if } N \text{ is even,} \\ \kappa_{N-1}(\rho_{N-1}e_{N-1} + \overline{a}_{N-1}e_N) & \text{if } N \text{ is odd,} \end{cases}$$

and, since $\psi(z) = 0$ μ-a.e., we get the desired result.

Let μ be a measure on $\mathbb{T}$, (φ_n) the corresponding OP and (χ_n) the related OLP. As a consequence of the whole previous discussion, if μ is associated with the sequence $\mathbf{a} \in \mathbb{D}^{\aleph_0}$ of Schur parameters, $\sigma(H(\mathbf{a})) = \operatorname{supp}\mu$ for $\mathbf{a} \notin \ell^2$ while $\sigma(C(\mathbf{a})) = \operatorname{supp}\mu$ always happens. Also, if μ is the finitely supported measure associated with $\mathbf{a}_N \in \mathbb{D}^{N-1} \times \mathbb{T}$, then $\sigma(H(\mathbf{a}_N)) = \sigma(C(\mathbf{a}_N)) = \operatorname{supp}\mu$. Similar relations hold between the mass points of the measure and the eigenvalues of the related matrices. As for the eigenvectors associated with a mass point λ, since $\langle \mathscr{X}_\lambda, \varphi_n\rangle_\mu = \mu(\{\lambda\})\overline{\varphi_n(\lambda)}$ and $\langle \mathscr{X}_\lambda, \chi_n\rangle_\mu = \mu(\{\lambda\})\overline{\chi_n(\lambda)}$, we find that $\sum_n \overline{\varphi_n(\lambda)}e_n$ is an eigenvector of the corresponding Hessenberg matrix when it represents the full operator U_μ, while $\sum_n \overline{\chi_n(\lambda)}e_n$ is always an eigenvector of the related five-diagonal matrix.

Let λ be a mass point of μ. Using the decomposition of $\mathscr{X}_\lambda$ with respect to the OLP basis we find that $\mu(\{\lambda\}) = \langle \mathscr{X}_\lambda, \mathscr{X}_\lambda\rangle_\mu = \mu(\{\lambda\})^2 \sum_n |\chi_n(\lambda)|^2$ and, so, since $\lambda \in \mathbb{T}$, we get from the relation (4) between OP and OLP that $\mu(\{\lambda\}) = 1/\sum_n |\varphi_n(\lambda)|^2$. Thus, when μ has infinite support, $\mathbf{\Phi}(\lambda) \in \ell^2$ if λ is a mass point.

Conversely, if $\lambda \in \mathbb{T}$ is such that $\mathbf{\Phi}(\lambda) \in \ell^2$, then λ is a mass point since $C(\mathbf{a})^*\overline{\mathbf{X}(z)} = \overline{z}\overline{\mathbf{X}(z)}$ $\forall z \in \mathbb{C}$. Notice that these arguments also work using the OP basis but restricted to measures outside the Szegö class.

Among other things, the preceding results show that the minimal representations of unitary operators are at most five-diagonal, but, are they exactly five-diagonal?

Moreover, like any unitary operator, every Hessenberg matrix that is unitary must be unitarily equivalent to a five-diagonal one. However, a question remains if we want to complete the "five-diagonal reduction" of the spectral problem for any unitary Hessenberg matrix: which one is the five-diagonal matrix related to an arbitrary unitary Hessenberg one?

A deeper study of unitary five-diagonal and Hessenberg matrices will answer the above questions.

3. Five-diagonal and Hessenberg matrices

The five-diagonal matrices presented in the previous section are examples of the following kind of matrices, that can be considered an intermediate step between the five-diagonal and the tri-diagonal case.

Definition 3.1. A (finite or infinite) five-diagonal matrix $C = (c_{i,j})$ is called para-tridiagonal if $c_{2k,2k+2} = c_{2k+1,2k-1} = 0$ $\forall k \geqslant 1$, that is,

$$C = \begin{pmatrix} c_{1,1} & c_{1,2} & c_{1,3} & & & & & & \\ c_{2,1} & c_{2,2} & c_{2,3} & 0 & & & & & \\ 0 & c_{3,2} & c_{3,3} & c_{3,4} & c_{3,5} & & & & \\ & c_{4,2} & c_{4,3} & c_{4,4} & c_{4,5} & 0 & & & \\ & & 0 & c_{5,4} & c_{5,5} & c_{5,6} & c_{5,7} & & \\ & & & c_{6,4} & c_{6,5} & c_{6,6} & c_{6,7} & 0 & \\ & & & & \ddots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}.$$

C is called irreducible if $c_{2,1} \neq 0$ and $c_{2k-1,2k+1}, c_{2k+2,2k} \neq 0$ $\forall k \geqslant 1$.

Unitary irreducible para-tridiagonal matrices, like unitary irreducible Hessenberg ones, have e_1 as a cyclic vector (in fact, any vector e_n, $n \in \mathbb{N}$, is cyclic in the first case). Therefore, any unitary irreducible para-tridiagonal matrix is the matrix representation of a multiplication operator on $\mathbb{T}$ and, hence, is unitarily equivalent to one with the form $C(\mathbf{a})$ or $C(\mathbf{a}_N)$. However we do not know how to describe this relation exactly. The following result is the first step to answer this and the previous questions, since it provides the general form of infinite unitary para-tridiagonal and Hessenberg matrices. The matrix representations introduced in the preceding section are indispensable guides for taking this step.

Theorem 3.2. *An infinite para-tridiagonal* (*Hessenberg*) *matrix is unitary* (*isometric*) *iff it has the form* $C(\mathbf{a},\mathbf{b})$ ($H(\mathbf{a},\mathbf{b})$), *where* $\mathbf{a},\mathbf{b}\in\mathbb{C}^{\aleph_0}$ *are such that* $|a_n|^2+|b_n|^2=1\ \forall n\in\mathbb{N}$, *and*

$$C(\mathbf{a},\mathbf{b}) := \begin{pmatrix} -a_1 & -\overline{b}_1a_2 & \overline{b}_1b_2 & & & & & \\ b_1 & -\overline{a}_1a_2 & \overline{a}_1b_2 & 0 & & & & \\ 0 & -\overline{b}_2a_3 & -\overline{a}_2a_3 & -\overline{b}_3a_4 & \overline{b}_3b_4 & & & \\ & \overline{b}_2b_3 & \overline{a}_2b_3 & -\overline{a}_3a_4 & \overline{a}_3b_4 & 0 & & \\ & & 0 & -\overline{b}_4a_5 & -\overline{a}_4a_5 & -\overline{b}_5a_6 & \overline{b}_5b_6 & \\ & & & \overline{b}_4b_5 & \overline{a}_4b_5 & -\overline{a}_5a_6 & \overline{a}_5b_6 & 0 \\ & & & & \ddots & \ddots & \ddots & \ddots & \ddots \end{pmatrix},$$

$$H(\mathbf{a},\mathbf{b}) := \begin{pmatrix} -a_1 & -\overline{b}_1a_2 & -\overline{b}_1\overline{b}_2a_3 & -\overline{b}_1\overline{b}_2\overline{b}_3a_4 & \cdots \\ b_1 & -\overline{a}_1a_2 & -\overline{a}_1\overline{b}_2a_3 & -\overline{a}_1\overline{b}_2\overline{b}_3a_4 & \cdots \\ & b_2 & -\overline{a}_2a_3 & -\overline{a}_2\overline{b}_3a_4 & \cdots \\ & & b_3 & -\overline{a}_3a_4 & \cdots \\ & & & b_4 & \cdots \\ & & & & \cdots \end{pmatrix}.$$

Proof. An infinite para-tridiagonal matrix C can be written in the way

$$C = \begin{pmatrix} C_1^{\mathrm{T}} & C_2 & 0 & \cdots \\ 0 & C_3^{\mathrm{T}} & C_4 & \cdots \\ 0 & 0 & C_5^{\mathrm{T}} & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix}, \quad \begin{array}{l} C_1\in\mathbb{C}^{(1,2)}, \\ C_n\in\mathbb{C}^{(2,2)},\ n\geqslant 2. \end{array}$$

It is unitary iff $C^*C=CC^*=I$, which is equivalent to

$$\begin{aligned} &C_1C_1^*=I_1, \\ &C_nC_n^*+(C_{n-1}^*C_{n-1})^{\mathrm{T}}=I_2, \quad n\geqslant 2, \\ &C_n^*C_{n-1}^{\mathrm{T}}=0, \quad n\geqslant 2. \end{aligned}$$

The first condition means that

$$C_1=\begin{pmatrix}-a_1 & b_1\end{pmatrix}, \quad |a_1|^2+|b_1|^2=1,$$

and, then, by induction, we find that the rest of the equations are satisfied iff

$$C_n=\begin{pmatrix} -\overline{b}_{n-1}a_n & \overline{b}_{n-1}b_n \\ -\overline{a}_{n-1}a_n & \overline{a}_{n-1}b_n \end{pmatrix}, \quad |a_n|^2+|b_n|^2=1, \ \ n\geqslant 2.$$

This proves the theorem in the para-tridiagonal case.

Now, let H be a Hessenberg matrix, that is, its nth column h_n belongs to span$\{e_1, e_2,\dots,e_{n+1}\}\subset\ell^2$. H is isometric iff $H^*H=I$, which means that $\{h_n\}_{n\in\mathbb{N}}$ is an orthonormal set of ℓ^2. We will see that this is equivalent to

$$
\begin{aligned}
&h_n = b_n e_{n+1} - a_n v_n, \quad n \in \mathbb{N} \\
&v_n = \sum_{i=1}^{n} \overline{a}_{i-1}\overline{b}_i\overline{b}_{i+1}\cdots\overline{b}_{n-1}e_i, \quad n \in \mathbb{N}, \\
&a_0 = 1; \quad |a_n|^2 + |b_n|^2 = 1, \ \ n \in \mathbb{N},
\end{aligned}
\tag{5}
$$

which proves the theorem for the Hessenberg case.

First of all, let us suppose that the columns of H have the form (5). From the expression of v_n we find that $v_{n+1} = \overline{b}_n v_n + \overline{a}_n e_{n+1}$ for $n \in \mathbb{N}$. Therefore, we get by induction that $v_n \perp \{h_1, h_2, \ldots, h_{n-1}\}$ and $(v_n, v_n) = 1$ for $n \in \mathbb{N}$. Then, the expression for h_n implies that $h_n \perp \{h_1, h_2, \ldots, h_{n-1}\}$ and $(h_n, h_n) = 1$ for $n \in \mathbb{N}$.

On the other hand, if the columns of H form an orthonormal set of ℓ^2, then we can write $h_1 = -a_1 e_1 + b_1 e_2$, $|a_1|^2 + |b_1|^2 = 1$, and, for $n \geqslant 2$, $h_n = b_n e_{n+1} + u_n$, $b_n \in \mathbb{C}$, $u_n \in \{h_1, h_2, \ldots, h_{n-1}\}^{\perp \operatorname{span}\{e_1, e_2, \ldots, e_n\}}$. So, h_1 has the form given by (5). Moreover, let us suppose that $h_1, h_2, \ldots, h_{n-1}$ satisfy (5). Then, $\{h_1, h_2, \ldots, h_{n-1}\}^{\perp \operatorname{span}\{e_1, e_2, \ldots, e_n\}} = \operatorname{span}\{v_n\}$ and we find that $u_n = -a_n v_n$, $a_n \in \mathbb{C}$. The condition $(h_n, h_n) = 1$ gives $|a_n|^2 + |b_n|^2 = 1$. This proves by induction that h_n has the form (5) for $n \in \mathbb{N}$. $\square$

A consequence of Theorem 3.2 is its analogue for finite matrices. The result for the Hessenberg case was already known [18]. In what follows, since the principal submatrix of order N of $C(\mathbf{a}, \mathbf{b})$ ($H(\mathbf{a}, \mathbf{b})$) only depends on $\mathbf{a}_N, \mathbf{b}_{N-1}$, this submatrix will be denoted by $C(\mathbf{a}_N, \mathbf{b}_{N-1})$ ($H(\mathbf{a}_N, \mathbf{b}_{N-1})$).

Corollary 3.3. *A finite para-tridiagonal* (*Hessenberg*) *matrix of order N is unitary iff it has the form $C(\mathbf{a}_N, \mathbf{b}_{N-1})$ ($H(\mathbf{a}_N, \mathbf{b}_{N-1})$), where $|a_n|^2 + |b_n|^2 = 1$ for $1 \leqslant n \leqslant N-1$ and $|a_N| = 1$.*

Proof. This result is just a direct consequence of Theorem 3.2 and the following facts: a finite square matrix M is unitary iff the infinite matrix $M \oplus I$ is unitary; the matrices $C(\mathbf{a}, \mathbf{b})$ and $H(\mathbf{a}, \mathbf{b})$, with $|a_n|^2 + |b_n|^2 = 1$ $\forall n \in \mathbb{N}$, decompose as a direct sum of their principal submatrices of order N and an infinite matrix iff $b_N = 0$. $\square$

Remark 3.4 (*Decomposition property*). Unitary para-tridiagonal and isometric Hessenberg matrices have similar decomposition properties. They decompose as a sum of diagonal blocks iff, for some N, $b_N = 0$, that is, $a_N \in \mathbb{T}$. Moreover, in this situation, the blocks must again be unitary para-tridiagonal and isometric Hessenberg matrices, respectively. In fact, if $a_N \in \mathbb{T}$,

$$
C(\mathbf{a}, \mathbf{b}) = \begin{cases} C(\mathbf{a}_N, \mathbf{b}_{N-1}) \oplus C\big(\overline{a}_N \mathbf{a}^{(N)}, \overline{a}_N \mathbf{b}^{(N)}\big), & \text{even } n, \\ C(\mathbf{a}_N, \mathbf{b}_{N-1}) \oplus C\big(\overline{a}_N \mathbf{a}^{(N)}, \overline{a}_N \mathbf{b}^{(N)}\big)^{\mathrm{T}}, & \text{odd } n, \end{cases}
$$

$$
C(\mathbf{a}_{N+M}, \mathbf{b}_{N+M-1}) = \begin{cases} C(\mathbf{a}_N, \mathbf{b}_{N-1}) \oplus C\big(\overline{a}_N \mathbf{a}_M^{(N)}, \overline{a}_N \mathbf{b}_{M-1}^{(N)}\big), & \text{even } n, \\ C(\mathbf{a}_N, \mathbf{b}_{N-1}) \oplus C\big(\overline{a}_N \mathbf{a}_M^{(N)}, \overline{a}_N \mathbf{b}_{M-1}^{(N)}\big)^{\mathrm{T}}, & \text{odd } n, \end{cases}
$$

where $\mathbf{a}^{(N)} = (a_{N+n})_{n\in\mathbb{N}}$, $\mathbf{a}_M^{(N)} = (a_{N+1}, a_{N+2}, \ldots, a_{N+M})$ and analogously for $\mathbf{b}$. In the Hessenberg case, if $a_N \in \mathbb{T}$,

$$H(\mathbf{a}, \mathbf{b}) = H(\mathbf{a}_N, \mathbf{b}_{N-1}) \oplus H\big(\overline{a}_N \mathbf{a}^{(N)}, \mathbf{b}^{(N)}\big),$$

$$H(\mathbf{a}_{N+M}, \mathbf{b}_{N+M-1}) = H(\mathbf{a}_N, \mathbf{b}_{N-1}) \oplus H\big(\overline{a}_N \mathbf{a}_M^{(N)}, \mathbf{b}_{M-1}^{(N)}\big).$$

Remark 3.5 (*Factorization property*). For $a, b \in \mathbb{C}$, let us define

$$\Theta(a, b) := \begin{pmatrix} -a & \overline{b} \\ b & \overline{a} \end{pmatrix}, \quad \hat{\Theta}_n(a, b) := I_{n-1} \oplus \Theta(a, b) \oplus I.$$

Then, for any bounded sequences $\mathbf{a}, \mathbf{b} \in \mathbb{C}^{\aleph_0}$,

$$C(\mathbf{a}, \mathbf{b}) = C_o(\mathbf{a}, \mathbf{b}) C_e(\mathbf{a}, \mathbf{b})^{\mathrm{T}}, \quad H(\mathbf{a}, \mathbf{b}) = \prod_{n=1}^{\infty} \hat{\Theta}_n(a_n, b_n),$$

where the infinite product, which has to be understood in the strong sense, is from the left to the right, and

$$C_e(\mathbf{a}, \mathbf{b}) = I_1 \oplus \left(\bigoplus_{n\in\mathbb{N}} \Theta(a_{2n}, b_{2n})\right), \quad C_o(\mathbf{a}, \mathbf{b}) = \bigoplus_{n\in\mathbb{N}} \Theta(a_{2n-1}, b_{2n-1}).$$

These factorizations show explicitly the isometric properties of the matrices given in Theorem 3.2 and Corollary 3.3.

We also denote $\Theta(a) := \Theta(a, \sqrt{1-|a|^2})$, $\hat{\Theta}_n(a) := I_{n-1} \oplus \Theta(a) \oplus I$, so that $H(\mathbf{a}) = \prod_{n=1}^{\infty} \hat{\Theta}_n(a_n)$ and $C(\mathbf{a}) = C_o(\mathbf{a}) C_e(\mathbf{a})$, where $C_e(\mathbf{a}) := I_1 \oplus (\bigoplus_{n\in\mathbb{N}} \Theta(a_{2n}))$ and $C_o(\mathbf{a}) := \bigoplus_{n\in\mathbb{N}} \Theta(a_{2n-1})$.

In the case of finite unitary Hessenberg matrices, the above properties have been used for spectral computations [18,19]. Notice that the factorization property in the Hessenberg case is much worse than in the para-tridiagonal one.

We know that any unitary matrix represents an orthogonal sum of multiplication operators on $\mathbb{T}$ and, hence, is unitarily equivalent to a direct sum of unitary irreducible para-tridiagonal matrices. However, if the initial matrix is also para-tridiagonal, the equivalence becomes an equality. This is just a consequence of Theorem 3.2, Corollary 3.3 and the decomposition property given in Remark 3.4. For the same reason, a similar result is also true for isometric Hessenberg matrices.

Corollary 3.6. *Every unitary para-tridiagonal matrix is a direct sum of irreducible and transposed irreducible ones. Every isometric Hessenberg matrix is a direct sum of irreducible ones.*

Even more, in the study of irreducible unitary para-tridiagonal (isometric Hessenberg) matrices, it is enough to consider those with the form $C(\mathbf{a})$ ($H(\mathbf{a})$) and their principal submatrices. More precisely, we have the following immediate result.

Lemma 3.7. *For any* $\mathbf{a} \in \mathbb{C}^{\aleph_0}, \mathbf{b} \in (\mathbb{C}\backslash\{0\})^{\aleph_0}$ *we have* $H(\mathbf{a}) = R^* H(\mathbf{a}, \mathbf{b}) R$ *and* $C(\mathbf{a}) = S^* C(\mathbf{a}, \mathbf{b}) S$, *where*

$$R = \begin{pmatrix} r_1 & & & \\ & r_2 & & \\ & & r_3 & \\ & & & \ddots \end{pmatrix}, \qquad \begin{array}{l} r_1 = 1, \\ r_{n+1}/r_n = b_n/|b_n|, \ n \geqslant 1, \end{array}$$

$$S = \begin{pmatrix} \overline{s}_1 & & & \\ & s_2 & & \\ & & \overline{s}_3 & \\ & & & \ddots \end{pmatrix}, \qquad \begin{array}{l} s_1 = 1, \\ s_2 = b_1/|b_1|, \\ s_{n+1}/s_{n-1} = \overline{b}_{n-1} b_n / |\overline{b}_{n-1} b_n|, \ n \geqslant 2. \end{array}$$

If R_N, S_N *are the principal submatrices of order* N *of* R, S *respectively*, $H(\mathbf{a}_N) = R_N^* H(\mathbf{a}_N, \mathbf{b}_{N-1}) R_N$ *and* $C(\mathbf{a}_N) = S_N^* H(\mathbf{a}_N, \mathbf{b}_{N-1}) S_N$.

Notice that Theorem 3.2 and the above lemma imply that an infinite Hessenberg matrix is unitary iff it has the form $H(\mathbf{a}, \mathbf{b})$ with $|a_n|^2 + |b_n|^2 = 1$ and $\mathbf{a} \notin \ell^2$.

The preceding results have the following consequence, that represents the "five-diagonal reduction" of the spectral problem for any unitary Hessenberg matrix. Without loss of generality we consider only the irreducible case.

Theorem 3.8. *Let* $H = (h_{i,j})$ *be a (finite or infinite) unitary irreducible Hessenberg matrix and let us define*

$$\tau_n := \begin{cases} 1, & n = 1, \\ \prod_{k=1}^{n-1} h_{k+1,k}, & n \geqslant 2. \end{cases}$$

Then, H is unitarily equivalent to a para-tridiagonal matrix $C = (c_{i,j})$ with the form $C(\mathbf{a})$ or $C(\mathbf{a}_N)$, where

$$a_n = -\frac{h_{1,n}}{\overline{\tau}_n}, \quad n \geqslant 1.$$

The unitary equivalence is given by $H = V^* C V$, where the columns of $V = (v_{i,j})$ can be recursively obtained by

$$v_{i,j} = \begin{cases} \frac{\overline{\tau}_j}{|\tau_j|} \delta_{i,j}, & j = 1, 2, \\ \frac{1}{h_{j,j-1}} \left(\displaystyle\sum_{k=i-2}^{\min\{i+2,2j-4\}} c_{i,k} v_{k,j-1} - \sum_{k=\min\left\{i, \left[\frac{i+3}{2}\right]\right\}}^{j-1} h_{k,j-1} v_{i,k} \right), & i \leqslant 2j-2, \ j \geqslant 3, \\ 0, & i \geqslant 2j-1, \ j \geqslant 3. \end{cases}$$

In the above expression the sums have to be understood only over those terms in which the matrix coefficients have indices between 1 *and the order of* H. *The eigen-*

vectors $x(\lambda) = \sum_n x_n(\lambda)e_n$ *of* H *and* $y(\lambda) = \sum_n y_n(\lambda)e_n$ *of* C *corresponding to the same eigenvalue* λ *are related by*

$$x_{2k-1}(\lambda) = \frac{\tau_{2k-1}}{|\tau_{2k-1}|}\lambda^{1-k}\overline{y_{2k-1}(\lambda)}, \quad x_{2k}(\lambda) = \frac{\tau_{2k}}{|\tau_{2k}|}\lambda^{1-k}y_{2k}(\lambda), \quad k \geqslant 1.$$

Proof. We will consider only the case of an infinite matrix H, the proof for the finite case being completely analogous. Then, from Theorem 3.2, H must have the form $H(\mathbf{a}, \mathbf{b})$, $\mathbf{a} \in \mathbb{D}^{\aleph_0}\backslash\ell^2$. So, according to Lemma 3.7, H is unitarily equivalent to $H(\mathbf{a})$ which, at the same time, is unitarily equivalent to $C(\mathbf{a})$ since they are representations of the same multiplication operator.

We know that $H = RH(\mathbf{a})R^*$, $R = (r_i\delta_{i,j})$, where $r_i = \tau_i/|\tau_i|$ since $h_{i+1,i} = b_i$. Besides, if μ is the measure related to the sequence $\mathbf{a}$ of Schur parameters and $(\varphi_n)_{n\geqslant 0}$, $(\chi_n)_{n\geqslant 0}$ are the corresponding OP, OLP respectively, then $H(\mathbf{a}) = U^*C(\mathbf{a})U$, $U = (u_{i,j})$, $u_{i,j} = \langle\varphi_{j-1}, \chi_{i-1}\rangle_\mu$. Therefore, $H = V^*C(\mathbf{a})V$, $V = UR^*$. For $j = 1, 2$, $\varphi_{j-1} = \chi_{j-1}$ and, so, $v_{i,j} = \overline{r}_j\delta_{i,j}$. For the rest of the columns in V, if $j \geqslant 2$, $\varphi_j \in \operatorname{span}\{1, z, z^{-1}, \ldots, z^{1-j}, z^j\} = \operatorname{span}\{\chi_0, \chi_1, \ldots, \chi_{2j-1}\}$ and, thus, $\langle\varphi_j, \chi_i\rangle_\mu = 0$ for $i \geqslant 2j$. Since R is diagonal, this implies that $v_{i,j} = 0$ for $i \geqslant 2j - 1$, $j \geqslant 3$. Moreover, from the equality $VH = C(\mathbf{a})V$ we get

$$\sum_{k=1}^{j} v_{i,k}h_{k,j-1} = \sum_{k=i-2}^{i+2} c_{i,k}v_{k,j-1}, \quad i \leqslant 2j-2,\ j \geqslant 3,$$

which completes the expression given for $v_{i,j}$, once the restriction $v_{i,j} = 0$, $i \geqslant \max\{j+1, 2j-1\}$, is taken into account in the above sums.

Given an eigenvalue λ of $H = RH(\mathbf{a})R^*$, the corresponding eigenvectors are spanned by $\sum_{n\in\mathbb{N}} r_n\overline{\varphi_{n-1}(\lambda)}e_n$, while the eigenvectors of $C(\mathbf{a})$ are spanned by $\sum_{n\in\mathbb{N}} \overline{\chi_{n-1}(\lambda)}e_n$. Hence, the referred relation between eigenvectors is just a consequence of the relation (4) between OP and OLP. □

The para-tridiagonal representations improve the Hessenberg representations of unitary operators because of their greater simplicity. Besides, as we pointed out in Remark 3.4, they have similar decomposition properties and, thus, "Divide and Conquer" algorithms [19] can be also developed for the spectral problem of a unitary para-tridiagonal matrix. Even more, the factorization given in Remark 3.5 allows to write the corresponding five-diagonal eigenvalue problem equivalently as a generalized eigenvalue problem for a tri-diagonal pair of unitary matrices.

Now we reach the announced result about the minimal representations of unitary operators.

Theorem 3.9. *A* (p, q)*-diagonal unitary matrix is a sum of diagonal blocks of order not greater than* $p + q$ *if* p *or* q *are equal to* 1.

Proof. We can restrict our attention to the case of $(1, q)$-diagonal matrices since, otherwise, we can deal with the adjoint matrix, that keeps the unitarity. Also, it is enough to prove that, if such a unitary band matrix has order greater than $q + 1$, then it must decompose as a sum of smaller diagonal blocks. Let us suppose that $\Omega = (\omega_{i,j})$ is of order greater that $q + 1$ and does not decompose. Ω is, in particular, an isometric Hessenberg matrix and, hence, $\Omega = H(\mathbf{a}, \mathbf{b})$ or $\Omega = H(\mathbf{a}_N, \mathbf{b}_{N-1})$, $N \geqslant q + 2$. Since it does not decompose, $b_n \neq 0$ for all n. Thus, for $j \geqslant q + 2$, $\omega_{1,j} = -\prod_{k=1}^{j-1} \overline{b}_k a_j = 0$ implies $a_j = 0$. Hence, if $i \leqslant j$, $\omega_{i,j} = -\overline{a}_{i-1} a_j \prod_{k=i}^{j-1} \overline{b}_k = 0$ for $j \geqslant q + 2$. Therefore, $\{\Omega^* e_1, \Omega^* e_2, \ldots, \Omega^* e_{q+2}\} \subset \operatorname{span}\{e_1, e_2, \ldots, e_{q+1}\}$, which is a contradiction with the unitarity of Ω. $\square$

A matrix that decomposes as a sum of finite diagonal blocks has always pure point spectrum. Therefore, the previous theorem shows that the only (p, q)-diagonal representations possible for any unitary operator are those where $p, q \geqslant 2$. Consequently, we have the following corollary.

Corollary 3.10. *The minimal representations of unitary operators are five-diagonal.*

4. Krein's Theorem

One of the advantages of band representations is that they make it easier to decide the "smallness" of a perturbation. For example, the compactness of an operator is equivalent to stating that the diagonals of a band representation converge to 0. This makes it quite simple, for example, to apply Weyl's Theorem [36,26,29] for the invariance of the essential spectrum. Also, it is easier to prove that a perturbation belongs to the trace class, which can be used to give a simple and elegant operator theoretic proof of Rakhmanov's lemma [31] using the Kato–Rosenblum Theorem [25,30,26,29] on the invariance of the absolutely continuous part of the spectrum of an operator.

In spite of the difficulties that appear, many results about the orthogonality measures of OP on $\mathbb{T}$ have been obtained using the Hessenberg representation [12,13,14, 16,34], mainly due to the efforts of Golinskii. The proofs of such results can be now simplified, but we want to show some new results and advantages provided by the para-tridiagonal representation in the analysis of the relation between measures and Schur parameters.

First of all, we will discuss the advantages found in the application of Krein's Theorem [1], getting new results for discrete measures whose support has a finite derived set. Krein's Theorem asserts that, given a measure μ with infinite support, it is equivalent to saying that $\operatorname{supp} \mu$ accumulates on the finite set $\{w_1, w_2, \ldots, w_N\}$ and that the operator $p_N(U_\mu)$ is compact, where $p_N(z) = \prod_{i=1}^{N} (z - w_i)$. This theorem was established in 1962 by Akhiezer and Krein [1] for measures on the real line with finite moments. Recently, the translation to the unit circle was given by Golinskii [12], who succeeded in characterizing in terms of the Schur parameters the measures whose support has one or two limit points, using the Hessenberg representation of U_μ. He also proved that the Schur parameters of any measure on $\mathbb{T}$ whose support has a finite number N of limit points must satisfy

$$\lim_n \rho_n \rho_{n+1} \cdots \rho_{n+N-1} = 0, \tag{6}$$

from which comes the property $\overline{\lim}_n |a_n| = 1$ for any measure whose support has a finite derived set.

However, with the Hessenberg representation it is hard to go further in this direction. The para-tridiagonal representation makes things easier, not only because of its band structure, but also due to its factorization properties. In the context of the para-tridiagonal representation, for the application of Krein's Theorem it is necessary to decide the compactness of $p_N(C(\mathbf{a}))$, where $\mathbf{a}$ is the sequence of Schur parameters associated with μ. This requires the calculation of the $4N + 1$ diagonals of $p_N(C(\mathbf{a}))$, some of them possibly giving redundant information. We can optimize the calculations using the factorization of Remark 3.5.

Proposition 4.1. *Given $w_1, w_2, \ldots, w_N \in \mathbb{T}$, let us define*

$$q_N(C(\mathbf{a})) = \begin{cases} C(\mathbf{a})^{*k} p_N(C(\mathbf{a})) & \text{if } N = 2k, \\ C_o(\mathbf{a})^* C(\mathbf{a})^{*k} p_N(C(\mathbf{a})) & \text{if } N = 2k+1, \end{cases}$$

where $p_N(z) = \prod_{i=1}^N (z - w_i)$. Then, $q_N(C(\mathbf{a}))$ is a $2N + 1$-diagonal matrix such that

$$q_N(C(\mathbf{a}))^* = \begin{cases} \left(\prod_{i=1}^N \overline{w}_i\right) q_N(C(\mathbf{a})) & \text{if } N \text{ is even}, \\ \overline{q_N(C(\mathbf{a}))} & \text{if } N \text{ is odd}, \end{cases}$$

and $p_N(C(\mathbf{a}))$ is compact iff $\lim_n q_N(C(\mathbf{a}))_{n+m,n} = 0$ for $m = 0, 1, \ldots, N$.

Proof. From the unitarity of $C(\mathbf{a})$ and $C_o(\mathbf{a})$, the equivalence between the compactness of $p_N(C(\mathbf{a}))$ and $q_N(C(\mathbf{a}))$ follows. The matrix $q_N(C(\mathbf{a}))$ is a linear combination of products of, at most, N tri-diagonal matrices, so, it is $2N + 1$-diagonal. Thus, $q_N(C(\mathbf{a}))$ is compact iff $\lim_n q_N(C(\mathbf{a}))_{n+m,n} = 0$ for $m = 0, \pm 1, \ldots, \pm N$. Hence, to finish the proof we just have to check the relations between $q_N(C(\mathbf{a}))$ and $q_N(C(\mathbf{a}))^*$.

When N is odd, $q_N(C(\mathbf{a}))$ is a linear combination of products of an odd number of alternate factors $C_o(\mathbf{a})$ and $C_e(\mathbf{a})$, or their adjoints. Since $C_o(\mathbf{a})$ and $C_e(\mathbf{a})$ are symmetric, $q_N(C(\mathbf{a}))$ is symmetric too.

In the case of even $N = 2k$, we can write $q_N(C(\mathbf{a})) = \prod_{i=1}^k r_i(C(\mathbf{a}))$, $r_i(C(\mathbf{a})) = (C(\mathbf{a}) - w_i)C(\mathbf{a})^*(C(\mathbf{a}) - w_{k+i})$. The result is just a consequence of the fact that $r_i(C(\mathbf{a}))^* = \overline{w}_i \overline{w}_{k+i} r_i(C(\mathbf{a}))$. $\square$

Therefore, we can apply Krein's Theorem imposing only that the main and lower diagonals of $q_N(C(\mathbf{a}))$ converge to 0, which will give in general $N+1$ asymptotic conditions for the Schur parameters of a measure whose support has N given limit points. For illustrative purposes we present the results achieved using this procedure when applied to the characterization of measures whose support has up to three limit points.

Proposition 4.2. *Let μ be the measure associated with the sequence $\mathbf{a}$ of Schur parameters. Then*:

1. $\{\operatorname{supp}\mu\}' = \{\alpha\}$ *iff*
   $\lim_n (\overline{a}_n a_{n+1} + \alpha) = 0.$
2. $\{\operatorname{supp}\mu\}' \subset \{\alpha, \beta\}$ *iff*
   $\lim_n \rho_n \rho_{n+1} = 0,$
   $\lim_n \rho_{n+1}(\overline{a}_n a_{n+2} - \alpha\beta) = 0,$
   $\lim_n (\overline{a}_n a_{n+1} + \alpha\beta a_n \overline{a}_{n+1} + \alpha + \beta) = 0.$
3. $\{\operatorname{supp}\mu\}' \subset \{\alpha, \beta, \gamma\}$ *iff*
   $\lim_n \rho_n \rho_{n+1} \rho_{n+2} = 0,$
   $\lim_n \rho_{n+1}\rho_{n+2}(\overline{a}_n a_{n+3} + \alpha\beta\gamma) = 0,$
   $\lim_n \rho_{n+1}(\overline{a}_n a_{n+1} + \overline{a}_{n+1} a_{n+2} - \alpha\beta\gamma a_n \overline{a}_{n+2} + \alpha + \beta + \gamma) = 0,$
   $\lim_n (\overline{a}_n a_{n+1}^2 - \rho_{n+1}^2 a_{n+2} + \alpha\beta\gamma(a_n^2 \overline{a}_{n+1} - a_{n-1}\rho_n^2) + (\alpha\beta + \beta\gamma + \gamma\alpha)a_n + (\alpha + \beta + \gamma)a_{n+1}) = 0.$

The first result of the above proposition is the same one obtained in [12], but the second assertion simplifies the one given in [12]. Notice that the relations given in the two last cases of the proposition also imply

$$\lim_n \rho_{n+1}\big(\overline{a}_n a_{n+1} + \overline{a}_{n+1} a_{n+2} + \alpha + \beta\big) = 0,$$

if $\{\operatorname{supp}\mu\}' \subset \{\alpha, \beta\}$, while, for $\{\operatorname{supp}\mu\}' \subset \{\alpha, \beta, \gamma\}$,

$$\lim_n \rho_{n+1}\rho_{n+2}\big(\overline{a}_n a_{n+1} + \overline{a}_{n+1} a_{n+2} + \overline{a}_{n+2} a_{n+3} + \alpha + \beta + \gamma\big) = 0.$$

The above results suggest the following improvement of the property (6) that gives a common feature for measures whose support has a finite derived set.

Theorem 4.3. *Let μ be the measure associated with the sequence $\mathbf{a}$ of Schur parameters. If $\{\operatorname{supp}\mu\}' \subset \{w_1, w_2, \ldots, w_N\}$, then*

$$\lim_n \prod_{i=1}^{N} \rho_{n+i} = 0, \quad N \geqslant 1,$$

$$\lim_n (\overline{a}_n a_{n+N} - P) \prod_{i=1}^{N-1} \rho_{n+i} = 0, \quad N \geqslant 2,$$

$$\lim_n \left(\sum_{j=1}^{N} \overline{a}_{n+j-1} a_{n+j} + S \right) \prod_{i=1}^{N-1} \rho_{n+i} = 0, \quad N \geqslant 2,$$

$$\lim_n \left(\sum_{j=1}^{N-1} \overline{a}_{n+j-1} a_{n+j} + P a_n \overline{a}_{n+N-1} + S \right) \prod_{i=1}^{N-2} \rho_{n+i} = 0, \quad N \geqslant 2,$$

where $P := (-1)^N w_1 w_2 \cdots w_N$ *and* $S := w_1 + w_2 + \cdots + w_N$.

Proof. We will consider only the case of even $N = 2k$, since the analysis for odd N is analogous. Then, the operator $q_N(C(\mathbf{a}))$ associated with the points $w_1, w_2, \ldots, w_N$, given in Proposition 4.1, has the form

$$q_N(C(\mathbf{a})) = C(\mathbf{a})^k - SC(\mathbf{a})^{k-1} + \cdots - \overline{S} P C(\mathbf{a})^{*k-1} + PC(\mathbf{a})^{*k}.$$

For $N - 3 \leqslant m \leqslant N$, $q_N(C(\mathbf{a}))_{n+m,n} = q_N^{(1)}(C(\mathbf{a}))_{n+m,n}$, where

$$q_N^{(1)}(C(\mathbf{a})) = C(\mathbf{a})^k - SC(\mathbf{a})^{k-1} - \overline{S} P C(\mathbf{a})^{*k-1} + PC(\mathbf{a})^{*k}.$$

So, $\lim_n q_N^{(1)}(C(\mathbf{a}))_{n+m,n} = 0$, $N - 3 \leqslant m \leqslant N$, under the hypothesis for μ.

Let us examine the coefficients $q_N^{(1)}(C(\mathbf{a}))_{n+m,n}$ for $m = N, N-1, N-2$. We can write $C_o(\mathbf{a}) = A_o + VB_o + B_o V^*$ and $C_e(\mathbf{a}) = A_e + VB_e + B_e V^*$, V being the right shift, defined by $Ve_n = e_{n+1}$, and

$$A_o e_n = \begin{cases} -a_n e_n, & \text{odd } n, \\ \overline{a}_{n-1} e_n, & \text{even } n, \end{cases} \qquad A_e e_n = \begin{cases} \overline{a}_{n-1} e_n, & \text{odd } n, \\ -a_n e_n, & \text{even } n, \end{cases}$$

$$B_o e_n = \begin{cases} \rho_n e_n, & \text{odd } n, \\ 0, & \text{even } n, \end{cases} \qquad B_e e_n = \begin{cases} 0, & \text{odd } n, \\ \rho_n e_n, & \text{even } n. \end{cases}$$

Taking into account that V, V^* rise and lower the indices of the vectors e_n, respectively, $V^* e_1 = 0$, and B_o, B_e vanish over vectors with even and odd index n, respectively, we find that

$$\begin{aligned} \left(q_N^{(1)}(C(\mathbf{a}))e_n, e_{n+N}\right) &= \left(((VB_oVB_e)^k + P(VB_eVB_o)^k)e_n, e_{n+N}\right) \\ &= \begin{cases} \rho_n \rho_{n+1} \cdots \rho_{n+N-1}, & \text{even } n, \\ P\rho_n \rho_{n+1} \cdots \rho_{n+N-1}, & \text{odd } n, \end{cases} \end{aligned}$$

$$\begin{aligned} &\left(q_N^{(1)}(C(\mathbf{a}))e_n, e_{n+N-1}\right) \\ &\quad = \left(((VB_oVB_e)^{k-1} VB_o A_e + A_o VB_e (VB_oVB_e)^{k-1})e_n, e_{n+N-1}\right) \\ &\qquad + P\left(((VB_eVB_o)^{k-1} VB_e A_o^* + A_e^* VB_o (VB_eVB_o)^{k-1})e_n, e_{n+N-1}\right) \\ &\quad = \begin{cases} -\rho_n \cdots \rho_{n+N-2}(a_{n+N-1} - Pa_{n-1}), & \text{even } n, \\ -P\rho_n \cdots \rho_{n+N-2}(\overline{a}_{n+N-1} - \overline{P}\overline{a}_{n-1}), & \text{odd } n. \end{cases} \end{aligned}$$

Therefore, $\lim_n \prod_{i=1}^{N} \rho_{n+i} = 0$ and $\lim_n (a_{n+N} - Pa_n) \prod_{i=1}^{N-1} \rho_{n+i} = 0$, which is equivalent to the first and second equalities of the theorem.

Concerning the coefficients $(q_N^{(1)}(C(\mathbf{a}))e_n, e_{n+N-2})$, we have that

$$
\begin{aligned}
&\big(C(\mathbf{a})^k e_n, e_{n+N-2}\big)\\
&\quad = \big((VB_oVB_e \cdots VB_oVB_eA_oA_e)e_n, e_{n+N-2}\big)\\
&\qquad + \big((VB_oVB_e \cdots VB_oA_eA_oVB_e)e_n, e_{n+N-2}\big)\\
&\qquad + \cdots + \big((A_oVB_e \cdots VB_oVB_eVB_oA_e)e_n, e_{n+N-2}\big)\\
&\quad = \begin{cases} -\rho_n \cdots \rho_{n+N-3}(\overline{a}_{n-1}a_n + \cdots + \overline{a}_{n+N-3}a_{n+N-2}), & \text{even } n,\\ -\rho_n \cdots \rho_{n+N-3}\overline{a}_{n-1}a_{n+N-2}, & \text{odd } n, \end{cases}
\end{aligned}
$$

$$
\begin{aligned}
&\big(C(\mathbf{a})^{*k} e_n, e_{n+N-2}\big)\\
&\quad = \big((VB_eVB_o \cdots VB_eVB_oA_e^*A_o^*)e_n, e_{n+N-2}\big)\\
&\qquad + \big((VB_eVB_o \cdots VB_eA_o^*A_e^*VB_o)e_n, e_{n+N-2}\big)\\
&\qquad + \cdots + \big((A_e^*VB_o \cdots VB_eVB_oVB_eA_o^*)e_n, e_{n+N-2}\big)\\
&\quad = \begin{cases} -\rho_n \cdots \rho_{n+N-3}a_{n-1}\overline{a}_{n+N-2}, & \text{even } n,\\ -\rho_n \cdots \rho_{n+N-3}(a_{n-1}\overline{a}_n + \cdots + a_{n+N-3}\overline{a}_{n+N-2}), & \text{odd } n, \end{cases}
\end{aligned}
$$

$$
\begin{aligned}
\big(C(\mathbf{a})^{k-1} e_n, e_{n+N-2}\big) &= \big((VB_oVB_e)^{k-1}e_n, e_{n+N-2}\big)\\
&= \begin{cases} \rho_n\rho_{n+1} \cdots \rho_{n+N-3}, & \text{even } n,\\ 0, & \text{odd } n, \end{cases}
\end{aligned}
$$

$$
\begin{aligned}
\big(C(\mathbf{a})^{*k-1} e_n, e_{n+N-2}\big) &= \big((VB_eVB_o)^{k-1}e_n, e_{n+N-2}\big)\\
&= \begin{cases} 0, & \text{even } n,\\ \rho_n\rho_{n+1} \cdots \rho_{n+N-3}, & \text{odd } n. \end{cases}
\end{aligned}
$$

From these results the last equality of the theorem follows. The third relation is just a consequence of the other ones. □

5. Perturbations of the Schur parameters and mass points

In the previous discussion we have exploited the band structure and factorization properties of the para-tridiagonal representation. Now we will also show the advantages of its simple dependence on the Schur parameters, in particular, of the fact that, contrary to the Hessenberg representation, any Schur parameter appears in only a finite number of elements of the para-tridiagonal representation.

The application to the study of OP of standard results of operator theory, like the Weyl, Krein or Kato–Rosenblum theorems, gives information about the limit points of the support of the orthogonality measure. However, for the analysis of isolated mass points other tools are more appropriate. This last section illustrates the usefulness of the para-tridiagonal representation for this purpose too. Our aim is to study the behaviour of the isolated mass points of the measure under mono-parametric perturbations of the Schur parameters using the Hellmann–Feynman Theorem [8,20,21,23].

Let us suppose a sequence $\mathbf{a}(t) \in \mathbb{D}^{\aleph_0}$ depending on $t \in I$, where I is an interval of $\mathbb{R}$. A measure μ^t corresponds to each value of t. The related OP and OLP sequences will be denoted by $\mathbf{\Phi}^t := (\varphi_n^t)_{n\geqslant 0}$ and $\mathbf{X}^t := (\chi_n^t)_{n\geqslant 0}$ respectively.

Besides, let $u(t)$ be a function of $t \in I$ with values on $\mathbb{T}$. For each t we can consider the finitely supported measure μ_N^t corresponding to the parameters $(a_1(t), \ldots, a_{N-1}(t), u(t))$, whose finite segments of OP and OLP are respectively $\mathbf{\Phi}_N^t := (\varphi_n^t)_{n=0}^{N-1}$ and $\mathbf{X}_N^t := (\varphi_n^t)_{n=0}^{N-1}$. The importance of such discrete measures is that they weakly converge to μ^t and, thus, they provide the, so called, Szegö quadrature formulas [24] for the measure μ^t.

We are interested in the evolution with t of the isolated mass points of μ^t, that is, the isolated eigenvalues of $\mathscr{C}(t) := C(\mathbf{a}(t))$. We will also analyze the movement of the mass points of the discrete approximations μ_N^t, that is, the eigenvalues of $\mathscr{C}_N(t) := C(\hat{\mathbf{a}}_N(t))$, $\hat{\mathbf{a}}_N(t) := (a_1(t), \ldots, a_{N-1}(t), u(t))$.

Since the finite matrices $\mathscr{C}_N(t)$ have N different eigenvalues, in any interval where $\hat{\mathbf{a}}_N(t)$ is differentiable with respect to t, its eigenvalues are differentiable functions $\lambda(t)$ [26]. Moreover, the corresponding eigenvectors $\mathbf{X}_N^t(\lambda(t))$ are also differentiable in t, since $\mathbf{X}_N(z)$ is a differentiable function of $a_1, \ldots, a_{N-1}, z$.

Concerning the infinite matrix $\mathscr{C}(t)$, a similar result holds, but only locally. More precisely, let us suppose that $\mathscr{C}(t)$ is differentiable in norm with bounded derivative $\mathscr{C}'(t)$ and $\|\mathscr{C}'(t)\|$ locally bounded. Then, if λ_0 is an isolated eigenvalue of $\mathscr{C}(t_0)$, there exists a neighbourhood of t_0 where $\mathscr{C}(t)$ has an isolated eigenvalue $\lambda(t)$ which is differentiable and such that $\lambda(t_0) = \lambda_0$. Moreover, a related eigenvector can be chosen as a strongly differentiable function of t in a neighbourhood of t_0 [22].

This last discussion justifies the following lemma.

Lemma 5.1. *Let $\Omega(t) = (\omega_{i,j}(t))_{i,j\in\mathbb{N}}$ be a bounded band matrix depending on a parameter $t \in I$, I being an interval of $\mathbb{R}$. Assume that the coefficients $\omega_{i,j}(t)$ are twice differentiable and $\sup_{i,j\in\mathbb{N}} |\omega'_{i,j}(t)|$, $\sup_{i,j\in\mathbb{N}} |\omega''_{i,j}(t)|$ are locally bounded on I. Then, $\Omega(t)$ is differentiable in norm with bounded derivative $\Omega'(t) := (\omega'_{i,j}(t))_{i,j\in\mathbb{N}}$ and $\|\Omega'(t)\|$ is locally bounded on I.*

Proof. We can write $\Omega(t) = \Omega_0(t) + \sum_{k=1}^{N}(V^k\Omega_k(t) + \Omega_{-k}(t)V^{*k})$, where $\Omega_k(t)$, $|k| \leqslant N$, are diagonal matrices and V is the right shift. Hence, if the statement is true for the matrices $\Omega_k(t)$, it is also true for $\Omega(t)$. So, we just have to check the proposition for a diagonal matrix $\Omega(t) = \operatorname{diag}(\omega_1(t), \omega_2(t), \ldots)$. If $\omega_n(t)$ are differentiable and $\sup_{n\in\mathbb{N}} |\omega'_n(t)|$ is locally bounded on I, $\Omega'(t) := \operatorname{diag}(\omega'_1(t), \omega'_2(t), \ldots)$ is bounded with $\|\Omega'(t)\| = \sup_{n\in\mathbb{N}} |\omega'_n(t)|$ locally bounded on I. If, besides, $\omega_n(t)$

are twice differentiable and $\sup_{n\in\mathbb{N}} |\omega_n''(t)| \leqslant K$ in a neighbourhood of t_0, using the mean value theorem we get

$$\left\| \frac{\Omega(t)-\Omega(t_0)}{t-t_0} - \Omega'(t_0) \right\| = \sup_{n\in\mathbb{N}} \left| \frac{\omega_n(t)-\omega_n(t_0)}{t-t_0} - \omega_n'(t_0) \right| \leqslant K|t-t_0|,$$

for t in such a neighbourhood. This proves the differentiability in norm. □

Now we can state the following result for a differentiable mono-parametric perturbation of the Schur parameters.

Proposition 5.2. *Let $a_n : I \to \mathbb{D}$ be differentiable for $n \in \mathbb{N}$. Then:*

1. *If $u : I \to \mathbb{T}$ is differentiable, the mass points of μ_N^t are differentiable functions $\lambda : I \to \mathbb{T}$ satisfying*
$$\lambda'(t) = \mu_N^t(\{\lambda(t)\}) X_N^t(\lambda(t))^{\mathrm{T}} \mathscr{C}_N'(t) \overline{X_N^t(\lambda(t))}.$$
2. *If $a_n : I \to \mathbb{D}$ is twice differentiable for $n \in \mathbb{N}$ and $\sup_{n\in\mathbb{N}} |a_n'(t)|$, $\sup_{n\in\mathbb{N}} |a_n''(t)|$, $\sup_{n\in\mathbb{N}} |\rho_n'(t)|$, $\sup_{n\in\mathbb{N}} |\rho_n''(t)|$ are locally bounded on I, for any isolated mass point λ_0 of μ^{t_0} there exists a differentiable function $\lambda : J \to \mathbb{T}$ on a neighbourhood J of t_0 such that $\lambda(t)$ is an isolated mass point of μ^t for $t \in J$ and $\lambda(t_0) = \lambda_0$. This function satisfies*
$$\lambda'(t) = \mu^t(\{\lambda(t)\}) X^t(\lambda(t))^{\mathrm{T}} \mathscr{C}'(t) \overline{X^t(\lambda(t))}.$$

Proof. From the previous discussions and Lemma 5.1 we find that the referred differentiable functions $\lambda(t)$ exist under the conditions of the theorem. The expression for $\lambda'(t)$ follows from the Hellmann–Feynman Theorem for normal operators. Let us consider the infinite case since the analysis of the finite case is analogous. The mass points $\lambda(t)$ are simple eigenvalues of $\mathscr{C}(t)$ with associated eigenspace spanned by $\overline{X^t(\lambda(t))}$. We know that there exists a strongly differentiable eigenvector $Y(t)$ of $\mathscr{C}(t)$ with respect to $\lambda(t)$. Therefore, just differentiating the equality $Y(t)^*\mathscr{C}(t)Y(t) = \lambda(t)Y(t)^*Y(t)$ and bearing in mind the unitarity of $\mathscr{C}(t)$, we get $Y(t)^*\mathscr{C}'(t)Y(t) = \lambda'(t)Y(t)^*Y(t)$. This relation is also true when substituting $Y(t)$ by $\overline{X^t(\lambda(t))}$ since they are proportional. The statement 2 is then a consequence of the equality $X^t(\lambda(t))^{\mathrm{T}}\overline{X^t(\lambda(t))} = \sum_{n\in\mathbb{N}} |\varphi_n(\lambda(t))|^2 = 1/\mu^t(\{\lambda(t)\})$. □

It is natural to expect a qualitatively different behaviour of the measure under rotations or dilatation of the Schur parameters. Therefore, it could be interesting to examine the preceding result when we decompose the mono-parametric perturbation in the way $a_n(t) = r_n(t)\mathrm{e}^{\mathrm{i}\alpha_n(t)}$, $r_n(t)$, $\alpha_n(t)$ being real functions.

Theorem 5.3. *Let $a_n(t) = r_n(t)\mathrm{e}^{\mathrm{i}\alpha_n(t)}$, where $r_n : I \to (-1, 1)$, $\alpha_n : I \to \mathbb{R}$ are differentiable for $n \in \mathbb{N}$. We define the functions*

$$\Gamma_n^t(z) := \frac{2}{\rho_n(t)^2}\operatorname{Im}\big(\mathrm{e}^{-\mathrm{i}\alpha_n(t)}z^{2-n}(\varphi_{n-1}^t(z))^2\big), \quad n \in \mathbb{N},$$

$$\Delta_n^t(z) := \left|\varphi_{n-1}^t(z)\right|^2 - \left|\varphi_n^t(z)\right|^2, \quad n \in \mathbb{N}.$$

1. *If* $u(t) = \mathrm{e}^{\mathrm{i}\beta(t)}$, *being* $\beta : I \to \mathbb{R}$ *differentiable, the mass points of* μ_N^t *have the form* $\lambda(t) = \mathrm{e}^{\mathrm{i}\theta(t)}$, *where* $\theta : I \to \mathbb{R}$ *is a differentiable function that satisfies*

$$\theta'(t) = \mu_N^t(\{\lambda(t)\})\left\{\sum_{n=1}^{N-1}(r_n'(t)\Gamma_n^t(\lambda(t)) + \alpha_n'(t)\Delta_n^t(\lambda(t))) + \beta'(t)|\varphi_{N-1}(\lambda(t))|^2\right\}.$$

2. *Let* $r_n : I \to (-1, 1)$, $\alpha_n : I \to \mathbb{R}$ *be twice differentiable. Assume that* $\sup_{n\in\mathbb{N}}|r_n'(t)|$, $\sup_{n\in\mathbb{N}}|r_n''(t)|$, $\sup_{n\in\mathbb{N}}|\alpha_n'(t)|$, $\sup_{n\in\mathbb{N}}|\alpha_n''(t)|$ *are locally bounded on* I *and there exists* $r < 1$ *such that* $|r_n(t)| < r$ *whenever* $r_n'(t) \neq 0$. *Then, if* λ_0 *is an isolated mass point of* μ^{t_0}, *there exists a differentiable function* $\lambda : J \to \mathbb{T}$ *on a neighbourhood* J *of* t_0 *such that* $\lambda(t)$ *is an isolated mass point of* μ^t *for* $t \in J$ *and* $\lambda(t_0) = \lambda_0$. $\lambda(t) = \mathrm{e}^{\mathrm{i}\theta(t)}$, *where* $\theta : I \to \mathbb{R}$ *is a differentiable function that satisfies*

$$\theta'(t) = \mu^t(\{\lambda(t)\})\sum_{n=1}^{\infty}\big(r_n'(t)\Gamma_n^t(\lambda(t)) + \alpha_n'(t)\Delta_n^t(\lambda(t))\big).$$

Remark 5.4. Notice that the above series converges due to the suppositions about the sequences $(r_n(t))_{n\in\mathbb{N}}$, $(\alpha_n(t))_{n\in\mathbb{N}}$ and the fact that $\mathbf{\Phi}^t(\lambda(t)) \in \ell^2$ since $\lambda(t)$ is a mass point of μ^t.

Proof. The conditions given for the sequences $(r_n(t))_{n\in\mathbb{N}}$, $(\alpha_n(t))_{n\in\mathbb{N}}$ and the function $\beta(t)$ are enough to apply Proposition 5.2. Therefore, the referred differentiable functions $\lambda(t)$ exist. Since $\mathbb{R}$ is the universal covering space of $\mathbb{T}$, with the imaginary exponential as a covering map, there exists a unique continuous real valued function $\theta(t)$ such that $\lambda(t) = \mathrm{e}^{\mathrm{i}\theta(t)}$, $\theta(t_0) = \operatorname{Arg}(\lambda(t_0))$. Moreover, the imaginary exponential is locally invertible with differentiable inverse, so, $\theta(t)$ must be differentiable too. From Proposition 5.2 we know that

$$\theta'(t) = \frac{\lambda'(t)}{\mathrm{i}\lambda(t)} = \frac{\mu^t(\{\lambda(t)\})}{\mathrm{i}\lambda(t)}X^t(\lambda(t))^{\mathrm{T}}\mathscr{C}'(t)\overline{X^t(\lambda(t))}.$$

The rest of the proof is just the calculation of the right hand side of the above expression, which we will do only for the infinite case since the arguments in finite case are similar. We can easily do this calculation using the factorization $\mathscr{C}(t) = \mathscr{C}_o(t)\mathscr{C}_e(t)$,

$\mathscr{C}_o(t) = C_o(\mathbf{a}(t)), \mathscr{C}_e(t) = C_e(\mathbf{a}(t))$, given by Remark 3.5. As a consequence of (2) and (4), $\mathscr{C}_e(t)\overline{X^t(\lambda(t))} = X^t(\lambda(t))$ and $\mathscr{C}_o(t)X^t(\lambda(t)) = \lambda(t)\overline{X^t(\lambda(t))}$. Therefore,

$$X^t(\lambda(t))^{\mathrm{T}}\mathscr{C}'(t)\overline{X^t(\lambda(t))} = X^t(\lambda(t))^{\mathrm{T}}\mathscr{C}'_o(t)X^t(\lambda(t)) + \lambda(t)X^t(\lambda(t))^*\mathscr{C}'_e(t)\overline{X^t(\lambda(t))},$$

which, using (4), gives

$$\theta'(t) = \mathrm{i}\mu^t(\{\lambda(t)\})\sum_{n=1}^{\infty}\lambda(t)^{-n}\left\{a'_n(t)\varphi^{t*}_{n-1}(\lambda(t))^2 - \overline{a}'_n(t)\varphi^t_n(\lambda(t))^2 - 2\rho'_n(t)\varphi^{t*}_{n-1}(\lambda(t))\varphi^t_n(\lambda(t))\right\}.$$

Finally, the expression given in the theorem for $\theta'(t)$ follows from the above one, taking into account (2) and the relations

$$a'_n(t) = r'_n(t)\mathrm{e}^{\mathrm{i}\alpha_n(t)} + \mathrm{i}\alpha'_n(t)a_n(t), \quad \rho'_n(t) = -\frac{r_n(t)}{\rho_n(t)}r'_n(t). \qquad \square$$

From the above theorem we directly get a bound for the angular velocity of the isolated mass points.

Corollary 5.5. *Under the conditions of Theorem 5.3*

$$|\theta'(t)| \leqslant \frac{2}{1-r^2}\sup_{n\geqslant 1}|r'_n(t)| + \sup_{n\geqslant 1}|\alpha'_n(t) - \alpha'_{n-1}(t)|,$$

where $\alpha_0 = 0$ and, in the case of μ^t_N, the sums run from 1 to N and $\alpha_N = \beta$.

The particular case of uniform rotations of the Schur parameters is specially interesting. It has been previously considered in [15] and by the authors in [4–6].

Corollary 5.6. *Let $\mathbf{a} \in \mathbb{D}^{\aleph_0}$, $u \in \mathbb{T}$ and $\alpha : I \to \mathbb{R}$ differentiable. If $a_n(t) = a_n\mathrm{e}^{\mathrm{i}\alpha(t)}$ for $n \in \mathbb{N}$ and $u(t) = u\mathrm{e}^{\mathrm{i}\alpha(t)}$. Then:*

1. *The differentiable arguments of the mass points of μ^t_N satisfy*

   $$\theta'(t) = \mu^t_N(\{\lambda(t)\})\alpha'(t).$$

2. *If $\alpha''(t)$ exists and is locally bounded on I, the differentiable arguments of the isolated mass points of μ^t satisfy*

   $$\theta'(t) = \mu^t(\{\lambda(t)\})\alpha'(t).$$

Proof. Apply Theorem 5.3 to $r_n(t) = |a_n|$, $\alpha_n(t) = \alpha(t) + \mathrm{Arg}(a_n)$ and $\beta(t) = \alpha(t) + \mathrm{Arg}(u)$. Notice that $\alpha'(t)$ is locally bounded on I if $\alpha(t)$ is twice differentiable. $\square$

This result is the generalization to arbitrary measures of the one founded in [5] for finitely supported measures using the Hessenberg representation. It says that under a uniform rotation of the Schur parameters the isolated mass points of the corresponding measure rotate in the same direction and the mass of each point gives its relative angular velocity with respect to the angular velocity of the Schur parameters. Therefore, a mass point rotates so much more quickly with the Schur parameters as its mass gets bigger. In fact, Theorem 5.3 suggests that, in general, the mass of an isolated mass point gives a measure of its instability under perturbations of the Schur parameters.

The study of the relation between Schur parameters and measures implies the attempt to find families of Schur parameters associated with measures with some common features. Theorem 5.3 opens a way to find mono-parametric families of Schur parameters whose measures have a common mass point. Among the ways to do this, we will just select some of them.

5.1. *Measures with a fixed mass point*

Let μ be the measure corresponding to a sequence $\mathbf{a} = (r_n e^{i\alpha_n})_{n\in\mathbb{N}}$ of Schur parameters and $(\varphi_n)_{n\geqslant 0}$ the associated OP. If $\lambda = e^{i\theta}$ is an isolated mass point of μ, our aim is to find mono-parametric perturbations $\mathbf{a}(t)$, $\mathbf{a}(t_0) = \mathbf{a}$, such that the corresponding measures μ^t have the same mass point, at least in a neighbourhood of t_0. We will also consider the analogous problem for the finitely supported measures μ_N associated with the parameters $(a_1, \dots, a_{N-1}, u)$, $u = e^{i\beta}$. In what follows we suppose that the perturbation satisfies the conditions given in Theorem 5.3.

Case 1. $a_k(t) = \begin{cases} a_k & \text{if } k \neq n \\ r(t)e^{i\alpha(t)} & \text{if } k = n \end{cases} \quad (r(t_0) = r_n, \alpha(t_0) = \alpha_n).$

This case corresponds to the perturbation of only the nth Schur parameter. So, the first n OP coincide with the unperturbed ones. Using (2) we get from Theorem 5.3 that $\lambda = e^{i\theta}$ is a fixed mass point if

$$\begin{aligned} &r'(t)\,\mathrm{Im}\big(e^{-i\alpha(t)}\lambda^{2-n}(\varphi_{n-1}(\lambda))^2\big) \\ &\quad = \alpha'(t)r(t)\big\{\mathrm{Re}(e^{-i\alpha(t)}\lambda^{2-n}(\varphi_{n-1}(\lambda))) + r(t)^2|\varphi_{n-1}(\lambda)|^2\big\}. \end{aligned}$$

If $(\varphi_{n-1}(\lambda))^2 = |\varphi_k(\lambda)|^2 e^{i\xi}$, the above equation becomes

$$d(r\sin(\alpha + (n-2)\theta - \xi)) + r^2 d\alpha = 0,$$

whose solution for the conditions $r(t_0) = r_n, \alpha(t_0) = \alpha_n$ is

$$r = \frac{\sin c}{\sin(\alpha + (n-2)\theta - \xi - c)},$$

$$c = \arctan\left(\frac{r_n \sin(\alpha_n + (n-2)\theta - \xi)}{1 + r_n\cos(\alpha_n + (n-2)\theta - \xi)}\right).$$

The same solution appears in the case of a finitely supported measure μ_N, $N > n$, if we leave the parameter u unperturbed.

Case 2. $a_k(t) = \begin{cases} a_k & \text{if } k < n \\ r(t)e^{i\alpha(t)} & \text{if } k = n \\ e^{i(\alpha(t)-\alpha_n)}a_k & \text{if } k > n \end{cases} \quad (r(t_0) = r_n, \alpha(t_0) = \alpha_n)$

Again, the first n OP coincide with the unperturbed ones. The condition given by Theorem 5.3 for a fixed mass point $\lambda = e^{i\theta}$ is now

$$\frac{2dr}{1-r^2}\sin(\alpha + (n-2)\theta - \xi) - d\alpha = 0,$$

where ξ is the phase of $(\varphi_{n-1}(\lambda))^2$. The solution for the conditions $r(t_0) = r_n$, $\alpha(t_0) = \alpha_n$ is

$$r = \frac{\sin\frac{1}{2}(\alpha + (n-2)\theta - \xi - c)}{\sin\frac{1}{2}(\alpha + (n-2)\theta - \xi + c)},$$

$$c = 2\arctan\left(\frac{1-r_n}{1+r_n}\tan\frac{1}{2}(\alpha_n + (n-2)\theta - \xi)\right).$$

This solution remains valid in the case of a measure μ_N, $N > n$, if we also include a perturbation $u(t) = e^{i(\alpha(t)-\alpha_n)}u$ of the parameter u.

Case 3. $a_k(t) = \begin{cases} a_k & \text{if } k < n \\ r(t)e^{i\alpha_n} & \text{if } k = n \\ e^{i\alpha(t)}a_k & \text{if } k > n \end{cases} \quad (r(t_0) = r_n, \alpha(t_0) = 0)$

As in the previous cases, the first n OP coincide with the unperturbed ones. From Theorem 5.3 and using (2) we find that the perturbations of this type with a fixed mass point $\lambda = e^{i\theta}$ are characterized by

$$2\sin(\alpha_n + (n-2)\theta - \xi)dr - (1 + 2r\cos(\alpha_n + (n-2)\theta - \xi) + r^2)d\alpha = 0,$$

where, again, ξ is the phase of $(\varphi_{n-1}(\lambda))^2$. The solution for the conditions $r(t_0) = r_n$, $\alpha(t_0) = 0$ is

$$r = -\frac{\sin\left(\frac{1}{2}\alpha + \alpha_n + (n-2)\theta - \xi - c\right)}{\sin(\frac{1}{2}\alpha - c)},$$

$$c = \arctan\left(\frac{\sin(\alpha_n + (n-2)\theta - \xi)}{r_n + \cos(\alpha_n + (n-2)\theta - \xi)}\right),$$

which is also valid in the case of a measure μ_N, $N > n$, if including a perturbation $u(t) = e^{i\alpha(t)}u$ of u.

If μ^{t_0} has an isolated point at $\lambda = e^{i\theta}$, the previous relations between r and α provide perturbations of the Schur parameters that give families of measures with the same mass point λ, at least for r, α in a neighbourhood of $r(t_0), \alpha(t_0)$. In the case of a finitely supported measure this neighbourhood is only restricted by the condition $|r| < 1$.

The simplest case of the above perturbations happens when $n = 1$, where always $\xi = 0$. Another particularly simple situation is the perturbation of a Geronimus measure, that corresponds to a constant sequence of Schur parameters.

Example: perturbations of Geronimus measures with a fixed mass point.

Let us consider the measure corresponding to a constant sequence of Schur parameters $a_n = a \in \mathbb{D}\backslash\{0\}$, $n \geqslant 1$ [10]. This measure has an isolated mass point at $\lambda = (1-a)/(1-\overline{a})$ if $|a-1/2| > 1/2$, that is, if $\mathrm{Re}(a) < |a|^2$. The related orthogonal polynomials are $\varphi_n(z) = \rho^{-n}(u_{n+1}(z) - (1-a)u_n(z))$, where $\rho = \sqrt{1-|a|^2}$, $u_n(z) = (w_1(z)^n - w_2(z)^n)/(w_1(z) - w_2(z))$ and $w_1(z)$, $w_2(z)$ are the solutions of $w^2 - (z+1)w + \rho^2 z = 0$ [11]. If $\lambda = \mathrm{e}^{\mathrm{i}\theta}$, the phase of $w_1(\lambda)$ and $w_2(\lambda)$ is $\theta/2$ and, thus, $(n-1)\theta/2$ is the phase of $u_n(\lambda)$. Hence, $\xi = (n-1)\theta$ and $\alpha + (n-2)\theta - \xi = \alpha - \theta$ for all n. In this case, the relations between r and α that give a fixed mass point for the three previous perturbations are independent of the index n of the Schur parameter where the perturbation starts.

Let us write $a = r_0 \mathrm{e}^{\mathrm{i}\alpha_0}$, $r_0, \alpha_0 \in \mathbb{R}$. Then, the condition for the existence of an isolated mass point is $\cos\alpha_0 < r_0$. Using the explicit form of the mass point we find that

$$\cos(\alpha_0 - \theta) = \frac{(1+r_0^2)\cos\alpha_0 - 2r_0}{1 + r_0^2 - 2r_0\cos\alpha_0}, \quad \sin(\alpha_0 - \theta) = \frac{(1-r_0^2)\sin\alpha_0}{1 + r_0^2 - 2r_0\cos\alpha_0}.$$

Taking into account these expressions we can find explicitly the relations between r and α that give a fixed mass point at $\lambda = (1-a)/(1-\overline{a})$ in the case of the three perturbations previously studied. We find the following results:

Case 1. $r = \frac{r_0 \sin\alpha_0}{\sin\alpha - r_0\sin(\alpha-\alpha_0)}$.

Case 2. $r = \frac{r_0 \sin\frac{1}{2}(3\alpha_0-\alpha) + \sin\frac{1}{2}(\alpha-\alpha_0)}{\sin\frac{1}{2}(\alpha_0+\alpha) - r_0\sin\frac{1}{2}(\alpha-\alpha_0)}$.

Case 3. $r = \frac{r_0 \sin\frac{1}{2}(2\alpha_0-\alpha) + \sin\frac{1}{2}\alpha}{\sin\frac{1}{2}(2\alpha_0-\alpha) + r_0\sin\frac{1}{2}\alpha}$.

Notice that in the first and second cases $r = r_0$ for $\alpha = \alpha_0$, due to the initial conditions $r(t_0) = r_0, \alpha(t_0) = \alpha_0$, while in the third case $r = r_0$ for $\alpha = 0$, since $r(t_0) = r_0, \alpha(t_0) = 0$. If $t_0 = 0$, we can choose $\alpha(t) = \alpha_0 + t$ in the first two cases and $\alpha(t) = t$ in the third one. Then, as a consequence of the previous results, we find that, in a neighbourhood of $t = 0$, the following families $\mathbf{a}(t)$ of Schur parameters are related to measures with a common mass point at $\lambda = (1-a)/(1-\overline{a})$ (we assume $a \in \mathbb{D}\backslash\mathbb{R}$ and $\mathrm{Re}(a) < |a|^2$):

Case 1. $a_k(t) = \begin{cases} a & \text{if } k \neq n, \\ \frac{\mathrm{Im}(a)}{\mathrm{Im}(a)\cos t - (|a|^2 - \mathrm{Re}(a))\sin t}\mathrm{e}^{\mathrm{i}t}a & \text{if } k = n. \end{cases}$

Case 2. $a_k(t) = \begin{cases} a & \text{if } k < n, \\ \frac{\mathrm{Im}(a)\cos\frac{t}{2} + (1-\mathrm{Re}(a))\sin\frac{t}{2}}{\mathrm{Im}(a)\cos\frac{t}{2} - (|a|^2-\mathrm{Re}(a))\sin\frac{t}{2}} \mathrm{e}^{\mathrm{i}t} a & \text{if } k = n, \\ \mathrm{e}^{\mathrm{i}t} a & \text{if } k > n. \end{cases}$

Case 3. $a_k(t) = \begin{cases} a & \text{if } k < n, \\ \frac{\mathrm{Im}(a)\cos\frac{t}{2} + (1-\mathrm{Re}(a))\sin\frac{t}{2}}{\mathrm{Im}(a)\cos\frac{t}{2} + (|a|^2-\mathrm{Re}(a))\sin\frac{t}{2}} a & \text{if } k = n, \\ \mathrm{e}^{\mathrm{i}t} a & \text{if } k > n. \end{cases}$

Acknowledgements

The work of the authors was supported by Project E-12/25 of DGA (Diputación General de Aragón) and by Ibercaja under grant IBE2002-CIEN-07.

References

[1] N.I. Akhiezer, M.G. Krein, Some questions in the theory of moments, Trans. Math. Mono., vol. 2, AMS, Providence, RI, 1962, Kharkov, 1938.
[2] M. Alfaro, El operador multiplicación en la teoría de polinomios ortogonales sobre la circunferencia unidad, in: Proceedings of the II Spanish-Portuguese Mathematical Conference, Madrid, 1973, Consejo Sup. Inv. Cient., Madrid, 1977, pp. 13–21.
[3] O. Bourget, J.S. Howland, A. Joye, Spectral analysis of unitary band matrices, Comm. Math. Phys. 234 (2003) 191–227.
[4] M.J. Cantero, Polinomios ortogonales sobre la circunferencia unidad. Modificaciones de los parámetros de Schur, Doctoral Dissertation, Universidad de Zaragoza, 1997.
[5] M.J. Cantero, L. Moral, L. Velázquez, Measures and para-orthogonal polynomials on the unit circle, East J. Approx. 8 (2002) 447–464.
[6] M.J. Cantero, L. Moral, L. Velázquez, Five-diagonal matrices and zeros of orthogonal polynomials on the unit circle, Linear Algebra Appl. 362 (2003) 29–56.
[7] P. Delsarte, Y. Genin, On the role of orthogonal polynomials on the unit circle in digital signal processing applications, in: P. Nevai (Ed.), Orthogonal Polynomials: Theory and Practice, NATO-ASI Series C, vol. 294, Kluwer, Dordrecht, 1990, pp. 115–133.
[8] R.P. Feynman, Forces in molecules, Phys. Rev. 56 (1939) 340–343.
[9] Ya.L. Geronimus, On polynomials orthogonal on the circle, on trigonometric moment problem, and on allied Carathéodory and Schur functions, Mat. Sb. 15 (1944) 99–130.
[10] Ya.L. Geronimus, Orthogonal Polynomials, Consultants Bureau, New York, 1961.
[11] L. Golinskii, Geronimus polynomials and weak convergence on a circular arc, Method. Appl. Analysis 6 (1999) 421–436.
[12] L. Golinskii, Singular measures on the unit circle and their reflection coefficients, J. Approx. Theory 103 (2000) 61–77.
[13] L. Golinskii, Operator theoretic approach to orthogonal polynomials on an arc of the unit circle, Matematicheskaya fizika, analiz, geometriya 7 (2000) 3–34.
[14] L. Golinskii, On the spectra of infinite Hessenberg and Jacobi matrices, Matematicheskaya fizika, analiz, geometriya 7 (2000) 284–298.
[15] L. Golinskii, P. Nevai, Szegö difference equations, transfer matrices and orthogonal polynomials on the unit circle, Commun. Math. Phys. 223 (2001) 223–436.

[16] L. Golinskii, P. Nevai, W. Van Assche, Perturbation of orthogonal polynomials on an arc of the unit circle, J. Approx. Theory 83 (1995) 392–422.
[17] W.B. Gragg, Positive definite Toeplitz matrices, the Arnoldi process for isometric operators, and Gaussian quadrature on the unit circle, J. Comput. Appl. Math. 46 (1993) 183–198; Numerical Methods of Linear Algebra, pp. 16–32, Moskov. Gos. Univ., Moskow, 1982.
[18] W.B. Gragg, The QR algorithm for unitary Hessenberg matrices, J. Comput. Appl. Math. 16 (1986) 1–8.
[19] W.B. Gragg, L. Reichel, A Divide and Conquer method for unitary and orthogonal eigenproblems, Numer. Math. 57 (1990) 695–718.
[20] H. Hellmann, Einführung in die Quantenchemie, Deuticke, Viena, 1937.
[21] E.K. Ifantis, A theorem concerning differentiability of eigenvectors and eigenvalues with some applications, Appl. Anal. 28 (1988) 257–283.
[22] E.K. Ifantis, Concavity and convexity of eigenvalues, Appl. Anal. 41 (1991) 209–220.
[23] M.E.H. Ismail, R. Zhang, On the Hellmann–Feynman theorem and the variation of zeros of certain special functions, Adv. Appl. Math. 9 (1988) 439–446.
[24] W.B. Jones, O. Njåstad, W.J. Thron, Moment theory, orthogonal polynomials, quadrature, and continued fractions associated with the unit circle, Bull. London Math. Soc. 21 (1989) 113–152.
[25] T. Kato, Perturbation of continuous spectra by trace class operators, Proc. Japan Acad. 33 (1957) 260–264.
[26] T. Kato, Perturbation Theory for Linear Operators, Springer, New York, 1966.
[27] F. Marcellán, E. Godoy, Orthogonal polynomials on the unit circle: distribution of zeros, J. Comput. Appl. Math. 37 (1991) 195–208.
[28] M. Reed, B. Simon, Methods of Modern Mathematical Physics, I. Functional Analysis, Academic Press, New York, 1972.
[29] M. Reed, B. Simon, Methods of Modern Mathematical Physics, IV. Analysis of Operators, Academic Press, New York, 1978.
[30] M. Rosenblum, Perturbation of the continuous spectrum and unitary equivalence, Pacific J. Math. 7 (1957) 997–1010.
[31] B. Simon, Orthogonal Polynomials on the Unit Circle, vols. 1 and 2, AMS Colloquim Series, American Mathematical Society, Providence, RI, 2005.
[32] M.H. Stone, Linear Transformation in Hilbert Space, AMS, Providence, RI, 1932.
[33] G. Szegö, Orthogonal Polynomials, AMS Colloq. Publ., vol. 23, 4th ed., AMS, Providence, RI, 1975.
[34] A.V. Teplyaev, The pure point spectrum of random polynomials orthogonal on the unit circle, Soviet Math. Dokl. 44 (1992) 407–411; Dokl. Akad. Nauk SSSR 320 (1991) 49–53.
[35] W.J. Thron, *L*-polynomials orthogonal on the unit circle, Nonlinear numerical methods and rational approximation (Wilrijk, 1987), pp. 271–278, Math. Appl., vol. 43, Reidel, Dordrecht, 1988.
[36] H. Weyl, Über gewöhnliche differentialgleichungen mit singularitäten und die zugehörigen entwicklungen willkürlicher funcktionen, Math. Ann. 68 (1910) 220–269.