Pré-Publicações do Departamento de Matemática
Universidade de Coimbra
Preprint Number 16–49

REFINED INTERLACING PROPERTIES FOR ZEROS OF PARAORTHOGONAL POLYNOMIALS ON THE UNIT CIRCLE

K. CASTILLO AND J. PETRONILHO

ABSTRACT: The purpose of this note is to extend in a simple and unified way the known results on interlacing of zeros of paraorthogonal polynomials on the unit circle. These polynomials can be regarded as the characteristic polynomials of any matrix similar to an unitary upper Hessenberg matrix with positive subdiagonal elements.

KEYWORDS: Paraorthogonal polynomials on the unit circle; zeros; unitary matrices; eigenvalues; interlacing; rank one perturbations.
MATH. SUBJECT CLASSIFICATION (2000): 15A42, 42C05.

1. Introduction and main result

The study of zeros of *orthogonal polynomials on the real line* (OPRL) can be regarded as an eigenvalue problem for Jacobi matrices[1]. This allows us to go back to one of the most important single books in the nineteenth century, *Cours d'analyse de l'École royale polytechnique* (1821) by Cauchy to deduce, at least in the weak sense, the zero interlacing property of consecutive OPRL from the simplest form of the nowadays called Cauchy interlacing theorem. The search of more refined eigenvalue interlacing properties of Jacobi matrices was probably initiated by Cauchy himself in his work *Sur l' Équation à l' Aide de Laquelle on Détermine les Inegalitées Séculaires des Mouvements des Planètes* (1829) and later continued by several authors, including in the second half of the last century Wilkinson [45], Kahan [29], Golub [20], Hill and Parlett [26], and Bar-On [6]. In the same spirit, this work

Received December 5, 2016.
KC is supported by the Portuguese Government through the Fundação para a Ciência e a Tecnologia (FCT) under the grant SFRH/BPD/101139/2014. This work is partially supported by the Centre for Mathematics of the University of Coimbra – UID/MAT/00324/2013, funded by the Portuguese Government through FCT/MCTES and co-funded by the European Regional Development Fund through the Partnership Agreement PT2020. JP is also partially supported by Dirección General de Investigación Científica y Técnica, Ministerio de Economía y Competitividad of Spain under the project MTM2015–65888–C4–4–P.

[1] A symmetric tridiagonal matrix whose next-to-diagonal elements are positive (cf. [27, p. 36]).

recovers one of the earliest approaches used to study zeros of *paraorthogonal polynomials on the unit circle* (POPUC), which is based on an eigenvalue problem for certain unitary matrices which bear many similarities with Jacobi matrices (cf. [31, 3, 23, 1, 25, 9, 16, 44, 7, 10, 11, 35, 36, 30, 39, 38, 40]).

Without wishing to delve into a historical discussion[2], as far as we know, the POPUC[3] were introduced (in a somewhat hidden form) and successfully developed in a serie of papers by Delsarte and Genin at the end of the 1980's [13, 15, 16], when they were working in signal processing. In [16], the authors focuses on the problem of computing the zeros of POPUC regarded as an eigenvalue problem for an unitary upper Hessenberg matrix with positive subdiagonal elements. Elegant and recent proofs of most interlacing properties of zeros of POPUC shared with OPRL are due to Simon [39] (cf. [40, Theorem 2.14.4]) where the theory of rank one perturbations plays a central role. However, before such work (and references therein) the zeros of POPUC were studied by the Linear Algebra community based on ideas close to those of Simon but supported on more elementary facts. Further analysis of these ideas will allow us to easily extend the known results. Indeed, our main purpose is to prove and improve, in connection with the works of Delsarte and Genin on the subject, the known zero interlacing properties of POPUC, based on the development of the ideas discussed by Arbenz and Golub in [4, Section 6][4].

Here and below, we mainly follow the notation of [35, 36, 40]. Denote by $\mathbb{D}$ the open unit disk and by $\mathbb{S}^1$ its boundary, i.e.,

$$\mathbb{D} := \{z \in \mathbb{C} : |z| < 1\} \, , \quad \mathbb{S}^1 := \{z \in \mathbb{C} : |z| = 1\} \, .$$

Let $(\alpha_0, \ldots, \alpha_{n-1}, \tau_n)$ with $\alpha_j \in \mathbb{D}$ $(j = 0, 1, \ldots, n-1)$ and $\tau_n \in \mathbb{S}^1$. Set

$$\Theta_j := \Theta(\alpha_j), \quad \Theta(\alpha) := \begin{pmatrix} \overline{\alpha} & \rho \\ \rho & -\alpha \end{pmatrix} \, , \quad \rho := \left(1 - |\alpha|^2\right)^{1/2} \, .$$

[2]The weakened orthogonality condition that POPUC satisfy appeared in [13, Equation 4.10] as far as we can tell. While it is true that in Geronimus' 1944 paper [17, Theorem IV] such polynomials were presented.

[3]In [13], Delsarte and Genin called to these polynomials (symmetric) *predictor polynomials* and its weakened orthogonality property *quasiorthogonality.* In [14], they refer to these polynomials as *quasiorthogonal polynomials on the unit circle.* This denomination could be also supported by the fact that in 1946 Geronimus regarding to these polynomials wrote that they *"...play the same role here as the quasi-orthogonal polynomials of M. Riesz in the Hamburger problem."* (cf. [18, Remark I]). The denomination POPUC was coined in [28].

[4]Such ideas were pioneering employed in the present context by Bohnhorst in her Ph.D. thesis [7] defended in 1993 at the Bielefeld University under the supervision of Elsner.

Define the $(n+1)$-by-$(n+1)$ matrix

$$\mathcal{C} := \mathcal{L}\,\mathcal{M}\,, \tag{1}$$

where $\mathcal{L}$ and $\mathcal{M}$ are given explicitly by

$$\mathcal{L} := \begin{cases} \Theta_0 \oplus \Theta_2 \oplus \cdots \oplus \Theta_{n-2} \oplus \overline{\tau}_n & \text{if } n \text{ is even} \\ \Theta_0 \oplus \Theta_2 \oplus \cdots \oplus \Theta_{n-1} & \text{if } n \text{ is odd}\,, \end{cases}$$

$$\mathcal{M} := \begin{cases} 1 \oplus \Theta_1 \oplus \Theta_3 \oplus \cdots \oplus \Theta_{n-1} & \text{if } n \text{ is even} \\ 1 \oplus \Theta_1 \oplus \Theta_3 \oplus \cdots \oplus \Theta_{n-2} \oplus \overline{\tau}_n & \text{if } n \text{ is odd.} \end{cases}$$

Any unitary $(n+1)$-by-$(n+1)$ upper Hessenberg matrix with positive subdiagonal elements is uniquely parameterized by $2n+1$ real numbers that compose the parameters of the array $(\alpha_0, \ldots, \alpha_{n-1}, \tau_n)$ [22] (cf. [24] and [2, Proposition 1]). The resulting matrix after this process is referred as the Schur parametric form of the original matrix. The factorization (1), which is unitarily similar to the Schur parametric form of an upper Hessenberg matrix with positive subdiagonal elements, was presented by Bunse-Gerstner and Elsner [9] (cf. [21, Section 12.2.10] and [7, Definition 3.3 and Lemma 3.4]). The explicit unitary pentadiagonal or double-staircase form of $\mathcal{C}$ (referred as *Doppel-Treppen-Matrix* in the original German source) was studied extensively by Bohnhorst [7], see Figure 1 for an 8-by-8 example (cf. [7, Equation 3.9] and [30, Figure 1.1]). The matrix $\mathcal{C}$ becomes a very popular object in the Mathematical Physics and Orthogonal Polynomials communities after the work [11], specially after Simon's monographs [35, 36] where it was called (improper) CMV matrix (cf. [38, 40]).

$$\begin{pmatrix}
\overline{\alpha}_0 & \rho_0\overline{\alpha}_1 & \rho_0\rho_1 & & & & & \\
\rho_0 & -\alpha_0\overline{\alpha}_1 & -\alpha_0\rho_1 & & & & & \\
 & \rho_1\overline{\alpha}_2 & -\alpha_1\overline{\alpha}_2 & \rho_2\overline{\alpha}_3 & \rho_2\rho_3 & & & \\
 & \rho_1\rho_2 & -\alpha_1\rho_2 & -\alpha_2\overline{\alpha}_3 & -\alpha_2\rho_3 & & & \\
 & & & \rho_3\overline{\alpha}_4 & -\alpha_3\overline{\alpha}_4 & \rho_4\overline{\alpha}_5 & \rho_4\rho_5 & \\
 & & & \rho_3\rho_4 & -\alpha_3\rho_4 & -\alpha_4\overline{\alpha}_5 & -\alpha_4\rho_5 & \\
 & & & & & \rho_5\overline{\alpha}_6 & -\alpha_5\overline{\alpha}_6 & \rho_6\overline{\tau}_7 \\
 & & & & & \rho_5\rho_6 & -\alpha_5\rho_6 & -\alpha_6\overline{\tau}_7
\end{pmatrix}$$

FIGURE 1. The matrix $\mathcal{C}$ for $n = 7$.

In order to make the notation more transparent, we write $\mathcal{C}(\alpha_0, \ldots, \alpha_{n-1}, \tau_n)$ instead of $\mathcal{C}$. We choose the representation (1) instead of their unitary similar upper Hessenberg matrix for a technical reason related to the manner in which Lemma 2.1 below is presented. In the next definition and subsequently, $\mathcal{I}$ denotes the identity matrix, whose order is made explicit or may be inferred from the context.

Definition 1.1 (cf. [39, Proposition 3.2]). *Let $\mathcal{C}(\alpha_0, \ldots, \alpha_{n-1}, \tau_n)$ be the matrix given by* (1)*, where $\alpha_j \in \mathbb{D}$ $(j = 0, 1, \ldots, n-1)$ and $\tau_n \in \mathbb{S}^1$. The (monic) polynomial P_{n+1} defined by*

$$P_{n+1}(z) := \det\big(z\mathcal{I} - \mathcal{C}(\alpha_0, \ldots, \alpha_{n-1}, \tau_n)\big)$$

is the POPUC of degree $n+1$ associated with the array $(\alpha_0, \ldots, \alpha_{n-1}, \tau_n)$.

It is not difficult to see that the eigenvalues of $\mathcal{C}(\alpha_0, \ldots, \alpha_{n-1}, \tau_n)$ are simple. This fact was observed in 1944 by Geronimus [17, Theorem IV] (cf. [18, Theorem III], [19, Theorem 9.1] and [5, Theorem 7.2.2]) using the connection between POPUC and OPUC. Note that if τ_n were in $\mathbb{D}$, then the corresponding characteristic polynomial would be an OPUC and their zeros would be in $\mathbb{D}$. A remarkable property of the eigenvectors of $\mathcal{C}(\alpha_0, \ldots, \alpha_{n-1}, \tau_n)$ is the fact that all their components are nonzero (cf. [35, Chapter 4] and references therein). This property is clearly valid also for the corresponding unitarily similar Hessenberg matrix.

Definition 1.2. *Two finite subsets $\{\zeta_1, \zeta_2, \ldots, \zeta_m\}$ and $\{\xi_1, \xi_2, \ldots, \xi_n\}$ $(1 \leq m \leq n)$ of $\mathbb{S}^1$ interlace (resp. strictly interlace) whenever there exist $n-m$ points $\zeta_{m+1}, \zeta_{m+2}, \ldots, \zeta_n \in \mathbb{S}^1$ such that any closed arc (resp. open arc) on $\mathbb{S}^1$ connecting two distinct elements of $\{\zeta_1, \zeta_2, \ldots, \zeta_n\}$ contains at last one element of $\{\xi_1, \xi_2, \ldots, \xi_n\}$, and vice versa.*

We can now formulate our main result.

Theorem 1.1. *Let $\mathcal{C}(\alpha_0, \ldots, \alpha_{n-1}, \tau_n)$ be a matrix given by* (1)*, where $\alpha_j \in \mathbb{D}$ $(j = 0, 1, \ldots, n-1)$ and $\tau_n \in \mathbb{S}^1$. The following sentences hold:*

(i) *Let $\beta \in \mathbb{S}^1 \backslash \{1\}$ and define $\mathcal{C}_m^\beta := \mathcal{C}(\alpha_0, \ldots, \alpha_{m-1}, \beta\alpha_m, \ldots, \beta\alpha_{n-1}, \beta\tau_n)$ $(0 \leq m < n)$ and $\mathcal{C}_n^\beta := \mathcal{C}(\alpha_0, \ldots, \alpha_{n-1}, \beta\tau_n)$. Then the eigenvalues of $\mathcal{C}(\alpha_0, \ldots, \alpha_{n-1}, \tau_n)$ and $\mathcal{C}_m^\beta$ strictly interlace on $\mathbb{S}^1$ for each $0 \leq m \leq n$.*

(ii) *For each* $0 \le m < n$, *let* $\tau_m \in \mathbb{S}^1$, *and let* $\mathcal{C}(\alpha_0, \ldots, \alpha_{n-1}, \tau_n)$ *be partitioned as*

$$\mathcal{C}(\alpha_0, \ldots, \alpha_{n-1}, \tau_n) = \begin{pmatrix} \mathcal{C}_{11} & \mathcal{C}_{12} \\ \mathcal{C}_{21} & \mathcal{C}_{22} \end{pmatrix}, \tag{2}$$

$\mathcal{C}_{11}$ *being the* $(m+1)$*-by-*$(m+1)$ *leading principal submatrix of* $\mathcal{C}(\alpha_0, \ldots, \alpha_{n-1}, \tau_n)$. *For each* $\zeta \in \mathbb{S}^1$, *define recursively the numbers*[5]

$$\tau_n(\zeta) := \tau_n, \quad \tau_j(\zeta) := \frac{\overline{\zeta}\ \alpha_j + \tau_{j+1}(\zeta)}{\overline{\alpha}_j \tau_{j+1}(\zeta) + \overline{\zeta}} \quad (j = n-1, \ldots, 1, 0). \tag{3}$$

Set[6]

$$A := C \cap \sigma(\mathcal{C}(\alpha_0, \ldots, \alpha_{m-1}, \tau_m)), \quad B := \sigma(\mathcal{N}) \setminus A,$$

where $C := \{\zeta \in \mathbb{S}^1 : \tau_m(\zeta) = \tau_m\}$, $\mathcal{N} := \mathcal{C}(\alpha_{m+1}, \ldots, \alpha_{n-1}, \tau_n)\,\mathcal{D}$ *with* $\mathcal{D} := \operatorname{diag}(\gamma_m, \mathcal{I})$, *and*

$$\gamma_m := \frac{\alpha_m - \tau_m}{\overline{\alpha}_m \tau_m - 1}. \tag{4}$$

Then $\mathcal{C}(\alpha_0, \ldots, \alpha_{n-1}, \tau_n)$ *and* $\mathcal{C}(\alpha_0, \ldots, \alpha_{m-1}, \tau_m)$ *have at most* $\min\{m+1, n-m\}$ *common eigenvalues. More precisely,* $\mathcal{C}(\alpha_0, \ldots, \alpha_{n-1}, \tau_n)$ *and* $\mathcal{C}(\alpha_0, \ldots, \alpha_{m-1}, \tau_m)$ *have* A *as the set of common eigenvalues,* A *being also given by the alternative expression*

$$A = \sigma(\mathcal{N}) \cap \sigma(\mathcal{C}(\alpha_0, \ldots, \alpha_{m-1}, \tau_m)).$$

Furthermore, the elements of the sets $\sigma\big(\mathcal{C}\ (\alpha_0, \ldots, \alpha_{n-1}, \tau_n)\big) \setminus A$ *and* $\sigma\big(\mathcal{C}(\alpha_0, \ldots, \alpha_{m-1}, \tau_m)\big) \cup B$ *strictly interlace on* $\mathbb{S}^1$.

Let P_{n+1} be the POPUC of degree $n+1$ associated to the array $(0, \ldots, 0, 1)$. Since $\mathcal{C}(0, \ldots, 0, 1)$ is a permutation matrix, it follows that $P_{n+1}(z) = z^{n+1} - 1$. The sequence $(P_j)_{j \ge 1}$ (all of whose zeros are roots of unity) produce, by geometric intuition, illuminating examples that fall within Theorem 1.1.

[5] In [15] (cf. [16, Equation 2.6]), Delsarte and Genin have shown that if the $\tau_j(\zeta)$'s (known as *pseudo reflection coefficients*), for a suitable choice of ζ, are given by (3), then the corresponding POPUC satisfy a three-term recurrence relation (cf. [12]). Bunse-Gerstner and He [10] have provided an illuminating discussion of the works of Delsarte and Genin on POPUC in matrix terms.

[6] $\sigma(\mathcal{A})$ denotes the spectrum of $\mathcal{A}$.

Example 1.1. *Let P_3 and P_6 be the POPUC associated to the arrays* $(0,0,1)$ *and* $(0,0,0,0,0,1)$*, respectively. In this situation,*

$$\mathcal{C}(0,0,1)=\begin{pmatrix}0&0&1\\1&0&0\\0&1&0\end{pmatrix},\quad \mathcal{C}(0,0,0,0,0,1)=\begin{pmatrix}0&0&1&0&0&0\\1&0&0&0&0&0\\0&0&0&0&1&0\\0&1&0&0&0&0\\0&0&0&0&0&1\\0&0&0&1&0&0\end{pmatrix},$$

and, therefore,

$$\sigma(\mathcal{C}(0,0,1))=\left\{1,e^{\pm i2\pi/3}\right\},\quad \sigma(\mathcal{C}(0,0,0,0,0,1))=\left\{\pm1,e^{\pm i2\pi/3},e^{\pm i\pi/3}\right\}.$$

In the notation of Theorem 1.1 we have $n=5$, $m=2$, $\tau_j(\zeta)=\zeta^{5-j}$ $(0\leq j\leq 5)$, $A=C=\sigma(\mathcal{C}(0,0,1))$, and $B=\emptyset$, where A is obtained by using any of the expressions outlined in Theorem 1.1. Clearly, $\mathcal{C}(0,0,0,0,0,1)$ and $\mathcal{C}(0,0,1)$ have A as the set of common eigenvalues and the elements of the sets $\sigma(\mathcal{C}(0,0,0,0,0,1))\setminus A$ and $\sigma(\mathcal{C}(0,0,1))$ strictly interlace on $\mathbb{S}^1$, in concordance with sentence (ii) of Theorem 1.1.

Regarding Theorem 1.1, as far as we know, sentence (i) for $m=0$ was proved by Ammar, Gragg and Reichel [1, Proposition 4.2], although the particular case $\beta=-1$ is known since Geronimus' work [17, Theorem IV] (cf. [18, Theorem III]). The sentence (i) for $m=n$ was proved by Bohnhorst in [7, Theorem 3.19] (cf. [8, Theorem 3.5]). In [39, Theorem 3.4], Simon proved a weaker version of sentence (ii) that reads as follows: *Strictly between any pair of eigenvalues of $\mathcal{C}(\alpha_0,\ldots,\alpha_{m-1},\tau_m)$ there is at least one eigenvalue of $\mathcal{C}(\alpha_0,\ldots,\alpha_{n-1},\tau_n)$.*

Corollary 1.1. *Let $\mathcal{C}(\alpha_0,\ldots,\alpha_{n-1},\tau_n)$ be a matrix given by* (1)*, where $\alpha_j\in\mathbb{D}$ $(j=0,1,\ldots,n-1)$ and $\tau_n\in\mathbb{S}^1$. Let $\tau_{n-1}\in\mathbb{S}^1$ and define γ_{n-1} as in* (4) *for $m=n-1$. Then $\mathcal{C}(\alpha_0,\ldots,\alpha_{n-1},\,\tau_n)$ and $\mathcal{C}(\alpha_0,\ldots,\alpha_{n-2},\tau_{n-1})$ have at most one common eigenvalue. More precisely, either $\mathcal{C}(\alpha_0,\ldots,\alpha_{n-1},\tau_n)$ and $\mathcal{C}(\alpha_0,\ldots,\alpha_{n-2},\tau_{n-1})$ have $\overline{\tau}_n\gamma_{n-1}$ as (only) common eigenvalue and the elements of $\sigma\big(\mathcal{C}(\alpha_0,\ldots,\alpha_{n-1},\tau_n)\big)\setminus\{\overline{\tau}_n\gamma_{n-1}\}$ and $\sigma\big(\mathcal{C}(\alpha_0,\ldots,\alpha_{n-2},\,\tau_{n-1})\big)$ strictly interlace on $\mathbb{S}^1$, or else $\mathcal{C}(\alpha_0,\ldots,\,\alpha_{n-1},\tau_n)$ and $\mathcal{C}(\alpha_0,\ldots,\alpha_{n-2},\tau_{n-1})$ have no common eigenvalues, and in such case $\overline{\tau}_n\gamma_{n-1}$ is not an eigenvalue of either, and the elements of the sets $\sigma\big(\mathcal{C}(\alpha_0,\,\ldots,\alpha_{n-1},\,\tau_n)\big)$ and $\sigma\big(\mathcal{C}(\alpha_0,\ldots,\alpha_{n-2},\tau_{n-1})\big)\cup\{\overline{\tau}_n\gamma_{n-1}\}$ strictly interlace on $\mathbb{S}^1$.*

Proof: Take $m = n-1$ in Theorem 1.1. Hence, (3) and (4) yield $C = \{\overline{\tau}_n\gamma_{n-1}\}$ which, in turn, is equal to $\sigma(\mathcal{N})$. Then either $A = C$ and $B = \emptyset$ if $\overline{\tau}_n\gamma_{n-1} \in \sigma\big(\mathcal{C}(\alpha_0,\ldots,\alpha_{n-2},\tau_{n-1})\big)$, or else $A = \emptyset$ and $B = C$ otherwise. The result follows immediately from sentence (iii) of Theorem 1.1. ■

Corollary 1.1 was proved by Bohnhorst [7, p. 48] (cf. [8, p. 819]) and rediscovered by Simon [39, Theorem 1.4]. It is worth noting that in view of Corollary 1.1 and besides the several and well-known practical consequences, POPUC answered the following open-ended question proposed by Turán as far as 1974 [42, Problem LXVI, p. 60]: *"It is known that the zeros of the nth orthogonal polynomial (with respect to a Lebesgue-integral function on an interval) separate the zeros of the $(n+1)$th polynomial. What corresponds to this fact on the unit circle?"*[7].

2. Proof of Theorem 1.1

2.1. Some preliminary lemmas. Theorem 1.1 will be proved through the following sequence of lemmas.

LEMMA 2.1. *Let $\mathcal{U}$ and $\mathcal{S}$ be unitary matrices of the same order and suppose that* rank $(\mathcal{I}-\mathcal{S}) = 1$ *. Then $\mathcal{U}$ and $\mathcal{US}$ have weakly interlacing eigenvalues on $\mathbb{S}^1$. Moreover, assume that $\mathcal{US}$ admits a decomposition $\mathcal{US} = \mathcal{U}_1 \oplus \mathcal{U}_2$, and let $\mathcal{U}$ be partitioned as*

$$\mathcal{U} = \begin{pmatrix} \mathcal{U}_{11} & \mathcal{U}_{12} \\ \mathcal{U}_{21} & \mathcal{U}_{22} \end{pmatrix},$$

$\mathcal{U}_{11}$ and $\mathcal{U}_1$ being of the same order. Set $U_1 := \sigma(\mathcal{U}_1)$, $U_2 := \sigma(\mathcal{U}_2)$, and $U := \sigma(\mathcal{U})$. Assume further that the eigenvalues of $\mathcal{U}_1$ and $\mathcal{U}_2$ are simple and $\sigma(\mathcal{U}_{11}) \cap U_1 = \sigma(\mathcal{U}_{22}) \cap U_2 = \emptyset$. Then the elements of the sets $U \setminus \big(U_1 \cap U_2\big)$ and $U_1 \cup \big(U_2 \setminus (U_1 \cap U_2)\big)$ strictly interlace on $\mathbb{S}^1$.

Proof: The first sentence of the lemma is the simplest form of a result due to Arbenz and Golub [4, Section 6] (cf. [7, Theorem 2.9] and [8, Theorem 2.7])[8]. In order to deduce the second one, we first claim that

$$U_1 \cap U_2 = U_1 \cap U = U_2 \cap U\,. \tag{5}$$

Indeed, since rank $(\mathcal{US}-\mathcal{U}) = 1$, there exist nonzero vectors $u, v \in \mathbb{C}^n$ (n being the common order of $\mathcal{U}$ and $\mathcal{S}$) such that $\mathcal{US} = \mathcal{U} + uv^T$. Using the

[7]We quote the English translation provided by Szüsz [43, Problem LXVI].

[8]It can be deduced directly using [32, p. 222] and [27, Corollary 4.3.9].

formula for the determinant of a rank one perturbation (cf. [34, Proposition 3.21]), we may write for each $\zeta \in \mathbb{C}$ [9]

$$\chi_{\mathcal{U}}(\zeta) = \chi_{\mathcal{US}}(\zeta) + v^T \operatorname{adj}(\zeta\mathcal{I} - \mathcal{US})u\,. \tag{6}$$

Let $\mathcal{US} = \mathcal{Z}\Lambda\mathcal{Z}^*$ be the spectral decomposition of $\mathcal{US}$ in which $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ and $\mathcal{Z} = (z_1 \ldots z_n)$. Thompson-McEnteggert's formula for the adjugate [41] (cf. [33, Theorem 2.1]) gives

$$\operatorname{adj}(\lambda_j\mathcal{I} - \mathcal{US}) = \chi'_{\mathcal{US}}(\lambda_j) z_j z_j^*\,, \tag{7}$$

where the prime denotes the derivative. Combining (6) with (7) yields[10]

$$\chi_{\mathcal{U}}(\lambda_j) = \Big(\chi'_{\mathcal{U}_1}(\lambda_j)\chi_{\mathcal{U}_2}(\lambda_j) + \chi_{\mathcal{U}_1}(\lambda_j)\chi'_{\mathcal{U}_2}(\lambda_j)\Big)\; z_j^* u v^T z_j\,. \tag{8}$$

We next claim that if $\lambda_j \in (U_1 - U_2) \cup (U_2 - U_1)$[11], then $z_j^* u v^T z_j \neq 0$. We only prove that $\lambda_j \in U_1 - U_2$ implies $v^T z_j \neq 0$. (To prove that $\lambda_j \in U_1 - U_2$ implies $z_j^* u \neq 0$, we proceed similarly, as well as for proving that $\lambda_j \in U_2 - U_1$ implies $z_j^* u v^T z_j \neq 0$.) Indeed, suppose that $\lambda_j \in U_1 - U_2$ and $v^T z_j = 0$. Since there is a normalized eigenvector v_j of $\mathcal{U}_1$ associated with λ_j such that $z_j = (v_j^T, 0, \ldots, 0)^T$, we deduce

$$\mathcal{U}_{11}\, v_j = \lambda_j v_j\,,$$

hence $\lambda_j \in \sigma(\mathcal{U}_{11}) \cap U_1$, contrary to $\sigma(\mathcal{U}_{11}) \cap U_1 = \emptyset$. Consequently, (5) follows from (8). Finally, it follows from (5) that the sets $U \setminus \big(U_1 \cap U_2\big)$ and $U_1 \cup \big(U_2 \backslash (U_1 \cap U_2)\big)$ have no common elements, thus the second sentence of the lemma follows from the first one. ■

LEMMA 2.2. *Let $\mathcal{U}$ be a unitary matrix and for a fixed k let $\mathcal{S}$ be the diagonal matrix obtained from the identity matrix by replacing the (k,k) entry with a number on $\mathbb{S}^1 \setminus \{1\}$. Assume that $\mathcal{U}$ and $\mathcal{S}$ have the same order. Assume further that the eigenvalues of $\mathcal{U}$ are simple and all its eigenvectors have a nonzero component at the position k. Then $\mathcal{U}$ and $\mathcal{US}$ have strictly interlacing eigenvalues on $\mathbb{S}^1$.*

Proof: Without loss of generality we can assume that $k = 1$, and so $\mathcal{S} = \operatorname{diag}(\beta, \mathcal{I})$ with $\beta \in \mathbb{S}^1 \setminus \{1\}$. Let $\mathcal{U} = \mathcal{Z}\Lambda\mathcal{Z}^*$ be the spectral decomposition

[9] $\chi_{\mathcal{A}}$ denotes the characteristic polynomial of $\mathcal{A}$.

[10] The eigenvalue interlacing already stated implies $U_1 \cap U_2 \subseteq U$, and so $U_1 \cap U_2 \subseteq U_1 \cap U$ and $U_1 \cap U_2 \subseteq U_2 \cap U$.

[11] Given a set E and $F, G \subseteq E$, we define $F - G := F \cap (E \backslash G)$; if $G \subseteq F$, then $F - G = F \backslash G$.

of $\mathcal{U}$ in which $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ and $\mathcal{Z} = (z_1 \ldots z_n)$. Arguing as in the proof of Lemma 2.1 we have

$$\chi_{\mathcal{US}}(\lambda_j) = \chi'_{\mathcal{U}}(\lambda_j)\, z_j^* u v^T z_j \,. \tag{9}$$

Let $a_j \neq 0$ be the first component of the vector z_j. Then

$$z_j^* u v^T z_j = z_j^* \mathcal{U}(\mathcal{I} - \mathcal{S}) z_j = \lambda_j (1-\beta)\, |a_j|^2 \neq 0 \,.$$

Thus the result follows from (9) and the first sentence of Lemma 2.1. ■

LEMMA 2.3. *Let $\alpha_j \in \mathbb{D}$ $(j = 0, 1, \ldots, n-1)$ and $\tau_n \in \mathbb{S}^1$. The following sentences hold:*

(i) *Let $\mathcal{S}$ be a diagonal matrix obtained from the $(n+1)$-by-$(n+1)$ identity matrix by replacing one of its diagonal entries with a number on $\mathbb{S}^1 \setminus \{1\}$. Then $\mathcal{C}(\alpha_0, \ldots, \alpha_{n-1}, \tau_n)$ and $\mathcal{C}(\alpha_0, \ldots, \alpha_{n-1}, \tau_n)\mathcal{S}$ have strictly interlacing eigenvalues on $\mathbb{S}^1$.*

(ii) *Let $\mathcal{C}(\alpha_0, \ldots, \alpha_{n-1}, \tau_n)$ be partitioned as in (2). Then, for each $0 \leq m < n$, $\mathcal{C}_{22}$ has no eigenvalues on $\mathbb{S}^1$.*

Proof: (i) The result follows directly from Lemma 2.2 and the fact that all the components of the eigenvectors of $\mathcal{C}(\alpha_0, \ldots, \alpha_{n-1}, \tau_n)$ are nonzero.

(ii) Assume that m is even. Note that $\mathcal{C}_{22}$ is the $(n-m)$-by-$(n-m)$ trailing principal submatrix of each of the matrices $\mathcal{C}(\alpha_m, \ldots, \alpha_{n-1}, \tau_n)$ and $\mathcal{C}(\alpha_m, \ldots, \alpha_{n-1}, \tau_n)\,\mathcal{S}$, where $\mathcal{S} := \operatorname{diag}(\beta, \mathcal{I})$. Suppose the assertion (ii) is false. Since $\mathcal{C}(\alpha_m, \ldots, \alpha_{n-1}, \tau_n)$ and $\mathcal{C}(\alpha_m, \ldots, \alpha_{n-1}, \tau_n)\,\mathcal{S}$ are unitary matrices, these matrices share all the eigenvalues of $\mathcal{C}_{22}$ on $\mathbb{S}^1$, which contradicts sentence (i). If m is odd, we argue in the same way noting that $\mathcal{C}_{22}^T$ is the $(n-m)$-by-$(n-m)$ trailing principal submatrix of each of the matrices $\mathcal{C}(\alpha_m, \ldots, \alpha_{n-1}, \tau_n)$ and $\mathcal{S}\,\mathcal{C}(\alpha_m, \ldots, \alpha_{n-1}, \tau_n)$. ■

LEMMA 2.4. *Let $\alpha_j \in \mathbb{D}$ $(j = 0, 1, \ldots, n-1)$ and $\tau_n \in \mathbb{S}^1$. Let $\mathcal{C}(\alpha_0, \ldots, \alpha_{n-1}, \tau_n)$ be partitioned as in (2), where $0 \leq m < n$. Let $\tau_m \in \mathbb{S}^1$ and define $\tau_m(\zeta)$ via (3) for each $\zeta \in \mathbb{S}^1$. Then $\mathcal{C}(\alpha_0, \ldots, \alpha_{n-1}, \tau_n)$ and $\mathcal{C}(\alpha_0, \ldots, \alpha_{m-1}, \tau_m)$ have at most $\min\{m+1, n-m\}$ common eigenvalues, which consist of the set of different solutions ζ of the equation $\tau_m(\zeta) = \tau_m$ on $\sigma(\mathcal{C}(\alpha_0, \ldots, \alpha_{m-1}, \tau_m))$.*

Proof: We begin by noting that

$$\det\big(\zeta\mathcal{I} - \mathcal{C}_n\big) = \det\big(\zeta\mathcal{I} - \mathcal{C}(\alpha_0, \ldots, \alpha_{m-1}, \tau_m(\zeta))\big)\det\big(\zeta\mathcal{I} - \mathcal{C}_{22}\big) \tag{10}$$

for each $\zeta \in \mathbb{S}^1$. Indeed, by sentence (ii) of Lemma 2.3, $\zeta\mathcal{I}-\mathcal{C}_{22}$ is nonsingular, hence (10) follows from the equality (cf. [7, Equation 3.41])

$$\mathcal{C}(\alpha_0,\ldots,\alpha_{m-1},\tau_m(\zeta)) = \mathcal{C}_{11} - \mathcal{C}_{12}(\mathcal{C}_{22}-\zeta\mathcal{I})^{-1}\mathcal{C}_{21}\,,$$

after taking into account that the Schur complement of $\zeta\mathcal{I}-\mathcal{C}_{22}$ in $\zeta\mathcal{I}-\mathcal{C}(\alpha_0,\ldots,\alpha_{n-1},\ \tau_n)$ is $\zeta\mathcal{I}-\big(\mathcal{C}_{11}-\mathcal{C}_{12}(\mathcal{C}_{22}-\zeta\mathcal{I})^{-1}\mathcal{C}_{21}\big)$. The result follows from (10) and the fact that for $\nu,\zeta\in\mathbb{S}^1$, with $\nu\neq\zeta$, $\mathcal{C}(\alpha_0,\ldots,\alpha_{m-1},\nu)$ and $\mathcal{C}(\alpha_0,\ldots,\alpha_{m-1},\zeta)$ have no common eigenvalues (see e.g. [40, Theorem 2.14.4]; alternatively, apply sentence (i) of Lemma 2.3). ∎

2.2. Proof of Theorem 1.1. (i) Let $\mathcal{S} := \operatorname{diag}(\mathcal{I}_m,\overline{\beta},\mathcal{I}_{n-m})$, $\mathcal{D} := \operatorname{diag}(\mathcal{I}_m,\ \mathcal{J}^{\beta}_{n-m+1})$, and $\mathcal{V} := \operatorname{diag}(\mathcal{I}_{m+1},\ \mathcal{J}^{\beta}_{n-m})$, where $\mathcal{J}^{\beta}_k := \operatorname{diag}(\beta,1,\beta,1,\ldots)$ is a k-by-k diagonal matrix. Note that

$$\begin{pmatrix}\overline{\beta} & 0\\ 0 & 1\end{pmatrix}\Theta(\alpha)\begin{pmatrix}1 & 0\\ 0 & \beta\end{pmatrix} = \Theta(\alpha\beta)\,. \tag{11}$$

Using (11) it is easily seen that [12]

$$\mathcal{D}^*\,\mathcal{C}(\alpha_0,\ldots,\alpha_{n-1},\tau_n)\,\mathcal{D}\,\mathcal{S} = \big(\mathcal{D}^*\,\mathcal{L}\,\mathcal{V}\big)\,\big(\mathcal{V}^*\,\mathcal{M}\,\mathcal{D}\,\mathcal{S}\big) = \mathcal{C}^{\beta}_m\,,$$

when m is even. Similarly, the transpose of (11) leads to

$$\mathcal{S}\,\mathcal{D}\,\mathcal{C}(\alpha_0,\ldots,\alpha_{n-1},\tau_n)\,\mathcal{D}^* = \big(\mathcal{S}\,\mathcal{D}\,\mathcal{L}\,\mathcal{V}^*\big)\,\big(\mathcal{V}\,\mathcal{M}\,\mathcal{D}^*\big) = \mathcal{C}^{\beta}_m\,,$$

when m is odd. The result follows from sentence (i) of Lemma 2.3.

(ii) Define the block diagonal matrix $S := \operatorname{diag}(\mathcal{I}_m,\mathcal{Z},\mathcal{I}_{n-m-1})$, where

$$\mathcal{Z} = \Theta^*_m\begin{pmatrix}\overline{\tau}_m & 0\\ 0 & \gamma_m\end{pmatrix}.$$

Hence

$$\mathcal{C}(\alpha_0,\ldots,\alpha_{m-1},\tau_m)\,\oplus\,\mathcal{N} = \mathcal{C}(\alpha_0,\ldots,\alpha_{n-1},\tau_n)\,\mathcal{S}\,,$$

when m is odd, and

$$\mathcal{C}(\alpha_0,\ldots,\alpha_{m-1},\tau_m)\,\oplus\,\mathcal{N}^T = \mathcal{S}^T\,\mathcal{C}(\alpha_0,\ldots,\alpha_{n-1},\tau_n)\,,$$

when m is even. Note that $\mathcal{N}$ has simple eigenvalues (on $\mathbb{S}^1$) by sentence (i) of Lemma 2.3. The result follows from Lemma 2.1, sentence (ii) of Lemma 2.3, and Lemma 2.4.

[12] A different proof is given in [37, Theorem 5.1].

Acknowledgment

The authors thank the Bielefeld University Library for kindly sending them a hard copy of Birgit Bohnhorst's Ph.D. Thesis.

References

[1] G. Ammar, W. Gragg, and L. Reichel, *Constructing a unitary Hessenberg matrix from spectral data*, in Numerical Linear Algebra, Digital Signal Processing and Parallel Algorithms (Leuven, 1988), vol. 70 of NATO Adv. Sci. Inst. Ser. F Compt. Systems Sci., Berlin, 1991, Springer, pp. 385–395.

[2] G. S. Ammar and W. B. Gragg, *Schur flow for orthogonal Hessenberg matrices*, in Hamiltonian and gradient flows, algorithms and control, vol. 3 of Fields Inst. Commun., Providence, RI, 1994, Amer. Math. Soc., pp. 27–34.

[3] G. S. Ammar, W. B. Gragg, and L. Reichel, *On the eigenproblem for orthogonal matrices*, in 25th IEEE Conference on Decision and Control, Athens, Greece, 1986, pp. 1963–1966.

[4] P. Arbenz and G. H. Golub, *On the spectral descomposition of Hermitian matrices modified by low rank perturbations with applications*, SIAM J. Matrix Anal. Appl., 9 (1988), pp. 40–58.

[5] F. V. Atkinson, *Discrete and continuous boundary problems*, vol. 8 of Mathematics in Science and Engineering, Academic Press, New York-London, 1964.

[6] I. Bar-On, *Interlacing properties of tridiagonal symmetric matrices with applications to parallel computing*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 548–562.

[7] B. Bohnhorst, *Beiträge zur numerischen Behandlung des unitären Eigenwertproblems*, PhD thesis, Fakultät für Mathematik, Universität Bielefeld, Bielefeld, Germany, 1993.

[8] B. Bohnhorst, A. Bunse-Gerstner, and H. Fassbender, *On the perturbation theory for unitary eigenvalue problems*, SIAM J. Matrix Anal. Appl., 21 (2000), pp. 809–824.

[9] A. Bunse-Gerstner and L. Elsner, *Schur parameter pencils for the solution of the unitary eigenproblem*, Linear Algebra Appl., 154/156 (1991), pp. 741–778.

[10] A. Bunse-Gerstner and C. He, *On a Sturm sequence of polynomials for unitary Hessenberg matrices*, SIAM J. Matrix Anal. Appl., 16 (1995), pp. 1043–1055.

[11] M. J. Cantero, L. Moral, and L. Velázquez, *Five-diagonal matrices and zeros of orthogonal polynomials on the unit circle*, Linear Algebra Appl., 362 (2003), pp. 29–56.

[12] K. Castillo, R. Cruz-Barroso, and F. Perdomo-Pío, *On a spectral theorem in paraorthogonality theory*, Pacific J. Math., 208 (2016), pp. 71–91.

[13] P. Delsarte and Y. Genin, *The tridiagonal approach to Szegő orthogonal polynomials, Toeplitz linear systems, and related interpolation problems*, SIAM J. Math. Anal., 19 (1988), pp. 718–735.

[14] ——, *On the role of orthogonal polynomials on the unit circle in digital signal processing applications*, in Orthogonal Polynomial: Theory and Practice (Columbos, OH, 1989), P. Nevai, ed., vol. 294 of NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci., Dordrecht, 1990, Kluwer Acad. Publ., pp. 115–133.

[15] ——, *Tridiagonal approach to the algebraic environment of Toeplitz matrices. I. Basic results*, SIAM J. Matrix Anal. Appl., 12 (1991), pp. 220–238.

[16] ——, *Tridiagonal approach to the algebraic environment of Toeplitz matrices. II. Zeros and eigenvalue problems*, SIAM J. Matrix Anal. Appl., 12 (1991), pp. 432–448.

[17] Y. L. Geronimus, *On polynomials orthogonal on the circle, on trigonometric moment-problem and on allied Carathéodory and Schur functions (in Russian)*, Rec. Math. [Mat. Sbornik] N.S., 57 (1944), pp. 99–130.

[18] Y. L. Geronimus, *On the trigonometric moment problem*, Ann. of Math., 47 (1946), pp. 742–761.

[19] Y. L. Geronimus, *Polynomials orthogonal on the unit circle and their applications*, in Series and Approximation, vol. 3 of Series One, Amer. Math. Soc., 1962, pp. 1–78.

[20] G. H. Golub, *Some uses of the Lanczos algorithm in numerical linear algebra*, Topics in numerical analysis (Proc. Roy. Irish Acad. Conf., Univ. Coll., Dublin, 1972), Academic Press, 1973, pp. 173–184.

[21] G. H. Golub and C. F. Van Loan, *Matrix Computation*, Johns Hopkins Studies in the Mathematical Sciences, The Johns Hopkins University Press, Baltimore, MD, fourth ed., 2013.

[22] W. B. Gragg, *Positive definite Toeplitz matrices, the Hessenberg process for isometric operators, and the Gauss quadrature on the unit circle (in Russian)*, Numerical Methods in Linear Algebra, Moskov. Gos. Univ., Moscow, 1982, pp. 16–32.

[23] ——, *The QR algorithm for unitary Hessenberg matrices*, J. Comp. Appl. Math., 16 (1986), pp. 1–8.

[24] ——, *Positive definite Toeplitz matrices, the Hessenberg process for isometric operators, and the Gauss quadrature on the unit circle*, J. Comp. Appl. Math., 46 (1993), pp. 183–198.

[25] W. B. Gragg and L. Reichel, *A divide and conquer method for unitary and orthogonal eigenproblems*, Numer. Math., 57 (1990), pp. 695–718.

[26] R. O. Hill Jr. and B. N. Parlett, *Refined interlacing properties*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 239–247.

[27] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, New York, second ed., 2013.

[28] W. B. Jones, O. Njåstad, and W. J. Thron, *Moment theory, orthogonal polynomials, quadrature, and continued fractions associated with the unit circle*, Bull. London Math. Soc., 21 (1989), pp. 113–152.

[29] W. Kahan, *Accurate eigenvalues of a symmetric tridiagonal matrix*, Tech. report CS41, Stanford University, Stanford, CA, 1966.

[30] R. Killip and I. Nenciu, *CMV: The unitary analogue of Jacobi matrices*, Comm. Pure Appl. Math., LX (2007), pp. 1148–1188.

[31] H. Kimura, *Generalized Schwarz form and lattice-ladder realizations of digital filters*, IEEE Trans. Circuits Systems, 32 (1985), pp. 1130–1139.

[32] C. Mehl, V. Mehrmann, A. C. M. Ran, and L. Rodman, *Eigenvalue perturbation theory of symplectic, orthogonal, and unitary matrices under generic structured rank one perturbations*, BIT Numer. Math., 54 (2014), pp. 219–255.

[33] D. S. Scott, *How to make the Lanczos algorithm converge slowly*, Math. Comp., 33 (1979), pp. 239–247.

[34] D. Serre, *Matrices: Theory and Applications*, Graduate Texts in Mathematics, Springer-Verlag, New York, second ed., 2010.

[35] B. Simon, *Orthogonal polynomials on the unit circle. Part I. Classical Theory*, vol. 54 of Amer. Math. Soc. Coll. Publ., Amer. Math. Soc., Providence, RI, 2005.

[36] ——, *Orthogonal polynomials on the unit circle. Part II. Spectral Theory*, vol. 54 of Amer. Math. Soc. Coll. Publ., Amer. Math. Soc., Providence, RI, 2005.

[37] ——, *Aizenman's theorem for orthogonal polynomials on the unit circle*, Constr. Approx., 23 (2006), pp. 229–240.

[38] ——, *CMV matrices: Five years after*, J. Comp. Appl. Math., 208 (2007), pp. 120–154.

[39] ——, *Rank one perturbations and the zeros of paraorthogonal polynomials on the unit circle*, J. Math. Anal. Appl., 329 (2007), pp. 376–382.

[40] ——, *Szegő's theorem and its descendants: Spectral theory for L^2 perturbations of orthogonal polynomials*, M. B. Porter Lectures, Princeton University Press, Princeton, 2011.

[41] R. C. Thompson and P. McEnteggert, *Principal submatrices. II: The upper and lower quadratic inequalities*, Linear Algebra Appl., 1 (1968), pp. 211–243.

[42] P. Turán, *Some open problems in approximation theory (in Hungarian)*, Mat. Lapok, 25 (1974), pp. 21–75.

[43] ——, *On some open problems of approximation theory. P. Turán memorial volume. Translated from the Hungarian by P. Szüsz*, J. Approx. Theory, 29 (1980), pp. 23–85.

[44] D. S. Watkins, *Some perspectives on the eigenvalue problem*, SIAM Rev., 35 (1993), pp. 430–471.

[45] J. H. Wilkinson, *The algebraic eigenvalue problem*, Monographs on Numerical Analysis, Oxford Science Publications. The Clarendon Press, Oxford University Press, New York, 1988.

K. Castillo
CMUC, Department of Mathematics, University of Coimbra, 3001-501 Coimbra, Portugal
E-mail address: kenier@mat.uc.pt

J. Petronilho
CMUC, Department of Mathematics, University of Coimbra, 3001-501 Coimbra, Portugal
E-mail address: josep@mat.uc.pt