

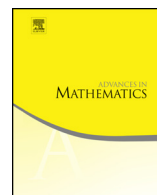


ELSEVIER

Contents lists available at ScienceDirect

Advances in Mathematics

www.elsevier.com/locate/aim



Poncelet's theorem, paraorthogonal polynomials and the numerical range of compressed multiplication operators

Andrei Martínez–Finkelshtein^{a,b,1}, Brian Simanek^a,
Barry Simon^{c,*,2}

^a Department of Mathematics, Baylor University, Waco, TX 76798, USA

^b Department of Mathematics, University of Almería, Almería 04120, Spain

^c Departments of Mathematics and Physics, Mathematics 253-37, California Institute of Technology, Pasadena, CA 91125, USA

ARTICLE INFO

ABSTRACT

Article history:

Received 31 October 2018

Received in revised form 27 March 2019

Accepted 7 April 2019

Available online xxxx

Communicated by D. Stroock

MSC:

47A12

42C05

34L05

Keywords:

Poncelet's theorem

Blaschke products

OPUC

There has been considerable recent literature connecting Poncelet's theorem to ellipses, Blaschke products and numerical ranges, summarized, for example, in the recent book [16]. We show how those results can be understood using ideas from the theory of orthogonal polynomials on the unit circle (OPUC) and, in turn, can provide new insights to the theory of OPUC.

© 2019 Elsevier Inc. All rights reserved.

* Corresponding author.

E-mail addresses: A_Martinez-Finkelshtein@baylor.edu (A. Martínez–Finkelshtein),

Brian_Simanek@baylor.edu (B. Simanek), bsimon@caltech.edu (B. Simon).

¹ Research supported in part by the Spanish Government and the European Regional Development Fund (grant MTM2017-89941-P), Junta de Andalucía (research group FQM-229 and Instituto Interuniversitario Carlos I de Física Teórica y Computacional), and by the University of Almería (Campus de Excelencia Internacional del Mar CEIMAR).

² Research supported in part by NSF grant DMS-1665526 and in part by Israeli BSF Grant No. 2014337.

POPUC
 Numerical range

1. Introduction

In 1813, Jean–Victor Poncelet [55] proved a remarkable theorem (see Halbeisen–Hunderbüler [44] for a simple proof) that says if K is an ellipse inside another ellipse, Q , so that there is a triangle with vertices in Q and sides tangent to K , then there are infinitely many such triangles, indeed, so many that their vertices fill Q and their tangent points fill K . There has been a huge literature motivated by this gem of projective geometry, even a recent book [22]. Our paper studies three different related developments.

Marden’s 1948 book [49, Ch. 1, §4], *Geometry of Polynomials*, popularized a theorem which he traces back to an 1864 paper of Jörg Siebeck [57].

Theorem A. *Let $\{w_j\}_{j=1}^p$ be the vertices of a convex polygon in \mathbb{C} ordered clockwise. Let $m_j \in \mathbb{R}$, and let*

$$M(z) = \sum_{j=1}^p \frac{m_j}{z - w_j} \tag{1.1}$$

Then the zero’s of M are the foci of a curve of class $p - 1$ which intersects each of the line segments $w_j w_k$; $j, k = 1, \dots, p$; $j \neq k$ at the point dividing the line in ratio m_j/m_k .

In this brief introduction, we are not going to try to give you the rather complicated definitions of the foci of a curve or of class nor will we use these notions later in this paper (but see [48,66]). We state this theorem to emphasize there is a n -gon version, that M and its zeros play a special role and that the ratios m_j/m_{j+1} occur.

The second set of results concern Blaschke products. Starting in 2002, Daepf, Gorkin and collaborators wrote a series of papers [14,15,17,36–38] considering finite Blaschke products³ of the form (for $\{z_j\}_{j=1}^n \subset \mathbb{D} := \{z \in \mathbb{C} : |z| < 1\}$, maybe not be all distinct)

$$B(z) = \prod_{j=1}^n \frac{z - z_j}{1 - \bar{z}_j z} \tag{1.2}$$

These are precisely the Schur functions (analytic maps of \mathbb{D} to itself) which are analytic in a neighborhood of $\overline{\mathbb{D}}$, of magnitude 1 on $\partial\mathbb{D}$, with n zeros (they actually consider zB and sometimes divide their basic function by z ; we prefer to take this B and sometime multiply it by z). Since $|zB(z)| < 1$ on \mathbb{D} and $|zB(z)| = 1$ on $\partial\mathbb{D}$, by the Cauchy–Riemann equations, the map $e^{i\theta} \mapsto e^{i\theta}B(e^{i\theta})$ is strictly increasing in θ and by the argument

³ When we want to specify n , we will refer to (1.2) as n fold Blaschke products.

principle is $n + 1$ to 1 so, for each $\lambda \in \partial\mathbb{D}$, there exist $n + 1$ solutions, w_j ; $j = 1, \dots, n + 1$, of $wB(w) = \bar{\lambda}$ (they take λ ; it will be clear later why we prefer $\bar{\lambda}$ as our label). We label the w_j with increasing arguments where arguments are taken in $[0, 2\pi)$.

The main result in this approach is

Theorem B. *For any $\{z_j\}_{j=1}^n \subset \mathbb{D}$ and any $\lambda \in \partial\mathbb{D}$, there exist $m_j(\lambda) > 0$ with $\sum_{j=1}^{n+1} m_j(\lambda) = 1$ so that*

$$\sum_{j=1}^{n+1} \frac{m_j(\lambda)}{z - w_j} = \frac{B(z)}{zB(z) - \bar{\lambda}} \tag{1.3}$$

The right side of this expression is a rational function of z which is $z^{-1} + O(|z|^{-2})$ at infinity and with poles exactly at the w_j so the left side is just a partial fraction expansion and $\sum_{j=1}^{n+1} m_j(\lambda) = 1$ follows from the asymptotics at infinity. The main issue is the proof of Daepf et al. [14] that $m_j > 0$ and they proved this by finding an explicit formula for the m_j in terms of the z 's and w 's,

$$m_j = \left[1 + \sum_{k=1}^{n+1} \frac{1 - |z_k|^2}{|w_j - z_k|^2} \right]^{-1} \tag{1.4}$$

It is left unmentioned that there is a probability measure and what its significance is. It has been noted (see, for example, footnote 5 on page 107 in [39]) that there is a converse of sorts to this result, that is, for every $\{m_j\}_{j=1}^{n+1}$ with $m_j > 0$ and $\sum_{j=1}^{n+1} m_j = 1$, there is a Blaschke product so that (1.3) holds.

The following theorem is natural to state in this $B(z)$ language

Theorem C. *Fix $\lambda \neq \mu$ both in $\partial\mathbb{D}$ and let $\{w_j\}_{j=1}^{n+1}$ (resp. $\{u_j\}_{j=1}^{n+1}$) be the solutions of $zB(z) = \bar{\lambda}$ (resp. $zB(z) = \bar{\mu}$). Then the w 's and u 's interlace. Conversely, if one is given such interlacing sets, there is a unique n fold Blaschke product so that the w 's and u 's are the solutions of a $zB(z)$ equation.*

This result was first proven by Gao and Wu [27] in the S_n framework below and the w 's and u 's enter as vertices of Poncelet $(n + 1)$ -gons. Their proof is long and involves lots of manipulations of determinants. The later, much shorter proof, of Daepf, Gorkin and Voss [17] constructs some rational Herglotz functions with given interlacing zeros and poles. We have a simple third proof. For reasons that will become obvious later, for now, we'll call this Wendroff's Theorem for Blaschke products. Parameter counting for this theorem is a little subtle. The set of w 's and u 's lie in a $2n + 2$ real dimensional manifold while it appears the equivalent set is only the n z_j 's in \mathbb{D} which is only $2n$ real parameters. But to get the w 's and u 's one needs two additional free parameters, namely λ and μ . Conversely these parameters are determined by the w 's and u 's since

$$\lambda = - \prod_{j=1}^{n+1} (-\bar{w}_j) \quad \mu = - \prod_{j=1}^{n+1} (-\bar{u}_j) \tag{1.5}$$

The final theme concerns a class of finite dimensional matrices, now called S_n , studied in a series of independent papers by Gau–Wu [24–30], also [76], and by Mirman [50–54] both series starting in 1998. Recall that an operator on a Hilbert space is called a *contraction* if its norm is at most 1. It is called *completely non-unitary* if it has no invariant subspace on which it is unitary, which in the finite dimensional case, is equivalent to there being no eigenvector with eigenvalue λ obeying $|\lambda| = 1$. In the finite dimensional case, the *defect index* of a contraction, A , is defined to be the dimension of the range of $\mathbf{1} - A^*A$. The space S_n is the set of completely non-unitary contractions on \mathbb{C}^n with defect index 1. One important theorem is

Theorem D. *For any $\{z_j\}_{j=1}^n \subset \mathbb{D}$ (maybe not different), there is an operator $A \in S_n$ whose eigenvalues (counting algebraic multiplicity) are the z_j . Any two elements in S_n are unitarily equivalent if and only if they have the same eigenvalues and multiplicities.*

Recall that if A is an operator on a Hilbert space, \mathcal{H} , then $N(A)$, the *numerical range* of an operator, A , on \mathcal{H} is the set of values $\langle \varphi, A\varphi \rangle$ where we run through all $\varphi \in \mathcal{H}$ with $\|\varphi\| = 1$ (not $\leq 1!$). It is a subtle fact that $N(A)$ is a convex subset of \mathbb{C} and an easy fact that it is compact when \mathcal{H} is finite dimensional. See [5,41,43] for more on numerical ranges. It is not hard to show that if A is normal, then (the closure of) $N(A)$ is the closed convex hull of the spectrum, so, in the finite dimensional normal case, $N(A)$ is the convex hull of the eigenvalues and so a convex polygon. In particular, if A is a k dimensional unitary operator with simple spectrum, then $N(A)$ is a convex k -gon inscribed in $\partial\mathbb{D}$.

Let $\mathcal{H} \subset \mathcal{K}$ be two Hilbert spaces and P the orthogonal projection from \mathcal{K} onto \mathcal{H} . If $A \in \mathcal{L}(\mathcal{H})$, $B \in \mathcal{L}(\mathcal{K})$, we are interested in the relation $A = PBP \upharpoonright \mathcal{H}$. If that holds we say that A is a *compression* of B and that B is a *dilation* of A . In case $\dim(\mathcal{K}) < \infty$, we call $\dim(\mathcal{K}) - \dim(\mathcal{H})$ the *rank* of the dilation. There is a huge literature on dilations, much of it involving families rather than single operators, for example, see [70,71]. Our usage limiting to a single operator and not demanding $A^n = PB^nP \upharpoonright \mathcal{H}$ (for $n = 1, \dots$) is more common; see, for example, the Wikipedia article [https://en.wikipedia.org/wiki/Dilation_\(operator_theory\)](https://en.wikipedia.org/wiki/Dilation_(operator_theory)).

Given a contraction, A on \mathcal{H} , one is interested in finding \mathcal{K} and $B \in \mathcal{L}(\mathcal{K})$ so that B is a unitary dilation of A . It is easy to construct such a dilation on $\mathcal{K} = \mathcal{H} \oplus \mathcal{H}$, so if $\dim(\mathcal{H}) = n$, a rank n unitary dilation, but one can show there is a one parameter family of rank one unitary dilations, $\{B_\lambda\}_{\lambda \in \partial\mathbb{D}}$ of any $A \in S_n$. For different λ , the eigenvalues are different and every point on the circle is an eigenvalue of exactly one B_λ .

The big theorem in these papers of Gau–Wu and Mirman is

Theorem E. *Let $A \in S_n$ and $\{B_\lambda\}_{\lambda \in \partial\mathbb{D}}$ its rank one unitary dilations. For each fixed λ , $N(B_\lambda)$ is a solid $n + 1$ -gon with vertices $\{w_j\}_{j=1}^{n+1}$ on $\partial\mathbb{D}$. Each edge of this polygon is*

tangent to $N(A)$ at a single point and as λ moves through all of $\partial\mathbb{D}$, these tangent points trace out the entire boundary of $N(A)$. Moreover

$$N(A) = \bigcap_{\lambda \in \partial\mathbb{D}} N(B_\lambda) \quad (1.6)$$

If, for a fixed λ , one forms

$$M(z) = \sum_{j=1}^{n+1} \frac{m_j}{z - w_j} \quad (1.7)$$

for suitable m_j , then the zeros of M are precisely the eigenvalues of A .

Remark. It was a conjecture of Halmos [42], proven by Choi–Li [13], that a formula like (1.6) holds for any contraction if the intersection is over all dilations of rank at most n .

We are not the first people to realize the relations between these themes; indeed, there is a very recent book on the subject [16] entitled: *Finding Ellipses: What Blaschke Products, Poncelet’s Theorem and the Numerical Range Know about Each Other*. The point of our paper is that while the authors of these works didn’t know it, they were discussing aspects of the theory of orthogonal polynomials on the unit circle (OPUC). Our realization of this additional connection will allow us to give new proofs and sometimes extend these results. These proofs are often quite simple (if one knows the OPUC background [60,61,59]!) and sometimes quite illuminating. For example, we’ll see that the left side of (1.3) is the matrix element of the resolvent of a unitary operator and the right side a Cramer’s rule ratio of determinants! Moreover, as we’ll explain, this formula, which is one of the main results of the 2002 paper of Daepf, Gorkin and Mortini [14], can be viewed as a special case of a general OPUC result of Khrushchev [46] published the year before!

We will also find that these lovely earlier ideas provide new results in the theory of OPUC of interest within that literature.

Because this paper is aimed at two disparate audiences (namely workers on the themes above and workers in OPUC) with rather different backgrounds, the presentation is more discursive than it might be if directed only at experts in a single area. In particular, Section 2 which officially sets notation and terminology presents a lot of results in OPUC with references to proofs. Section 3 states our main results. Some of these results will rely on a new theorem presented in Section 4 on GGT matrices. Sections 5–7 recover results on S_n and Sections 8–9 on Blaschke products. Section 10 discusses two new variants of Wendroff’s Theorem connected to Theorem C and the final three sections have additional remarks and results.

B.S. would like to thank Fritz Gesztesy and Lance Littlejohn for the invitation to visit Baylor where our collaboration was begun.

2. OPUC on one toe

A probability measure, $d\mu$ on $\partial\mathbb{D}$ is called *non-trivial* if it is not supported on a finite set of points. That is true if and only if $L^2(\partial\mathbb{D}, d\mu)$ is infinite dimensional and, in turn, is true if and only if $\{z^j\}_{j=0}^\infty$ are linearly independent as functions in that L^2 space. In that case, by using Gram–Schmidt on this set, we can define monic orthogonal polynomials, $\Phi_n(z; d\mu)$, and orthonormal polynomials, $\varphi_n(z; d\mu) = \Phi_n(z; d\mu)/\|\Phi_n\|$. Usually, we’ll drop the “ $d\mu$ ” unless it is needed for clarity.

The now standard reference for OPUC is Simon [60,61]; older references are parts of the books of Szegő [69], Geronimus [33] and Freud [21]. A summary of the high points is [59] which was named after a story from the Talmud. This even briefer summary is to set notation and terminology and emphasize those aspects of the theory we’ll need. We note that large swathes of the standard theory concern asymptotics of the Φ ’s and the relation of the Verblunsky coefficients to qualitative properties of the measure, none of which are relevant to our study here which concerns only $\{\Phi_n\}_{n=0}^N$ for fixed finite N . We note that there are issues we hope to study elsewhere of the large N behavior of the numerical range where these ideas may be important.

On $L^2(\partial\mathbb{D}, d\mu)$, one can define the antiunitary maps

$$\tau_n(f)(e^{i\theta}) = e^{in\theta} \overline{f(e^{i\theta})} \tag{2.1}$$

which takes z^k to z^{n-k} . Let \mathcal{P}_n be the $n + 1$ dimensional space of complex polynomials of degree at most n . Then τ_n maps \mathcal{P}_n to itself and has the form:

$$\tau_n(P_n(z)) = z^n P_n \left(\frac{1}{\bar{z}} \right) \tag{2.2}$$

so it just reverses the coefficients of P_n and complex conjugates them. We will follow the awful, but unfortunately universal, convention of usually writing P_n^* instead of $\tau_n(P_n)$ hoping the implicit n is clear so for example one writes

$$(z\Phi_n)^* = \Phi_n^*$$

where the $*$ on the left is τ_{n+1} while the one on the right is τ_n !

This map is important because of the second part of the elementary

Proposition 2.1. *In \mathcal{P}_n*

- (a) *Any f orthogonal to $\{z^j\}_{j=0}^{n-1}$ is a multiple of Φ_n*
- (b) *Any f orthogonal to $\{z^j\}_{j=1}^n$ is a multiple of Φ_n^**

Because we will need ideas in its proof later we’ll give the simple proof of the basic Szegő recursion

Theorem 2.2. Let $d\mu$ be a non-trivial probability measure on $\partial\mathbb{D}$. Then for each $n = 0, 1, \dots$ there exists $\alpha_n(d\mu) \in \mathbb{D}$ so that

$$\Phi_{n+1}(z) = z\Phi_n(z) - \bar{\alpha}_n\Phi_n^*(z) \tag{2.3}$$

Proof. Because they are monic, $\Phi_{n+1}(z) - z\Phi_n(z)$ is a polynomial of degree n which it is easy to see is orthogonal to $\{z^j\}_{j=1}^n$ and so a multiple of Φ_n^* which proves that (2.3) holds for some $\alpha_n \in \mathbb{C}$. By moving the second term on the right to the left and noting that the two terms now on the left are orthogonal one sees that $\|\Phi_{n+1}\|^2 + \|z\Phi_n\|^2 = |\alpha_n|^2\|\Phi_n^*\|^2$ so one has that

$$\|\Phi_{n+1}\| = \rho_n\|\Phi_n\|; \quad \rho_n \equiv \sqrt{1 - |\alpha_n|^2} \tag{2.4}$$

proving that $|\alpha_n| < 1$. \square

The α_n are called *Verblunsky coefficients*. The strange sign and complex conjugate in (2.3) are picked so that Theorem 2.6 below is true. Taking $z = 0$ in (2.3) and noting that Φ_n monic implies that $\Phi_n^*(0) = 1$, we see that

$$\alpha_n = -\overline{\Phi_{n+1}(0)} \tag{2.5}$$

(2.4) implies the important

$$\|\Phi_n\| = \rho_0 \dots \rho_{n-1} \tag{2.6}$$

Applying τ_{n+1} to (2.3) we get $\Phi_{n+1}^*(z) = \Phi_n^*(z) - \alpha_n z\Phi_n(z)$. This equation and (2.3) can be inverted to give what are called inverse Szegő recursion, which we'll write for the normalized OPUC:

$$z\varphi_n(z) = \rho_n^{-1}(\varphi_{n+1}(z) + \bar{\alpha}_n\varphi_{n+1}^*(z)) \tag{2.7}$$

$$\varphi_n^*(z) = \rho_n^{-1}(\varphi_{n+1}^*(z) + \alpha_n\varphi_{n+1}(z)) \tag{2.8}$$

An important basic fact about zeros of OPUC is the following (which the reader is right to suspect is connected to Theorem D!) is

Theorem 2.3 (*Wendroff's Theorem for OPUC*). All the zeros of $\Phi_n(z)$ lie in \mathbb{D} . Conversely, given any labelled set of n , not necessarily distinct, points in \mathbb{D} , there exists a measure so that those points are exactly the zeros (counting multiplicity) of the associated $\Phi_n(z)$. Any two such measures have the same $\{\alpha_j\}_{j=0}^{n-1}$ and so also the same $\{\varphi_j\}_{j=0}^n$.

Remarks. 1. The first statement goes back to Szegő's basic 1920-21 OPUC paper [68]. [60] has at least six proofs of this statement: [60, proofs of Theorem 1.7.1 containing equations (1.7.22), (1.7.43), (1.7.45), (1.7.46), (1.7.47) and (1.7.51)]. More generally,

Fejér [20] proved for any measure of compact support in \mathbb{C} whose support doesn't lie in a line that the zeros of all its OPs lie in the interior of the convex hull of the support.

2. The full result goes back to Geronimus [32] in 1946 long before the proofs of Theorem D.

3. The name comes from a theorem for orthogonal polynomials on the real line (OPRL) proven in Wendroff [74] in 1961: given $2n+1$ distinct points in \mathbb{R} thought of as an n point set interlacing an $n + 1$ point set, there is a probability measure on \mathbb{R} with all moments finite so that the two sets are the zeros of the OPRL P_n and P_{n+1} and all such measures have the same first $2n+1$ Jacobi parameters (i.e., coefficients of the three terms recurrence relation) and $\{p_j\}_{j=0}^{n+1}$, [60, Subsection 1.2.6]. The OPRL result without proof appears in a footnote of Geronimus' paper and [60] tried to push the name Geronimus–Wendroff Theorem for both results but the literature seems to have stuck with Wendroff which we'll follow in this paper for theorems of this type including new ones. Theorem C has the flavor of Wendroff's OPRL result and we'll see it can be viewed as an analog for paraorthogonal polynomials.

4. The uniqueness part, namely that Φ_n determines the $\{\alpha_j\}_{j=0}^{n-1}$ comes from (2.5) which determines α_{n-1} , and, then from inverse recursion for the Φ 's, we get Φ_{n-1} and so by iteration all the $\{\alpha_j\}_{j=0}^{n-1}$.

5. One way to show existence is a calculation [60, Proof of Theorem 1.7.5] that proves that if Q_n has all its zeros in \mathbb{D} , and if $d\mu(\theta) = c d\theta/|Q(e^{i\theta})|^2$ (where c is a normalization constant) then Q_n is orthogonal in $L^2(d\mu)$ to \mathcal{P}_{n-1} and so is the monic OPUC as required. This is the measure with Q_n as monic OPUC that has $\alpha_j = 0$ if $j \geq n$.

Two classes of functions (discussed in detail in [60, Section 1.3]) are *Carathéodory functions* (analytic function on \mathbb{D} which obey $\operatorname{Re}(F(z)) > 0$; $F(0) = 1$) and *Schur functions* (analytic functions on \mathbb{D} which obey $|f(z)| < 1$). Given a probability measure, $d\mu$, on $\partial\mathbb{D}$, we define two associated functions on \mathbb{D} :

$$F(z) = \int \frac{e^{i\theta} + z}{e^{i\theta} - z} d\mu(\theta); \quad F(z) = \frac{1 + zf(z)}{1 - zf(z)} \tag{2.9}$$

called the Carathéodory function and Schur function of $d\mu$.

Schur associated a set of parameters to any Schur function via $f_0 \equiv f$ and

$$\gamma_n(f) = f_n(0); \quad f_n(z) = \frac{\gamma_n + zf_{n+1}(z)}{1 - \bar{\gamma}_n f_{n+1}(z)} \tag{2.10}$$

If f is a finite degree m Blaschke product, then $\gamma_m \in \partial\mathbb{D}$ and the process terminates. If not (in which case we call f a *non-trivial Schur function*), we can define the *Schur iterates*, f_n , and *Schur parameters*, $\gamma_n(f) \in \mathbb{D}$, for all n .

Theorem 2.4 (*Schur's Theorem (1917)*). *There is a one-one correspondence between non-trivial Schur functions and sequences $\{\gamma_n\}_{n=1}^\infty$ in \mathbb{D} given by the map from f to its Schur parameters.*

This was proven by Schur [56]; see [60, Section 1.3.6].

Theorem 2.5 (*Verblunsky's Theorem (1935)*). *There is a one–one correspondence between non–trivial probability measures on $\partial\mathbb{D}$ and sequences $\{\alpha_n\}_{n=0}^\infty$ in \mathbb{D} given by the map from a measure to its OPUC and the Verblunsky coefficients defined via Szegő recursion.*

This was proven by Verblunsky [73] using an equivalent definition of his coefficients. [60] has four proofs of this result (see [60, Theorems 1.7.11, 3.1.3, 4.1.5 and 4.2.8]).

Theorem 2.6 (*Geronimus' Theorem (1944)*). *Let $d\mu$ be a non–trivial probability measure on $\partial\mathbb{D}$ and f its Schur function. Then*

$$\alpha_n(d\mu) = \gamma_n(f) \quad (2.11)$$

This theorem is due to Geronimus [31]. [60] has five proofs of this result (see [60, Theorems 3.1.4, 3.2.7, 3.2.10, 3.4.3 and 4.5.9]). This theorem explains why one writes Szegő recursion with the complex conjugate and minus sign on α .

Let $d\mu$ be a non–trivial probability measure with Verblunsky coefficients $\{\alpha_n\}_{n=0}^\infty$. The *second kind polynomials* for $d\mu$, written $\{\Psi_n\}_{n=0}^\infty$, are the OPUC for the measure whose Verblunsky coefficients are $\{-\alpha_n\}_{n=0}^\infty$. The following was given explicitly in Geronimus [31,32] and earlier implicitly in Verblunsky [73].

Theorem 2.7. *Let $d\mu_n$ be the measure with Verblunsky coefficients*

$$\alpha_j(d\mu_n) = \begin{cases} \alpha_j(d\mu), & \text{if } j \leq n-1 \\ 0, & \text{if } j \geq n \end{cases} \quad (2.12)$$

Then the Carathéodory function, F_n of $d\mu_n$ is Ψ_n^/Φ_n^* , and the Carathéodory function, F , of $d\mu$ is $\lim_{n \rightarrow \infty} F_n$ uniformly on compact subsets of \mathbb{D} .*

The proof (see [60, (3.2.21)]) depends on the formula

$$\Psi_k^*(z)\Phi_k(z) + \Phi_k^*(z)\Psi_k(z) = 2z^k \prod_{j=0}^{k-1} \rho_j^2 \quad (2.13)$$

In 2001, Khrushchev [46] found a new approach to Rahkmanoff's Theorem (which gives a sufficient condition for the Verblunsky coefficients to be asymptotically vanishing) with lots of other interesting stuff. A basic result he needed was the following

Theorem 2.8 (*Khrushchev's formula*). *Let f be the Schur function of some non–trivial probability measure, $d\mu$, on the unit circle and let f_n be its n th Schur iterate. Let $B_n(z) = \Phi_n(z)/\Phi_n^*(z)$. Then the Schur function of the probability measure $|\varphi_n(e^{i\theta})|^2 d\mu$ is $f_n(z)B_n(z)$.*

[60,61] have three proofs of this: Khrushchev’s original proof [61, Theorem 9.2.2], a proof using second kind polynomials [60, Corollary 4.4.2] and a proof using rank two perturbations of CMV matrices [60, Theorem 4.5.10].

This formula is an OPUC analog of the fact that the Green’s function for a whole line Schrödinger operator is the product of two suitably normalized Weyl solutions.

We let P_n be the projection in $L^2(d\mu)$ onto \mathcal{P}_{n-1} , the polynomials of degree at most $n - 1$. We use the subscript n because operators on \mathcal{P}_{n-1} are represented by $n \times n$ matrices. Given a non-trivial measure, let M_z be multiplication on by z on $L^2(\partial\mathbb{D}, d\mu)$. By a *compressed multiplication operator*, we mean the compression of the unitary M_z to polynomials of degree at most $n - 1$, i.e. $A = P_n M_z P_n$ restricted to \mathcal{P}_{n-1} . We’ll write A^μ when we want to be explicit about the measure.

Theorem 2.9. A^μ depends only on the Verblunsky coefficients $\{\alpha_j(d\mu)\}_{j=0}^{n-1}$. That is

$$A^\mu = A^\nu \iff \forall_{0 \leq j \leq n-1} \alpha_j(d\mu) = \alpha_j(d\nu) \tag{2.14}$$

Remarks. 1. We don’t merely mean unitarily equivalent. These are operators on an explicit space of polynomials and we mean equality. That said, we’ll see shortly that if we only know that A^μ is unitarily equivalent to A^ν , it is still true that $\forall_{0 \leq j \leq n-1} \alpha_j(d\mu) = \alpha_j(d\nu)$.

2. While this result is not unexpected, we sketch the proof since we don’t know any precise references.

Proof. If the relevant Verblunsky coefficients are equal, then they define the same set of $\{\varphi_j\}_{j=0}^{n-1}$ and $\{\varphi_j^*\}_{j=0}^{n-1}$ and the first is an orthonormal basis. Moreover, they both have

$$A\varphi_j = \begin{cases} \rho_j \varphi_{j+1} + \bar{\alpha}_j \varphi_j^* & \text{if } 0 \leq j \leq n - 2 \\ \bar{\alpha}_j \varphi_j^* & \text{if } j = n - 1 \end{cases} \tag{2.15}$$

and so they are equal.

Conversely, for any μ , we have that $Q_\mu((A^\mu)^{n-1}\mathbf{1}) = \Phi_{n-1}(\cdot; \mu)$ where Q_μ is the projection onto the orthogonal complement of $\{(A^\mu)^j \mathbf{1}\}_{j=1}^{n-2}$ so, by Wendroff’s Theorem for OPUC, equality of the A ’s implies equality of the α_j ’s for $0 \leq j \leq n - 2$. Then (2.15) for $j = n - 1$ implies the equality of the α_{n-1} ’s. \square

Fix z_0 . When does $(A^\mu - z_0)q = 0$ have a non-zero solution $q \in \mathcal{P}_{n-1}$? Clearly, only if $(z - z_0)q(z)$, which is a degree n polynomial, is killed by P_n , i.e. if $(z - z_0)q(z)$ is a multiple of $\Phi_n(z)$ and this happens if and only if $\Phi_n(z_0) = 0$. Thus the eigenvalues are precisely the zeros of Φ_n . A closer look shows that all the eigenvalues have geometric multiplicity 1 and algebraic multiplicity the order of the zero. Thus:

Theorem 2.10. *The eigenvalues of A are the zeros of Φ_n including up to algebraic multiplicity. Thus*

$$\det(z - A) = \Phi_n(z) \tag{2.16}$$

In particular, since for an $n \times n$ matrix, C , $\det(-C) = (-1)^n \det(C)$, we have that, by (2.5)

$$\det(A) = (-1)^{n+1} \bar{\alpha}_{n-1} \tag{2.17}$$

This implies that if A^μ is unitarily equivalent to A^ν , then μ and ν have the same Φ_n and so, by Wendroff’s Theorem, the same $\{\alpha_j\}_{j=0}^{n-1}$.

Suppose now that $d\mu$ is a trivial measure on $\partial\mathbb{D}$, say with $n + 1$ pure points, $\{w_j\}_{j=1}^{n+1}$. Then $\{z^k\}_{k=0}^n$ are still independent, so one can use Gram–Schmidt to form $\{\Phi_j\}_{j=0}^n$. As the norm minimizer, one also has that

$$\Phi_{n+1} = \prod_{j=1}^{n+1} (z - w_j) \tag{2.18}$$

Since this has norm 0, one expects (see (2.6)) and indeed finds that Φ_{n+1} is given by Szegő recursion but with $|\alpha_n| = 1$. That is, trivial measures are described by sets of $n + 1$ Verblunsky coefficients, the first n in \mathbb{D} and the last in $\partial\mathbb{D}$. The corresponding multiplication operators are precisely the $n + 1$ –dimensional unitaries with a cyclic vector. Moreover, the Schur and Geronimus Theorems extend. The Schur function of such an $n + 1$ point measure is, up to a leading phase factor, an n –fold Blaschke product (whose zeros are those of Φ_n) and which has $n + 1$ Schur parameters, n in \mathbb{D} and the last in $\partial\mathbb{D}$.

This motivates, the following: Suppose, we are given a non–trivial measure with Verblunsky coefficients, $\{\alpha_j\}_{j=0}^\infty$, and we consider, $\Phi_n(z)$. Given $\lambda \in \partial\mathbb{D}$, we define the *paraorthogonal polynomial* (POPUC) by

$$\Phi_{n+1}(z; \lambda) = z\Phi_n(z) - \bar{\lambda}\Phi_n^*(z) \tag{2.19}$$

The definition goes back to Delsarte–Genin [19] and Jones et al. [45]; among later papers, we mention [7,9–11,18,34,58,63,75]. One can show that the $n + 1$ point measure, $d\nu_\lambda$, whose first n Verblunsky coefficients are the first n α_j and with $\alpha_n = \lambda$ has $\Phi_{n+1}(z; \lambda)$ as its $n + 1$ –st monic OPUC (which has norm 0!). We’ll use U_λ as multiplication by z on $L^2(\partial\mathbb{D}, d\nu_\lambda)$. This L^2 space has dimension $n + 1$ and can be viewed as the space of polynomials of degree at most n . It will be useful to sometimes view it as an $n + 1 \times n + 1$ matrix in some convenient basis – most often the normalized OPUC but sometimes the eigenvectors.

Theorem 2.11. Fix $\{\alpha_j\}_{j=0}^{n-1}$ all in \mathbb{D} and let A be the corresponding compressed multiplication operator. The POPUC of degree $n + 1$ are in one to one correspondence with the rank one unitary dilations of A . The eigenvalues of the unitary, U_λ , associated to $\Phi_{n+1}(z; \lambda)$ are the zeros of that polynomial so that

$$\det(z - U_\lambda) = \Phi_{n+1}(z; \lambda) \tag{2.20}$$

In particular,

$$\det(U_\lambda) = (-1)^n \bar{\lambda} \tag{2.21}$$

We note that since $\Phi_n^*(0) = 1$, the constant term in $\Phi_{n+1}(z)$ is $-\bar{\lambda}$ so we have that

$$\lambda = (-1)^n \prod_{j=0}^{n+1} \bar{w}_j \tag{2.22}$$

and we note that

$$\Phi_{n+1}^*(z, \lambda) = -\lambda \Phi_{n+1}(z, \lambda) \tag{2.23}$$

We will need an explicit matrix representation for M_z and so A^μ which [60] calls the GGT representation after Geronimus [31], Gragg [40] and Teplyaev [72]. We note there is another matrix representation called the CMV representation (after [8]) discussed in [60, Section 4.2] which has advantages for the study of M_z on its infinite dimensional space but for our needs on the $n \times n$ operator, A^μ , the GGT representation is simpler and more natural.

$\{\varphi_j\}_{j=0}^{n-1}$ is an orthonormal basis for \mathcal{P}_{n-1} and the matrix elements of A^μ in that basis are

$$\mathcal{G}_{k\ell} = \langle \varphi_k, z\varphi_\ell \rangle; \quad k, \ell = 0, \dots, n-1 \tag{2.24}$$

Their explicit calculation is (see [60, (4.15)]) with $\alpha_{-1} = -1$

$$\mathcal{G}_{k\ell} = \begin{cases} -\bar{\alpha}_\ell \alpha_{k-1} \prod_{j=k}^{\ell-1} \rho_j & 0 \leq k \leq \ell \\ \rho_\ell & k = \ell + 1 \\ 0 & k \geq \ell + 2 \end{cases} \tag{2.25}$$

In [60], this is calculated using $\langle \Phi_n^*, P \rangle = \|\Phi_n\|^2 P(0)$ if $\deg P \leq n$. An easier alternative taken from [6] is to use Szegő recursion and inverse Szegő recursion in the form

$$z\varphi_n(z) = \rho_n \varphi_{n+1}(z) + \bar{\alpha}_n \varphi_n^*(z) \tag{2.26}$$

$$\varphi_j^*(z) = \rho_{j-1} \varphi_{j-1}^*(z) - \alpha_{j-1} \varphi_j(z) \tag{2.27}$$

so

$$\begin{aligned}
 z\varphi_n(z) &= \rho_n\varphi_{n+1}(z) - \bar{\alpha}_n\alpha_{n-1}\varphi_n(z) + \bar{\alpha}_n\rho_{n-1}\varphi_{n-1}^*(z) \\
 &= \rho_n\varphi_{n+1}(z) - \bar{\alpha}_n\alpha_{n-1}\varphi_n(z) - \bar{\alpha}_n\rho_{n-1}\alpha_{n-2}\varphi_{n-1}(z) \\
 &\quad + \bar{\alpha}_n\rho_{n-1}\rho_{n-2}\varphi_{n-2}^*(z)
 \end{aligned}
 \tag{2.28}$$

which upon iterating (using that $\varphi_0^* = 1 = \varphi_0$) yields

$$z\varphi_n(z) = \rho_n\varphi_{n+1}(z) + \sum_{k=0}^n \mathcal{G}_{kn}\varphi_k(z)
 \tag{2.29}$$

with \mathcal{G} given by (2.25). In other words,

$$\mathcal{G}(\{\alpha_j\}_{j=0}^{n-1}) = \begin{pmatrix} \bar{\alpha}_0 & \bar{\alpha}_1\rho_0 & \bar{\alpha}_2\rho_0\rho_1 & \dots & \bar{\alpha}_{n-1}\rho_0 \dots \rho_{n-2} \\ \rho_0 & -\bar{\alpha}_1\alpha_0 & -\bar{\alpha}_2\alpha_0\rho_1 & \dots & -\bar{\alpha}_{n-1}\alpha_0\rho_1 \dots \rho_{n-2} \\ 0 & \rho_1 & -\bar{\alpha}_2\alpha_1 & \dots & -\bar{\alpha}_{n-1}\alpha_1\rho_2 \dots \rho_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & -\bar{\alpha}_{n-1}\alpha_{n-2} \end{pmatrix}
 \tag{2.30}$$

We call the GGT representation of a compressed multiplication operator a *GGT contraction*. The GGT representation of the matrix associated to a POPUC we'll call a *GGT unitary*.

3. Main results

Here are the main results of this paper. We begin with three theorems that capture the main structure theorem for S_n (Theorem D).

Theorem 1. *Every compressed multiplication operator lies in S_n .*

Theorem 2. *Every element in S_n is unitarily equivalent to a compressed multiplication operator.*

Theorem 3. *For any set of n elements (with multiplicity) in \mathbb{D} , there is a compressed multiplication operator with those eigenvalues. Two compressed multiplication operators with the same characteristic polynomial are unitarily equivalent.*

Thus, we can parametrize equivalence classes of S_n by two sets of n elements from \mathbb{D} :

- (i) eigenvalues of a representative $A \in S_n$ of the equivalence class, or
- (ii) Verblunsky coefficients $\{\alpha_j\}_{j=0}^{n-1}$, associated to the compressed multiplication operator.

Theorem 1 will be easy and we'll see that given the first two theorems, Theorem 3 is essentially a restatement of Wendroff's Theorem. Thus the key will be Theorem 2. We will have two proofs, neither so short. The one in Section 5 will rely on turning a GGT unitary on its head, i.e. noting that a GGT unitary flipped along the secondary diagonal is again a GGT unitary but with different Verblunsky coefficients. We'll develop this idea in Section 4. Our second proof of Theorem 2 in Section 6 will involve constructing an orthonormal basis for the space on which $A \in S_n$ acts that will show it is acting as a compressed multiplication operator. Our construction of this basis is motivated by inverse Szegő recursion.

Arlinskii has four papers [1–4], one with coauthors, that are related to the above three theorems. These papers deal with a related, but different, class of operators which he calls truncated CMV matrices. He gets these by starting with a finite or infinite unitary CMV matrix and stripping off their first row and column. If the original unitary is a finite $(n + 1) \times (n + 1)$ matrix, this is related to, but distinct from, our compressed multiplication operators in CMV basis. It is cleaner if one describes the difference in GGT basis. Then he gets his operators by removing the first row and column and we remove the last row and column. For his objects, he has analogs of Theorems 1 and 2. He also has a theorem like Theorem 3 except that, because truncated CMV matrices whose Verblunsky coefficients are related up to an overall phase factor are unitarily equivalent, the uniqueness result is a little more involved. Because of our Theorem 4.2, one can establish a simple unitary equivalence between truncated CMV matrices and compressed multiplication operators and then obtain our Theorems 1–3 from his theorems or vice-versa. We note that although he doesn't seem to know about the earlier work around Theorem D, like us, he has an OPUC proof of it.

Next, in Section 7, we'll turn to the study of the numerical range of compressed multiplication operators and prove two theorems.

For each $\lambda \in \partial\mathbb{D}$, let U_λ be the associated unitary, Φ_{n+1} the associated POPUC, $\{w_j\}_{j=1}^{n+1}$ the zeros of Φ_{n+1} , cyclically ordered, η_j the associated normalized eigenvectors, so $\eta_j(z) = N_j^{-1}\Phi_{n+1}(z)/(z - w_j)$ and $d\mu$ the spectral measure

$$d\mu = \sum_{j=1}^{n+1} |\langle \eta_j, 1 \rangle|^2 \delta_{w_j} \tag{3.1}$$

Let $m_j = |\langle \eta_j, \varphi_n \rangle|^2 > 0$ (since $\deg(\eta_j) = n$) so $\sum_{j=1}^{n+1} m_j = 1$ (since $\|\varphi_n\| = 1$). Let A be the dimension n compressed multiplication operator and $w_{n+2} \equiv w_1$.

Theorem 4. *For $j = 1, \dots, n + 1$, the line from w_j to w_{j+1} intersects $N(A)$ in a single point, ζ_j , and $|\zeta_j - w_j|/|\zeta_j - w_{j+1}| = m_j/m_{j+1}$. In particular, $\prod_{j=1}^{n+1} |\zeta_j - w_j| = \prod_{j=1}^{n+1} |\zeta_j - w_{j+1}|$.*

Remark. If the line is between w_j and w_k , the corresponding point also lies in $N(A)$ but it is not, in general, the only point in $N(A)$.

Theorem 5. For each λ , we have that $N(U_\lambda)$ is a solid $(n+1)$ -gon whose sides are tangent to $N(A)$. $\partial N(A)$ is a strictly convex analytic curve and one has that

$$N(A) = \cap_{\lambda \in \partial \mathbb{D}} N(U_\lambda) \tag{3.2}$$

The proofs, which rely on operator theory, will not be difficult.

With the above definition of the m_j , the spectral measure for φ_n and the operator U_λ is

$$d\nu = \sum_{j=1}^{n+1} m_j \delta_{w_j} \tag{3.3}$$

Theorem 6. Fix a degree $n + 1$ POPUC and let $d\nu$ be the measure constructed above. Then one has that

$$\int \frac{1}{z - e^{i\theta}} d\nu(\theta) = \frac{\Phi_n(z)}{z\Phi_n(z) - \bar{\lambda}\Phi_n^*(z)} = \frac{\Phi_n(z)}{\Phi_{n+1}(z, \lambda)} \tag{3.4}$$

Conversely, let $d\nu$ be an $n + 1$ point probability measure on $\partial \mathbb{D}$. Then (3.4) holds for the POPUC defined by $d\nu$.

If $\{z_j\}_{j=1}^n$ are the zeros of Φ_n , then $\Phi_n(z)/\Phi_n^*(z) = \prod_{j=1}^n \frac{z-z_j}{1-\bar{z}_jz}$, so Theorem 6 is the Blaschke product theorem of Daep et al. (1.3) and its converse as discussed after (1.4) with a very different proof of the positivity of the m_j and of the formula. As we'll discuss this result can be proven from Khrushchev's formula by taking limits to extend his formula to trivial measures. It can also be obtained from general formulae for M -functions in [60]. We will give two simple direct proofs. From the OPUC point of view, the picture isn't complete until we find the Verblunsky coefficients for the trivial measure $d\nu$.

Theorem 7. Let $\{\alpha_j\}_{j=0}^{n-1}$ be the Verblunsky coefficients for the spectral measure, $d\mu$, of (3.1), that is the measure for which φ_j are the OPUC. Then the Verblunsky coefficients $\{\alpha_j(d\nu)\}_{j=0}^{n-1}$ for the measure, $d\nu$ of (3.3) are given by

$$\alpha_j(d\nu) = -\lambda \bar{\alpha}_{n-1-j}, \quad j = 0, \dots, n - 1; \quad \alpha_n(d\nu) = \lambda \tag{3.5}$$

The $\lambda = 1$ case of this result is implicit in a remark on rank two decoupling of CMV matrices in [60] but we'll give two more direct proofs: one using the results from Section 4 on turning a CMV matrix on its head and one proof using Geronimus' Theorem and Szegő recursion. Theorems 6 and 7 will be proven in Sections 8 and 9.

Section 10 has two Wendroff type theorems (Section 10 will discuss previous literature related to this result but we note here that Theorem 8 appears previously in [35]).

Theorem 8 (Wendroff’s Theorem for (P)OPUC). *The zero’s of POPUC’s for two values of λ interlace. Conversely, given two sets of $n + 1$ interlacing points on $\partial\mathbb{D}$, there exist unique $\{\alpha_j\}_{j=0}^{n-1}$ in \mathbb{D} and λ, μ in $\partial\mathbb{D}$ so these are zeros of the associated POPUCs.*

We put the (P) in “(P)OPUC” because, as we’ll discuss, there is a rather different Wendroff type theorem for POPUC in the literature which only involves some POPUC parameters whereas we discuss both the POPUC parameter and the OPUC parameters.

Theorem 9 (Wendroff’s Theorem for Second Kind POPUC). *Let $\{w_j\}_{j=1}^{n+1}$ be the zeros of a POPUC, Φ_{n+1} , ordered clockwise, and $\{y_j\}_{j=1}^{n+1}$ be the zeros of associated second kind POPUC, Ψ_{n+1} , also ordered clockwise, so that y_1 is the first zero after w_1 going clockwise. Then the w ’s and y ’s strictly interlace and one has that*

$$\prod_{j=1}^{n+1} y_j = - \prod_{j=1}^{n+1} w_j \tag{3.6}$$

Conversely, if $\{w_j\}_{j=1}^{n+1}$ and $\{y_j\}_{j=1}^{n+1}$ are strictly interlacing and obey (3.6), then there is a unique set of Verblunsky coefficients $\alpha_0, \dots, \alpha_{n-1} \in \mathbb{D}$ and $\lambda \in \partial\mathbb{D}$ so that $\{w_j\}_{j=1}^{n+1}$ is the set of zeros of the associated POPUC and $\{y_j\}_{j=1}^{n+1}$ the zeros of the associated second kind POPUC.

Recall that for OPUC, second kind OPUC were defined by putting a minus sign in front of all the Verblunsky coefficients. We define the second kind POPUC by also putting a minus sign in front of the λ parameter unlike the convention in [63]. Theorem 8 is equivalent to the result of Gau–Wu quoted as Theorem C. We have a new proof. The first halves of both of our Wendroff type theorems are already in the OPUC literature; we’ll say more about that in Section 10.

Section 11 discusses a theorem of Gau–Wu [25] about zeros of derivatives of polynomials all of whose zeros are on $\partial\mathbb{D}$ and the question of Gorkin–Skubak [37] asking when a finite subset of \mathbb{D} is the set of zeros of the derivative of a polynomial all of whose zeros lie on $\partial\mathbb{D}$.

Section 12 discusses Poncelet’s Theorem. Theorem A shows that the eigenvalues of $A \in S_n$ are the foci of a curve containing set $\partial N(A)$. While we don’t have an OPUC way to understand the eigenvalues as foci, Section 13 explains two other ways of going from $N(A)$ to the eigenvalues.

4. Turning a unitary GGT matrix on its head

In this section we answer a question about unitary $n \times n$ GGT matrices that will be relevant to one of our proofs of Theorem 2 and 7 (although the later will need $(n + 1) \times (n + 1)$ matrices). We begin with

Theorem 4.1. *Let U be an $n \times n$ unitary matrix with cyclic vector ψ_0 . Then there is a unique basis $\{e_j\}_{j=1}^n$ in which U is a GGT unitary $\mathcal{G}(\{\alpha_j\}_{j=0}^{n-1})$ and $\psi_0 = e_1$. The $\{\alpha_j\}_{j=0}^{n-1}$ are uniquely determined by the pair U, ψ_0 .*

Remarks. 1. We emphasize that the $\{\alpha_j\}_{j=0}^{n-1}$ are not unitary invariants of U alone but of U and its cyclic vector. The unitary invariants of U are the n eigenvalues so n real parameters, but the α 's are $n - 1$ in \mathbb{D} and one in $\partial\mathbb{D}$ (since $\alpha_{n-1} \in \partial\mathbb{D}$) and so $2n - 1$ real parameters. To determine \mathcal{G} , all that matters from the cyclic vector is the spectral measure, so $n - 1$ independent real weights leading to the correct parameter count. That the eigenvalues of U aren't enough to determine the α 's is indicated by the form of the Wendroff type theorems in Section 10.

2. While we won't need it below, there are explicit formulae for the weights of the spectral measure $d\mu = \sum_{j=1}^n q_j \delta_{w_j}$ for U and ψ_0 in terms of the OPUC $\{\varphi_j\}_{j=0}^{n-1}$ (which are determined by $\{\alpha_j\}_{j=0}^{n-2}$). Namely $q_j = 1/\lambda_{n-1}(w_j)$; $\lambda_{n-1}(w) = \sum_{j=0}^{n-1} |\varphi_j(w)|^2$. The λ are called Christoffel numbers. This formula goes back to Jones et al. [45] (see [60, Theorem 2.2.12] for a quick proof) although it is just the POPUC analog of well-known formulae for OPRL known under the name Gauss quadrature or Gauss–Jacobi quadrature.

Proof. The pair U, ψ_0 determine the spectral measure and so a set of $n + 1$ monic OPUC whose last element is a POPUC and whose first n elements normalized are the basis obtained by using Gram–Schmidt on $\{U^k \psi\}_{k=0}^{n-1}$. The representation in this matrix is precisely the claimed GGT matrix. Uniqueness is easy. \square

Here is the main result of this section:

Theorem 4.2. *Let U be an $n \times n$ unitary matrix with cyclic vector ψ_0 with $\det(U) = (-1)^{n+1}$. Then there is a unique basis $\{e_j\}_{j=1}^n$ in which U is a GGT unitary $\mathcal{G}(\{\beta_j\}_{j=0}^{n-1})$ and $\psi_0 = e_n$. The $\{\beta_j\}_{j=0}^{n-1}$ are uniquely determined by the pair U, ψ_0 and the β 's are related to the α 's of Theorem 4.1 by*

$$\beta_j = \bar{\alpha}_{n-2-j}, \quad j = 0, \dots, n - 2; \quad \beta_{n-1} = -1 \tag{4.1}$$

Remarks. 1. At first sight, it may be puzzling that U is unitarily equivalent to GGT matrices with, in general, two different Verblunsky coefficients. After all, Wendroff's Theorem for OPUC implies that a GGT contraction determines uniquely all its Verblunsky coefficients but as pointed out in Remark 1 to Theorem 4.1 this is not true for GGT unitaries as this theorem dramatically demonstrates.

2. The intuition is simple. A look at (2.30) shows that when $\alpha_{n-1} = -1$, there is a covariance under reflection about the reverse main diagonal so long as one relabels the α 's via (4.1).

The covariance under reflection suggests we want to reverse the order of the GGT basis. That has the advantage of taking e_1 to e_n as we want but such a reversal not only reflects, it also gives a transpose so the following is critical

Lemma 4.3. *Given any set of Verblunsky coefficients $\{\alpha_j\}_{j=0}^{n-1}$, with $\alpha_j \in \mathbb{D}, j = 0, \dots, n - 2$, and $\alpha_{n-1} \in \overline{\mathbb{D}}$, the matrix representation of A , the associated compressed multiplication operator, in the basis $\{\chi_j\}_{j=0}^{n-1}$ where $\chi_j = \bar{\varphi}_j$ is the transpose of the GGT matrix.*

Remark. It is perhaps surprising a finite GGT matrix is unitarily equivalent to its transpose. Using the so called AGR factorization (see [64]) one can write the unitary operator explicitly. If $U_1 \dots U_n$ is the AGR factorization and $W_j = U_j \dots U_n$ and $W = W_n \dots W_2$, then one can show that $W\mathcal{G}W^{-1} = \mathcal{G}^t$.

Proof. Clearly

$$\begin{aligned} \langle \chi_j, z\chi_k \rangle &= \int \bar{\chi}_j(e^{i\theta})e^{i\theta}\chi_k(e^{i\theta})d\mu(\theta) \\ &= \int \varphi_j(e^{i\theta})e^{i\theta}\bar{\varphi}_k(e^{i\theta})d\mu(\theta) \\ &= \mathcal{G}_{kj} \quad \square \end{aligned} \tag{4.2}$$

Proof of Theorem 4.2. As in the proof of Theorem 4.1, realize the space on which U acts as $L^2(\partial\mathbb{D}, d\mu)$ where $d\mu$ is the spectral measure for ψ_0 and U . Let $\varphi_j, j = 0, \dots, n - 1$ be the orthonormal polynomials for $d\mu$, so, in particular, $\psi_0 = \varphi_0$. Let $e_j = \bar{\varphi}_{n-j}$ so $e_n = \psi_0$. By either the explicit formulae, (2.25) for $\mathcal{G}_{k\ell}$ or the form of the matrix (2.30), one sees that the transpose of $\mathcal{G}(\{\alpha_j\}_{j=0}^{n-1})$ in a basis run backwards is $\mathcal{G}(\{\beta_j\}_{j=0}^{n-1})$ with β given by (4.1). Given Lemma 4.3, we see that the matrix of U in basis $\{e_j\}_{j=1}^n$ is $\mathcal{G}(\{\beta_j\}_{j=0}^{n-1})$ as claimed. \square

We supposed that $\det(U) = (-1)^{n+1}$ for simplicity of calculation (which implies that $\alpha_{n-1} = -1$). It is easy to see that, more generally, one has that

Theorem 4.4. *Let U be an $n \times n$ unitary matrix with cyclic vector ψ_0 . Then there is a unique basis $\{e_j\}_{j=1}^n$ in which U is a GGT unitary $\mathcal{G}(\{\beta_j\}_{j=0}^{n-1})$ and $\psi_0 = e_n$. The $\{\beta_j\}_{j=0}^{n-1}$ are uniquely determined by the pair U, ψ_0 and the β 's are related to the α 's of Theorem 4.2 by*

$$\beta_j = -\alpha_{n-1}\bar{\alpha}_{n-2-j}, j = 0, \dots, n - 2; \quad \beta_{n-1} = \alpha_{n-1} \tag{4.3}$$

5. $S_n =$ compressed multiplication via GGT matrices

In this Section, we give our first proof of Theorem 2. Let $A \in S_n$. Pick once and for all a unit vector $x_{n-1} \in \text{ran}(\mathbf{1} - A^*A)$. Then there exists some $\rho_{n-1} \in (0, 1]$ such that

$$A^*A = \mathbf{1} - \rho_{n-1}^2 \langle x_{n-1}, \cdot \rangle x_{n-1} \tag{5.1}$$

(for now, ρ_{n-1} is just some number in $(0, 1]$ but eventually we will see that it is the same as a ρ_{n-1} associated to OPUC and defined in (2.4)). By using the polar decomposition, we can write $A = UB$ where U is unitary with $\det(U) = (-1)^{n+1}$ and B is diagonal in a basis including x_{n-1} as its last element so that B has 1’s along diagonal except for the last diagonal element which is some complex number a with $|a|^2 + \rho_{n-1}^2 = 1$. To see this start with $A = V|A|$ as polar decomposition, and shift to a basis where $|A|$ is diagonal. Then pick $U = VD$ where D is diagonal and unitary with ones along diagonal except for the last slot whose phase is picked to arrange $\det(U) = (-1)^{n+1}$ and then take $B = D^*|A|$.

Lemma 5.1. *For any A obeying (5.1), with the above decomposition $A = UB$, we have that x_{n-1} is cyclic for U if and only if A has no eigenvalue of magnitude 1.*

Proof. If x_{n-1} was not cyclic for U , let \mathcal{K} be the orthogonal complement of the cyclic subspace generated by U and x_{n-1} . Then U^{-1} is a polynomial in U (look at the secular equation) so U^* leaves \mathcal{K}^\perp invariant and thus U leaves \mathcal{K} invariant. Since \mathcal{K} is orthogonal to x_{n-1} , we have that $B \upharpoonright \mathcal{K} = \mathbf{1}$. It follows that any eigenvector of $U \upharpoonright \mathcal{K}$ is an eigenvector of A with the same eigenvalue so A has eigenvalues of magnitude 1 as claimed.

Conversely, if $A\eta = \lambda\eta$ with $\|\eta\| = |\lambda| = 1$, then $\|A\eta\| = \|\eta\|$ so η is orthogonal to x_{n-1} , implying that $B\eta = \eta$, and is thus also an eigenvector of U . It follows that $\langle \eta, U^k x_{n-1} \rangle = \langle (U^*)^k \eta, x_{n-1} \rangle = 0$ so η is orthogonal to the cyclic subspace of U and x_{n-1} so x_{n-1} is not cyclic for U . \square

First Proof of Theorem 2. Let $A \in S_n$. Then, as above $A = UB$ where, by Lemma 5.1, U has x_{n-1} as cyclic vector. By Theorem 4.2, U is a GGT unitary where $\beta_{n-1} = -1$ and B is diagonal with 1’s along the diagonal except for a in the lower right corner. It follows that in this basis A is a GGT contraction with Verblunsky coefficients

$$\gamma_j = \begin{cases} \beta_j, & \text{if } 0 \leq j \leq n - 2 \\ -a, & \text{if } j = n - 1 \end{cases} \tag{5.2}$$

Thus A is a compressed multiplication operator. \square

6. $S_n =$ compressed multiplication via inverse Szegő recursion

In this Section, we give our second proof of Theorem 2. Let $A \in S_n$. We will carefully pick a basis $\{x_j\}_{j=0}^{n-1}$ with x_{n-1} a unit vector as in Section 5, for which

$$A^*A = \mathbf{1} - \rho_{n-1}^2 x_{n-1} x_{n-1}^* \tag{6.1}$$

and show that this basis is the basis of OPUC for a compressed multiplication operator. Our motivation for the construction is inverse Szegő recursion.

In (6.1), the unit vector $x_{n-1} \in \mathcal{H}_n$, the space on which A , which we'll now denote by A_n , acts, so $x_{n-1} \in \text{ran}(\mathbf{1} - A_n^* A_n)$, and $\rho_{n-1} \in (0, 1]$.

Suppose that $n \geq 2$ (Theorem 2 is trivial when $n = 1$). Let \mathcal{H}_{n-1} be the $n - 1$ -dimensional orthogonal complement of x_{n-1} so

$$\mathcal{H}_{n-1} = \ker(\mathbf{1} - A_n^* A_n) \tag{6.2}$$

and let P_{n-1} be the projection on \mathcal{H}_{n-1} .

Since A_n is an isometry on \mathcal{H}_{n-1} , $A_n[\mathcal{H}_{n-1}]$, the image of \mathcal{H}_{n-1} under A_n has dimension $n - 1$ so we can pick a unit vector $y_{n-1} \in A_n[\mathcal{H}_{n-1}]^\perp$ unique up to overall phase factor. For now we won't specify that phase but we will later (essentially so that if x_{n-1} is φ_{n-1} , then y_{n-1} is φ_{n-1}^*). Let Q_{n-1} be the projection onto $A_n[\mathcal{H}_{n-1}]$.

Proposition 6.1. *First, $Q_{n-1}x_{n-1} \neq 0$. Secondly there exists $a_{n-1} \in \mathbb{C}$ so that*

$$\rho_{n-1}^2 + |a_{n-1}|^2 = 1 \tag{6.3}$$

$$A_n x_{n-1} = a_{n-1} y_{n-1} \tag{6.4}$$

Thirdly, there exists a unique unit vector $x_{n-2} \in \mathcal{H}_{n-1}$ so that $A_n x_{n-2}$ is a positive multiple of $Q_{n-1}x_{n-1}$. One has that

$$A_n x_{n-2} = \rho_{n-2}^{-1} [x_{n-1} + a_{n-2} y_{n-1}] \tag{6.5}$$

for some $\rho_{n-2} \in (0, 1]$ and $a_{n-2} \in \mathbb{C}$ where

$$\rho_{n-2}^2 + |a_{n-2}|^2 = 1 \tag{6.6}$$

Let $A_{n-1} = P_{n-1} A_n P_{n-1}$ as an operator on \mathcal{H}_{n-1} . Then $A_{n-1} \in S_{n-1}$ with

$$\mathbf{1} - A_{n-1}^* A_{n-1} = \rho_{n-2}^2 \langle x_{n-2}, \cdot \rangle x_{n-2} \tag{6.7}$$

Proof. If $Q_{n-1}x_{n-1} = 0$, then A_n leaves \mathcal{H}_{n-1} invariant. Since A_n is an isometry on that space, the restriction of A_n to that invariant subspace would be unitary which means that

A_n would have eigenvalues of magnitude one contradicting the hypothesis that $A \in S_n$ (it is interesting that this is how the eigenvalue condition enters in the proof!).

By (6.1), if $u \in \mathcal{H}_{n-1}$, then $\langle A_n u, A_n x_{n-1} \rangle = \langle u, x_{n-1} \rangle = 0$ so $A_n x_{n-1} \in A_n[\mathcal{H}_{n-1}]^\perp$ proving (6.4). Formula (6.1), which says that $\|A_n x_{n-1}\|^2 = 1 - \rho_{n-1}^2$ then implies (6.3).

Since A_n is an isometry on \mathcal{H}_{n-1} , there is a unique vector $u \in \mathcal{H}_{n-1}$ so that $A_n u = Q_{n-1} x_{n-1}$ and one has that $\rho_{n-2} \equiv \|u\| \in (0, 1]$. Let $x_{n-2} = \rho_{n-2}^{-1} u$ and note that $(\mathbf{1} - Q_{n-1})x_{n-1} = -a_{n-2} y_{n-1}$ for some $a_{n-2} \in \mathbb{C}$. This proves (6.5). Writing (6.5) as

$$x_{n-1} = \rho_{n-2} A_n x_{n-2} - a_{n-2} y_{n-1} \tag{6.8}$$

and noting that y_{n-1} is orthogonal to Ax_{n-2} implies (6.6).

Since A_n is an isometry on \mathcal{H}_{n-1} , it maps the orthogonal complement of u in \mathcal{H}_{n-1} into \mathcal{H}_{n-1} so A_{n-1} is an isometry on that orthogonal complement which proves (6.7). Moreover, if $A_{n-1}v = cv$ with v a unit vector and $|c| = 1$, then $\|P_{n-1}A_nv\| = 1 \geq \|A_nv\|$ so $P_{n-1}A_nv = A_nv$ and v is an eigenvector of A_n and there is a contradiction, i.e. $A_{n-1} \in S_{n-1}$. \square

If $n \geq 2$, we define, consistently with viewing $A_{n-1} \in S_{n-1}$, \mathcal{H}_{n-2} to be the complement of x_{n-2} in \mathcal{H}_{n-1} , i.e.

$$\mathcal{H}_{n-2} = \ker(\mathbf{1} - A_{n-1}^* A_{n-1}) \tag{6.9}$$

and Q_{n-2} , an operator on \mathcal{H}_{n-1} , the projection onto $A_{n-1}[\mathcal{H}_{n-2}]$. If $n = 2$, then we take $Q_0 = 0$.

Proposition 6.2. *One can choose a unit vector $y_{n-2} \in \mathcal{H}_{n-1}$ with $Q_{n-2}y_{n-2} = 0$ (i.e. $y_{n-2} \in \mathcal{H}_{n-1} \cap [A_{n-1}[\mathcal{H}_{n-2}]]^\perp$) so that*

$$Ax_{n-2} = \rho_{n-2}x_{n-1} + a_{n-2}y_{n-2} \tag{6.10}$$

$$\langle y_{n-1}, y_{n-2} \rangle = \rho_{n-2} > 0 \tag{6.11}$$

Remarks. 1. (6.10) is, of course, (direct) Szegő recursion (2.3) and, as noted, (6.5) is inverse Szegő recursion, (2.7). We note that (6.11) is motivated by [60, (1.5.61)] while (6.12) below is its * dual and (6.15) below is a special case of [60, (1.5.60)].

2. This also holds when $n = 2$. \mathcal{H}_{n-1} is then one dimensional so y_{n-1} is a multiple of x_{n-1} since they are both vectors in \mathcal{H}_{n-1} .

Proof. We start taking the inner product of $A_n x_{n-2}$ with (6.8) which given that $Ax_{n-2} \perp y_{n-1}$ and $\|A_n x_{n-2}\| = 1$ implies that

$$\langle Ax_{n-2}, x_{n-1} \rangle = \rho_{n-2} \tag{6.12}$$

Since $\mathbf{1} - P_{n-1} = \langle x_{n-1}, \cdot \rangle x_{n-1}$, where P_{n-1} be the projection on \mathcal{H}_{n-1} , this implies that

$$(\mathbf{1} - P_{n-1})Ax_{n-2} = \rho_{n-2}x_{n-1} \tag{6.13}$$

On the other hand, let $w \equiv P_{n-1}Ax_{n-2}$ and let $u \in \mathcal{H}_{n-2}$. Since A_n preserves inner products on \mathcal{H}_{n-1} , we see that $\langle A_nu, A_nx_{n-2} \rangle = 0$. Since $A_nu \in \mathcal{H}_{n-1}$, we see that $P_{n-1}A_nu = A_nu$. Thus

$$\begin{aligned} \langle A_nu, w \rangle &= \langle A_nu, P_{n-1}Ax_{n-2} \rangle = \langle P_{n-1}A_nu, A_nx_{n-2} \rangle \\ &= \langle A_nu, A_nx_{n-2} \rangle = 0 \end{aligned}$$

Since $w \in \mathcal{H}_{n-1}$, we conclude that w is a multiple of y_{n-2} . Therefore, for some b , we have that

$$Ax_{n-2} = (\mathbf{1} - P_{n-1})Ax_{n-2} + P_{n-1}Ax_{n-2} = \rho_{n-2}x_{n-1} + by_{n-2} \tag{6.14}$$

Since $y_{n-2} \perp x_{n-1}$, we have that $b^2 + \rho_{n-1}^2 = 1$, so we can pick the phase of y_{n-2} so that (6.10) holds.

With this choice, we need to prove (6.11). Actually, if $a_{n-2} = 0$ the above argument doesn't fix the phase of y_{n-2} so in that case we will show that (6.11) can be used to determine the phase. Taking the inner product of (6.8) with y_{n-1} and using that $y_{n-1} \perp Ax_{n-2}$ we get that

$$\langle y_{n-1}, x_{n-1} \rangle = -a_{n-2} \tag{6.15}$$

Taking the inner product of (6.10) with y_{n-1} and again using $y_{n-1} \perp Ax_{n-2}$ we get that

$$0 = \rho_{n-2}\langle y_{n-1}, x_{n-1} \rangle + a_{n-2}\langle y_{n-1}, y_{n-2} \rangle \tag{6.16}$$

Multiplying (6.15) by ρ_{n-2} we conclude that

$$a_{n-2}\langle y_{n-1}, y_{n-2} \rangle = a_{n-2}\rho_{n-2} \tag{6.17}$$

which implies (6.11) if $a_{n-1} \neq 0$.

If $a_{n-2} = 0$, i.e. $\rho_{n-2} = 1$, then by (6.15), $y_{n-1} \perp x_{n-1}$ and, thus, $y_{n-1} \in \mathcal{H}_{n-1}$. Since it is orthogonal to $A_n[\mathcal{H}_{n-1}]$, it is also orthogonal to $A_n[\mathcal{H}_{n-2}]$ and thus we can pick $y_{n-1} = y_n$ so that (6.11) holds. \square

Second Proof of Theorem 2. By iterating the above construction, having made a choice of phase for x_{n-1} and y_{n-1} , we get $\mathcal{H}_1 \subset \mathcal{H}_2 \subset \dots \subset \mathcal{H}_n$ with $\dim(\mathcal{H}_j) = j$ and $x_j, y_j \in \mathcal{H}_{j+1}$, $j = 0, 1, \dots, n - 1$. Moreover $x_j \perp \mathcal{H}_j$ so $\{x_j\}_{j=0}^{n-1}$ is an orthonormal basis.

Changing the choice of phase of y_{n-1} by replacing it by $e^{i\theta}y_{n-1}$ replaces each y_j by $e^{i\theta}y_j$ (same θ) since we rely on $\langle y_j, y_{j-1} \rangle > 0$. Moreover, $x_0, y_0 \in \mathcal{H}_1$, a one dimensional space, so we can choose $e^{i\theta}$ so that $y_0 = x_0$. We make that choice once and for all, which changes the phases of the a_j and then we let $\alpha_j = \bar{a}_j$.

By construction, we have for $j = 0, \dots, n - 2$

$$Ax_j = \rho_j x_{j+1} + \bar{\alpha}_j y_j \tag{6.18}$$

$$x_{j+1} = \rho_j Ax_j - \bar{\alpha}_j y_{j+1} \tag{6.19}$$

and by (6.4) that

$$Ax_{n-1} = \bar{\alpha}_{n-1} y_{n-1} \tag{6.20}$$

Multiplying (6.18) by ρ_j and substituting in (6.19) gives

$$x_{j+1} = \rho_j^2 x_{j+1} + \rho_j \bar{\alpha}_j y_j - \bar{\alpha}_j y_{j+1} \tag{6.21}$$

Since $1 - \rho_j^2 = \alpha_j \bar{\alpha}_j$, if $\alpha_j \neq 0$, we can divide it out and get

$$y_{j+1} = \rho_j y_j - \alpha_j x_{j+1} \tag{6.22}$$

If $\alpha_j = 0$, we saw that $y_{j+1} = y_j$ so (6.22) still holds.

We thus have (6.18)/(6.20)/(6.22) which are the same as (2.26)/(2.27) which can be used with $x_0 = y_0$ to show that, in $\{x_j\}_{j=0}^{n-1}$ basis, A is given by a GGT matrix proving Theorem 2. \square

Proof of Theorem 3. The result is essentially a restatement of Wendroff Theorem for OPUC. Given the eigenvalues $\{z_j\}_{j=1}^n$, take $\Phi_n(z) = \prod_{j=1}^n (z - z_j)$. By Wendroff’s Theorem, this is an OPUC for measure and if A is the corresponding compressed multiplication operator, then its eigenvalues are the zeros of Φ_n , so the required set. In fact, one proof of Wendroff’s theorem [60, Theorem 1.7.5] shows that one can take the measure to be $d\theta/|\varphi_n|^2$, the so-called Bernstein–Szegő measure associated to φ_n .

By Theorem 2.10, two compressed multiplication operators with the same eigenvalues have the same Φ_n and are the exact same operator acting on \mathcal{P}_{n-1} . Thus, they are unitarily equivalent. \square

7. The numerical range

In this section, we prove Theorems 4 and 5. Since we’ve seen that S_n agrees with compressed multiplication operators, we suppose A acts on \mathcal{P}_{n-1} as a compressed multiplication operators with Verblunsky coefficients $\{\alpha_j\}_{j=0}^{n-1}$. We will realize the unitary operators $\{U_\lambda\}_{\lambda \in \partial\mathbb{D}}$ associated to the degree $n + 1$ POPUC as operators on \mathcal{P}_n . We have

$\{\varphi_j\}_{j=0}^n$ as an orthonormal basis for \mathcal{P}_n in $L^2(\partial\mathbb{D}, d\mu_\lambda)$ for the measures, $d\mu_\lambda$, associated to each of U_λ . For $\psi \in \mathcal{P}_n$, we have that

$$\psi \in \mathcal{P}_{n-1} \iff \langle \varphi_n, \psi \rangle = 0 \tag{7.1}$$

Szegő recursion says that

$$A\varphi_j = \begin{cases} \rho_j\varphi_{j+1} + \bar{\alpha}_j\varphi_j^*, & j = 0, \dots, n-2 \\ \bar{\alpha}_j\varphi_j^*, & j = n-1 \end{cases} \tag{7.2}$$

while

$$U_\lambda\varphi_j = \begin{cases} \rho_j\varphi_{j+1} + \bar{\alpha}_j\varphi_j^*, & j = 0, \dots, n-1 \\ \bar{\lambda}_j\varphi_j^*, & j = n \end{cases} \tag{7.3}$$

U_λ has eigenvalues at those $\{w_j\}_{j=1}^{n+1}$, labelled so that $0 \leq \arg w_j < \arg w_{j+1} < 2\pi, j = 1, \dots, n$, with $\Phi_{n+1}(w_j) = 0$, where $\Phi_{n+1}(z) = z\Phi_n(z) - \bar{\lambda}\Phi_n^*(z)$. The eigenvectors are

$$\eta_j(z) = N_j^{-1}\Phi_{n+1}(z)/(z - w_j) = N_j^{-1} \prod_{k \neq j} (z - w_k) \tag{7.4}$$

where $N_j > 0$ is a normalization factor. Of course $\Phi_{n+1}, w_j, \eta_j, N_j$ are all λ dependent but we suppress this dependence unless we need to be explicit.

We define

$$m_j(\lambda) = |\langle \varphi_n, \eta_j \rangle|^2 \tag{7.5}$$

the Fourier coefficients of φ_n in the orthonormal basis $\{\eta_j\}_{j=1}^{n+1}$ so

$$m_j(\lambda) > 0, \quad \sum_{j=1}^{n+1} m_j(\lambda) = 1 \tag{7.6}$$

The spectral measure for φ_n and U_λ is

$$d\nu^\lambda(z) = \sum_{j=1}^{n+1} m_j \delta_{w_j} = |\varphi_n(z)|^2 d\mu^\lambda(z) \tag{7.7}$$

where

$$d\mu^\lambda(z) = \sum_{j=1}^{n+1} q_j \delta_{w_j}; \quad q_j = 1 / \sum_{k=0}^{n+1} |\varphi_k(w_j)|^2 \tag{7.8}$$

(the Christoffel numbers, q_j^{-1} , here sum to $n + 1$ while in Section 4, they only summed to n because there we were discussing $n \times n$ unitaries whereas here our unitary operators are $(n + 1) \times (n + 1)$).

We won't need them but we note there are explicit formulae for N_j and m_j , viz

$$N_j = q_j^{1/2} \prod_{k \neq j} |w_j - w_k|; \quad m_j = q_j |\varphi_n(w_j)|^2 \tag{7.9}$$

We will need that these and the w_j are real analytic in λ although one can also get that from eigenvalue perturbation theory [65].

Lemma 7.1. *Up to phase factor, the span of η_j and η_k , $j \neq k$ has a unique unit vector, ψ , in \mathcal{P}_{n-1} and it is given by*

$$\psi = [\langle \varphi_n, \eta_j \rangle \eta_k - \langle \varphi_n, \eta_k \rangle \eta_j] / \sqrt{m_j + m_k} \tag{7.10}$$

One has that

$$\zeta_{jk} := \langle \psi, U_\lambda \psi \rangle = \langle \psi, A \psi \rangle = \frac{m_j z_k + m_k z_j}{m_j + m_k} \tag{7.11}$$

Proof. (7.10) is a consequence of (7.1). (7.11) for U_λ follows from the Fourier expansion for U_λ and that this equals $\langle \psi, A \psi \rangle$ follows from $A = P_n U_\lambda P_n$. \square

Theorem 7.2 (=Theorem 4). *For any $j, k \in \{1, \dots, n + 1\}$, $j \neq k$, the point, ζ , on the line between w_j and w_k with*

$$|\zeta - w_j| / |\zeta - w_k| = m_j / m_k \tag{7.12}$$

lies in $N(A)$. For $k = j + 1$ (or $j = n + 1, k = 1$), this is the only point on the line which lies in $N(A)$ and that line is tangent to $\partial N(A)$ at the point ζ .

Remark. The polygon with vertices $\{w_j\}_{j=1}^{n+1}$ is $\partial N(U_\lambda)$ and circumscribes $N(A)$, so it is a kind of Poncelet polygon.

Proof. Given the formula (7.11) for ζ , (7.12) is a direct calculation. Since $\psi \in \mathcal{P}_{n-1}$, we have that $\zeta \in N(A)$.

Since U_λ is normal, $N(U_\lambda)$ is the convex hull of the eigenvalues $\{w_j\}_{j=1}^{n+1}$ and $\langle \psi, U_\lambda \psi \rangle = \sum_{j=1}^{n+1} w_j |\langle \psi, \eta_j \rangle|^2$ is in the segment from w_j to w_{j+1} if and only if ψ is the space spanned by η_j and η_{j+1} . Thus the only point on that line in $N(A)$ comes from ψ given by (7.10) with $\langle \psi, A \psi \rangle$ given by (7.11). Since $N(A) \subset N(U_\lambda)$, $N(A)$ lies on one side of the line segments and therefore the line segment is tangent. \square

Theorem 7.3 (=Theorem 5). $\partial N(A)$ is a real analytic curve. Each $\zeta \in \partial N(A)$ is a tangent point of an edge of some $N(U_\lambda)$. Moreover

$$N(A) = \bigcap_{\lambda \in \partial \mathbb{D}} N(U_\lambda) \tag{7.13}$$

Proof. As noted above, the w_j and m_j are real analytic functions of $\bar{\lambda}$, so ζ is real analytic. Since $\partial N(A)$ is a convex curve, there is, in the sense of a line with the curve on one side of the line, a tangent at each point (although, a priori, it might not be unique). This tangent must meet $\partial \mathbb{D}$ at a point which is a zero of some $\Phi_{n+1}(z; \lambda)$ and so the tangent is an edge of that $N(U_\lambda)$. That means that the function ζ fills out the entire set $\partial N(A)$. The analyticity then implies uniqueness of tangent and so proves strict convexity. Thus, we need only prove (7.13).

Since A is a restriction of each U_λ , we have that $N(A) \subset N(U_\lambda)$, so the \subseteq of (7.13) is immediate.

Pick $\xi \in N(A)^{int}$, the interior of $N(A)$ (which is non–empty by the strict convexity). Let $u \notin N(A)$. The line segment from ξ to u meets $\partial N(A)$ in a unique point, ζ . ζ lies on the edge of some $\partial N(U_\lambda)$. Since this edge is tangent to $\partial N(A)$, it must be distinct from the line from ξ to u which implies that $u \notin N(U_\lambda)$. Thus $N(A)^c \subset \cup N(U_\lambda)^c$ proving the \supseteq half of (7.13). \square

8. The Schur functions associated to OPUC

In this section and the next, we will prove Theorems 6 and 7 and the following related result

Theorem 8.1. Fix Verblunsky coefficients $\{\alpha_j\}_{j=0}^{n-1}$ and corresponding OPUC $\{\Phi_j\}_{j=0}^n$. Let

$$B_n(z) = \frac{\Phi_n(z)}{\Phi_n^*(z)} \tag{8.1}$$

Then B_n is a Schur function and for any $\lambda \in \partial \mathbb{D}$, the Schur iterates of λB_n are $\lambda B_{n-1}, \lambda B_{n-2}, \dots, \lambda B_0 = \lambda$ and the Schur parameters are

$$\gamma_j(\lambda B_n) = \begin{cases} -\lambda \bar{\alpha}_{n-1-j}, & j = 0, \dots, n-1 \\ \lambda, & j = n \end{cases} \tag{8.2}$$

The Carathéodory function for the associated measure is

$$F_n(z) = -\frac{\Phi_{n+1}(z; -\lambda)}{\Phi_{n+1}(z; \lambda)} \tag{8.3}$$

The measure associated to λB_n is the measure $d\nu^\lambda$ of (7.7).

In this section, we'll prove this theorem and use it to give the first proof of Theorem 7. In the next section, we'll use it to give one proof of Theorem 6. In that section we'll also give GGT matrix proofs of Theorems 6 and 7. Formulae like (8.2) are implicit in the work of Khrushchev [46]. We'll say more about his work and its relation to the first proof of Theorem 7 in the next Section.

Start of Proof of Theorem 8.1. If $e^{i\theta} \in \partial\mathbb{D}$, then we have that $|\Phi_n(e^{i\theta})| = |\Phi_n^*(e^{i\theta})|$. Moreover, since Φ_n has all of its zeros in \mathbb{D} , Φ_n^* has none, so $B_n(z)$ is analytic in \mathbb{D} with $|B_n(e^{i\theta})| = 1$ for $e^{i\theta} \in \partial\mathbb{D}$. By the maximum principle, B_n is a Schur function; indeed, up to a phase factor, it is the Blaschke product of the zeros of Φ_n .

By Szegő recursion, (2.3) and the result of applying τ_{n+1} to it, we have that

$$\begin{aligned} \lambda B_n(z) &= \lambda \frac{z\Phi_{n-1}(z) - \bar{\alpha}_{n-1}\Phi_{n-1}^*(z)}{\Phi_{n-1}^*(z) - \alpha_{n-1}z\Phi_{n-1}(z)} \\ &= \frac{-\lambda\bar{\alpha}_{n-1} + z(\lambda B_{n-1}(z))}{1 - \bar{\lambda}\alpha_{n-1}z(\lambda B_{n-1}(z))} \end{aligned} \tag{8.4}$$

which precisely says that $\gamma_0(\lambda B_n) = -\lambda\bar{\alpha}_{n-1}$ and that λB_{n-1} is the first Schur iterate. By the obvious repetition, we get (8.2) and the claimed full list of Blaschke iterates.

To get (8.3), we note that

$$\begin{aligned} F_n(z) &= \frac{1 + z\lambda B_n(z)}{1 - z\lambda B_n(z)} \\ &= \frac{z\Phi_n(z) + \bar{\lambda}\Phi_n^*(z)}{-(z\Phi_n(z) - \bar{\lambda}\Phi_n^*(z))} \end{aligned} \tag{8.5}$$

proving (8.3). This proves the entire theorem except for the identification of the measure to which we now turn. \square

We have just shown that the Carathéodory function of $d\nu$, the measure with Schur function λB_n is

$$F(z) = \frac{1 + \lambda z B_n(z)}{1 - \lambda z B_n(z)} = -\frac{z\varphi_n(z) + \bar{\lambda}\varphi_n^*(z)}{z\varphi_n(z) - \bar{\lambda}\varphi_n^*(z)} \tag{8.6}$$

On the other hand, the Carathéodory function of the measure defined by $\Phi_{n+1}(z; \lambda)$ is, by [60, (3.2.4)],

$$G(z) = \frac{\Psi_{n+1}^*(z)}{\Phi_{n+1}^*(z)} = -\frac{\Psi_{n+1}(z)}{\Phi_{n+1}(z)} \tag{8.7}$$

where we used (2.23). Here $\Psi_{n+1}(z; \lambda) \equiv z\Psi_n(z) + \bar{\lambda}\Psi_n^*(z)$ with Ψ_n the second kind OPUC; see Theorem 2.7.

Lemma 8.2. Fix λ and z_0 in $\partial\mathbb{D}$. Suppose that $\Phi_{n+1}(z_0; \lambda) = 0$. Then

$$\frac{1}{\Psi_{n+1}(z_0)} = \frac{\overline{\varphi_n(z_0)}}{2z_0 \prod_{j=0}^{n-1} \rho_j} \tag{8.8}$$

Proof. By (2.13), we have that

$$\Psi_n^*(z_0)\Phi_n(z_0) + \Phi_n^*(z_0)\Psi_n(z_0) = 2z_0^n \prod_{j=0}^{n-1} \rho_j^2 \tag{8.9}$$

By hypothesis, we have that $\Phi_n(z_0) = \overline{z_0\lambda}\Phi_n^*(z_0)$, so (8.9) becomes

$$\overline{z_0}\Phi_n(z_0)^*\Psi_{n+1}(z_0) = 2z_0^n \prod_{j=0}^{n-1} \rho_j^2 \tag{8.10}$$

Since $\Phi_n^*(z_0) = z_0^n \overline{\varphi_n(z_0)} \prod_{j=0}^{n-1} \rho_j$ we get (8.8). \square

End of the Proof of Theorem 8.1. Let z_0 be a zero of the POPUC $\Phi_{n+1}(z; \lambda)$. It suffices to show that for any such z_0 , one has that

$$\lim_{\epsilon \downarrow 0} F((1 - \epsilon)z_0)/G((1 - \epsilon)z_0) = |\varphi_n(z_0)|^2 \tag{8.11}$$

We note that by (8.6) and (8.7), we have that

$$\frac{F(z)}{G(z)} = \frac{z\Phi_n(z) + \bar{\lambda}\Phi_n^*(z)}{\Psi_{n+1}(z)} \tag{8.12}$$

The right side of (8.12) has a limit at z_0 which is $2z_0\Phi_n(z_0)/\Psi_{n+1}(z_0)$. By (8.8), this is $|\varphi_n(z_0)|^2$. \square

9. M-functions of POPUC

In this section, we'll begin with a proof of Theorem 6 following up on the last section and then provide a totally different approach to proving Theorems 6 and 7 using GGT matrices. Finally, we'll discuss the relation of these theorems to earlier work on OPUC.

We'll call a function like

$$\int \frac{1}{z - e^{i\theta}} d\nu(\theta) \tag{9.1}$$

which appears on the left side of (3.4) an M -function in analogy with the Weyl m -function of OPRL and the theory of second order ODEs (although those functions have $(x - z)^{-1}$, not $(z - x)^{-1}$ where x is the variable of integration).

First Proof of Theorem 6. Since $\frac{w+z}{w-z} = 1 - 2z\frac{1}{z-w}$, we have that

$$F(z) = 1 - 2zM(z) \Rightarrow M(z) = \frac{1 - F(z)}{2z} \tag{9.2}$$

where F , the Carathéodory of $d\nu$, is given by (8.6). (9.2) and (8.5) imply (3.4).

This provides a relation between a POPUC and the measure, $d\nu$, associated to it. Given a Blaschke product or the POPUC associated to it (the one with the same zeros), one sees from this that the combination on the right of (1.3) or (3.4) has the form on the left of these equations. Conversely, given a set of $n + 1$ points on $\partial\mathbb{D}$ and probability weights, we can form the associated measure ν and its M -function M as in (9.1). By (2.9) and (9.2), M can be expressed in terms of a Schur function. It remains to use Theorem 8.1 to show that M has the form of the right hand side of (1.3), or equivalently, of (3.4). \square

The following direct proof of (3.4) does not rely on identifying a Schur function. We regard it as the simplest proof of (1.3). After we'd found this proof, we found almost the identical argument in [35, Section 2.2].

Second Proof of Theorem 6. Let U_λ be given by Theorem 2.11. Then by the spectral theorem and the definition of $d\nu$, we have that

$$\int \frac{1}{z - e^{i\theta}} d\nu(\theta) = \langle \varphi_n, (z - U_\lambda)^{-1} \varphi_n \rangle \tag{9.3}$$

In the GGT representation, φ_n is the vector $(0, 0, \dots, 1)^t$, so Cramer's rule says that

$$\langle \varphi_n, (z - U_\lambda)^{-1} \varphi_n \rangle = \frac{\det(z - \mathcal{G}_n(\{\alpha_j\}_{j=0}^{n-1}))}{\det(z - \mathcal{G}_{n+1}(\{\alpha_j\}_{j=0}^{n-1} \cup \{\lambda\}))} \tag{9.4}$$

since the result of dropping the last row and column of GGT matrix is a GGT matrix of one degree less. By (2.20), the denominator of (9.4) is $\Phi_{n+1}(z; \lambda)$ and by (2.16), the numerator is $\Phi_n(z)$. These facts imply (3.4). Once one has this, the rest of the proof is the same as the last paragraph of the first proof. \square

Second Proof of Theorem 7. Except for replacing n by $n + 1$ and α_{n-1} by λ , this is just Theorem 4.4. \square

Finally, we turn to the connection of Theorems 6, 7 and 8.1 to earlier work of Khrushchev [46] and Simon [60]. When $\lambda = 1$, (8.2), proven as we do using Szegő recursion, is in Khrushchev's paper. He uses it in part to prove what [61, Theorem 9.2.4] calls Khrushchev's formula that if $d\eta$ is a non-trivial probability measure, then the Schur function of $|\varphi_n(e^{i\theta})|^2 d\eta(\theta)$ is $f_n(z)B_n(z)$ where f_n is the n th Schur iterate of the Schur function of $d\eta$, i.e. the Schur function with Schur parameters $\gamma_j(f_n) = \alpha_{n+j}(d\eta)$, $j = 0, 1, \dots$, and B_n is given by (8.1).

If now, we take $\alpha_n \rightarrow \lambda \in \partial\mathbb{D}$, the measure $d\eta$ converges to the measure we called $d\mu^\lambda$, so $|\varphi_n(e^{i\theta})|^2 d\eta$ converges to the measure that we called $d\nu^\lambda$ and the Khrushchev formula in the limit says that the Schur function of that measure is λB_n which is the final assertion in Theorem 8.1 which we saw above is the essence of the first proof of Theorem 6.

In [60, Section 4.4], Simon, using rank two perturbation theory computed matrix elements of $(\mathcal{C} - z)^{-1}$, where \mathcal{C} is the CMV matrix, which can be used to also find matrix elements of $(\mathcal{G} - z)^{-1}$ which provides another proof of Theorem 6.

10. Wendroff type theorems

Simon [63] has three theorems about zero interlacing involving POPUC which he proves using rank one perturbation theory (recently Castillo–Petronilho [12] have extended some of these results and recast them).

(1) (proven earlier by Cantero et al. [8]; see also [35]). If $\lambda, \mu \in \partial\mathbb{D}$ are different, then the zeros of $\Phi_{n+1}(z; \lambda)$ and $\Phi_{n+1}(z; \mu)$ strictly interlace.

(2) (proven independently by Wong [75]) If $\lambda \in \partial\mathbb{D}$, then the zeros of the first and second kind POPUC, $\Phi_{n+1}(z; \lambda)$ and $\Psi_{n+1}(z; \lambda)$, strictly interlace (we remark that here we define $\Psi_{n+1}(z; \lambda)$ to have all the opposite sign Verblunsky coefficients to $\Phi_{n+1}(z; \lambda)$, including changing $\alpha_n = \lambda$ to $\alpha_n = -\lambda$, which [63,75] call $\Psi_{n+1}(z; -\lambda)$).

(3) (special case appeared earlier in Golinskii [34] and a refined version in [12]). If $\lambda, \mu \in \partial\mathbb{D}$, perhaps equal, the zeros of the POPUC $\Phi_{n+1}(z; \lambda)$ and $\Phi_n(z; \mu)$ are either all distinct in which case $\Phi_n(z; \mu)$ has exactly one zero in exactly n of the $n + 1$ intervals obtained by removing the zeros of $\Phi_{n+1}(z; \lambda)$ from $\partial\mathbb{D}$ or else they have exactly one zero in common in which case $\Phi_n(z; \mu)$ has no zero in the two intervals obtained by removing the zeros of $\Phi_{n+1}(z; \lambda)$ from $\partial\mathbb{D}$ closest to the common zero and exactly one zero in the other $n - 1$ intervals.

Wendroff type theorems mean suitable unique converses and, in this regard, parameter counting is important. The various parameters lie in real manifolds and if there is to be existence and uniqueness, the two manifolds must have the same real dimension.

In this regard (1) is fine. The set of possible zeros has real dimension $2n + 2$ and they are determined by n complex Verblunsky coefficients $\{\alpha_j\}_{j=0}^{n-1}$, λ and μ , also $2n + 2$ real parameters. Indeed Theorem 8 is such a Wendroff theorem which we noted is essentially identical to a theorem of Gau–Wu [27] which we called Theorem C. It was later proven in a simpler way by Daepf et al. [17]. The Daepf et al. proof is essentially a POPUC analog of a standard proof of Wendroff for OPRL. Our proof (below) is new, albeit close to the earlier OPUC proof of Golinskii–Kudryavtsev [35], and we feel illuminating.

As stated, (2) doesn't have the right parameter counting for a Wendroff theorem. The zeros are again $2n + 2$ parameters but n complex Verblunsky coefficients $\{\alpha_j\}_{j=0}^{n-1}$ and λ are only $2n + 1$ real parameters. That's because there is a restriction on the zeros, namely (3.6). With this restriction, the set of zeros is only $2n + 1$ parameters so parameter counting is fine and there is a Wendroff theorem, Theorem 9, new here.

The parameter counting is also wrong for there to be a strict converse in case (3). There are $2n + 1$ real parameters for the zeros but n complex Verblunsky coefficients $\{\alpha_j\}_{j=0}^{n-1}$, λ and μ , are $2n + 2$ real parameters. Indeed some simple examples when $n = 1$ show that, in general, there exists a one parameter family of choices for $(\alpha_0, \alpha_1, \lambda, \mu)$ leading to zeros (w_1, w_2, y_1) . Nevertheless, Castillo et al. [11] do have a Wendroff type theorem in the context of (3)! They consider sequences of monic polynomials $\{\Xi_j\}_{j=0}^\infty$ obeying a three term recurrence relation

$$\Xi_{n+1}(z) = (z + \beta_n)\Xi_n(z) - \gamma_n\Xi_{n-1}(z) \tag{10.1}$$

with $n = 1, 2, \dots$ with $\Xi_0 = 1$ and $\Xi_1(z) = z + \beta_0$ for suitable $\beta_j \in \partial\mathbb{D}, j \geq 0$ and $\gamma_j \in \mathbb{C} \setminus \{0\}, j \geq 1$. [11] show that, under suitable hypotheses, there exist Verblunsky coefficients $\{\alpha_j\}_{j=0}^\infty \in \mathbb{D}^\infty$ and $\{\lambda_j\}_{j=0}^\infty \in \partial\mathbb{D}^\infty$ so that Ξ_j are the associated POPUC but, when they exist, $\{\alpha_j\}_{j=0}^\infty \in \mathbb{D}^\infty$ and $\{\lambda_j\}_{j=0}^\infty \in \partial\mathbb{D}^\infty$ are not unique but depend on an arbitrary choice of $\lambda_0 \in \partial\mathbb{D}$. They proved that given sets of zeros as in case (3), there is a unique choice of $\{\beta_j\}_{j=0}^n$ and $\{\gamma_j\}_{j=0}^n$ and so $\{\Xi_j\}_{j=0}^{n-1}$ for which Ξ_n and Ξ_{n+1} have the required zeros. Thus there is a kind of Wendroff theorem for the sequence of POPUC but the OPUC are not determined. In contradistinction, Theorem 8 does determine all the lower order OPUC from the two POPUC which is why we call it Wendroff’s Theorem for (P)OPUC.

We now turn to the proof Theorem 8. We begin with uniqueness. Since

$$\Phi_{n+1}(z, \lambda) = z\Phi_n(z) - \bar{\lambda}\Phi_n^*(z); \quad \Phi_{n+1}(z, \mu) = z\Phi_n(z) - \bar{\mu}\Phi_n^*(z) \tag{10.2}$$

we have that

$$\Phi_n(z) = \frac{\lambda\Phi_{n+1}(z; \lambda) - \mu\Phi_{n+1}(z; \mu)}{(\lambda - \mu)z} \tag{10.3}$$

If $z_1, z_2, \dots, z_{n+1}, w_1, w_2, \dots, w_{n+1}$ are all in $\partial\mathbb{D}$ and are the zeros of $\Phi_{n+1}(z, \lambda)$ and $\Phi_{n+1}(z, \mu)$, then

$$\Phi_{n+1}(z; \lambda) = \prod_{j=1}^{n+1} (z - z_j); \quad \Phi_{n+1}(z; \mu) = \prod_{j=1}^{n+1} (z - w_j) \tag{10.4}$$

so by (10.3), the z ’s and w ’s determine $\Phi_n(z)$ and so, by Wendroff’s Theorem for OPUC (Theorem 2.3), they determine $\{\alpha_j\}_{j=0}^{n-1}$. By (2.22), the zeros also determine λ and μ . Thus, we have the uniqueness claim.

For existence, we fix $2n + 2$ points $z_1, z_2, \dots, z_{n+1}, w_1, w_2, \dots, w_{n+1}$ all in $\partial\mathbb{D}$. For now, they need not be different or in any order but we do need λ and μ below to be unequal.

Define

$$Q_{n+1}(z) = \prod_{j=1}^{n+1} (z - z_j); \quad R_{n+1}(z) = \prod_{j=1}^{n+1} (z - w_j) \tag{10.5}$$

$$\lambda = - \prod_{j=1}^{n+1} (-\bar{z}_j) \quad \mu = - \prod_{j=1}^{n+1} (-\bar{w}_j) \tag{10.6}$$

so that

$$Q_{n+1}(0) = -\bar{\lambda}; \quad R_{n+1}(0) = -\bar{\mu} \tag{10.7}$$

We revert to τ^n notation for the * map since we'll be using it with different implicit n 's. We claim that

$$\tau_{n+1}(Q_{n+1}) = -\lambda Q_{n+1}; \quad \tau_{n+1}(R_{n+1}) = -\mu R_{n+1} \tag{10.8}$$

One way of seeing this is to note both sides have the same zeros, are polynomials of degree $n + 1$ and have the same leading coefficients. Another way is to note that

$$z \left(\frac{1}{z} - \bar{z}_j \right) = 1 - z\bar{z}_j = -\bar{z}_j(z - z_j)$$

Define now

$$P_n(z) = \frac{\lambda Q_{n+1}(z) - \mu R_{n+1}(z)}{(\lambda - \mu)z} \tag{10.9}$$

Of course, this is motivated by (10.3). We note that by (10.7) the numerator vanishes at $z = 0$ so P_n is a polynomial of degree n and the $\lambda - \mu$ in the denominator makes it a monic polynomial. We compute:

$$\begin{aligned} \tau_n(P_n) &= \tau_{n+1}(zP_n) = \frac{\bar{\lambda}(-\lambda)Q_{n+1} - \bar{\mu}(-\mu)R_{n+1}}{\bar{\lambda} - \bar{\mu}} \\ &= \mu\lambda \frac{Q_{n+1} - R_{n+1}}{\lambda - \mu} \end{aligned} \tag{10.10}$$

where we used

$$\frac{1}{\bar{\lambda} - \bar{\mu}} = \frac{1}{\frac{1}{\lambda} - \frac{1}{\mu}} = -\frac{\lambda - \mu}{\mu\lambda}$$

This implies, by direct computation from (10.9) and (10.10), that

Lemma 10.1. *We have that*

$$zP_n - \bar{\lambda}\tau_n(P_n) = Q_{n+1}; \quad zP_n - \bar{\mu}\tau_n(P_n) = R_{n+1} \tag{10.11}$$

Lemma 10.2. *Let $z_0 \in \partial\mathbb{D}$. Then*

$$P_n(z_0) = 0 \iff Q_{n+1}(z_0) = R_{n+1}(z_0) = 0 \tag{10.12}$$

Proof. Since $z_0 \in \partial\mathbb{D}$, one has that $P_n(z_0) = 0 \iff \tau_n(P_n)(z_0) = 0$. Given this, (10.12) is immediate from (10.9), (10.10) and (10.11). \square

With this algebra under our belt

Proof of Theorem 8. That the zero’s interlace follows from the fact that $B_n(z) = \Phi_n(z)/\Phi_n^*$ has a strictly increasing argument. Uniqueness is proven above.

Finally, given interlacing z ’s and w ’s, let \tilde{z}_j be the roots of $z^{n+1} - \bar{\lambda}$ and \tilde{w}_j be the roots of $z^{n+1} - \bar{\mu}$. The P for this set of roots is clearly z^n which has all its roots in \mathbb{D} . One can continuously deform these sets keeping the corresponding λ and μ unchanged through strictly interlacing sets and so deform from z^n to P_n . The zeros of P_n move continuously and, by Lemma 10.2 and strict interlacing, never go through $\partial\mathbb{D}$. Thus P_n has all its zeros inside \mathbb{D} so, by Wendroff’s theorem for OPUC (Theorem 2.3), it is a Φ_n . By Lemma 10.1, Q_{n+1} and R_{n+1} are the POPUC for the Φ_n . \square

Theorem 8 appears already in Golinskii–Kudryavtsev [35, Theorem 3.2]. Our proof of uniqueness is essentially the same as theirs. Our proof of existence is related to theirs but instead of our perturbation argument above, they use a Theorem of Simon [61, Theorem 11.5.6]. We believe that the role of interlacing is clearer in our new argument than in this earlier argument of Simon.

We now turn to the proof of Theorem 9. We will need (8.7) which expresses the Carathéodory function, $G(z)$, of the measure we called $d\mu^\lambda$ in terms of $\Phi_{n+1}(z; \lambda)$ and $\Psi_{n+1}(z; \lambda)$. The Verblunsky coefficients of that measure are precisely what the second half of the Theorem claims are determined by the zeros obeying (3.6). The proof will depend on an analysis of what we’ll call “quasi-Carathéodory” functions (because the c ’s below may not be positive). A degree $n + 1$ quasi-Carathéodory function is one of the form, given $\{z_j\}_{j=1}^{n+1} \subset \partial\mathbb{D}$:

$$f(z) = \sum_{j=1}^{n+1} c_j \frac{z_j + z}{z_j - z}; \quad c_j \in \mathbb{R} \setminus \{0\}; \quad \sum_{j=1}^{n+1} c_j = 1 \tag{10.13}$$

Proposition 10.3. *Let f be a quasi-Carathéodory function (10.13) whose poles are written in cyclic order with $z_{n+2} \equiv z_1$, $c_{n+2} \equiv c_1$. Then*

- (a) $f(z)$ is purely imaginary on the unit circle with the z_j ’s removed.

(b) $f(0) = 1; f(\infty) = -1$

(c) For each $j = 1, \dots, n+1$, if c_j and c_{j+1} have the same (resp. different) signs, then f has an odd (resp. even) number of zeros (counting multiplicity) in the arc between z_j and z_{j+1} .

Remark. If all c_j 's are of the same sign, by the pigeonhole principle, all the zeros of f lie on $\partial\mathbb{D}$ and there is a single zero in each arc between z_j and z_{j+1} . However, if we allow for a sign changes, it is easy to construct examples where not all the zeros of f lie on $\partial\mathbb{D}$ and examples where all of them do.

Proof. (a) and (b) are immediate. (c) By noting that $\text{Im}((1+z)/(1-z)) = 2\text{Im}(z)/|1-z|^2$, one sees that if $c_j > 0$, then $f \rightarrow i\infty$ as z approaches z_j from larger argument. This implies that if c_j and c_{j+1} has opposite (resp. same) signs, there must be an even (resp. odd) number of sign changes of $\text{Im}(f)$ in the arc between z_j and z_{j+1} . \square

Proposition 10.4. Let $\{z_j\}_{j=1}^{n+1}$ and $\{w_j\}_{j=1}^{n+1}$ be two sets of points, all distinct, on the unit circle for which (3.6) holds. Let

$$f(z) = \frac{\prod_{j=1}^{n+1}(1 - \bar{w}_j z)}{\prod_{j=1}^{n+1}(1 - \bar{z}_j z)} \tag{10.14}$$

Then f is a quasi-Carathéodory function.

Proof. We begin by noting that if $|u| = 1$, then

$$\overline{1 - \bar{u}(1/\bar{z})} = (z - u)z^{-1} = -u(1 - \bar{u}z)z^{-1}$$

which implies that

$$f\left(\frac{1}{\bar{z}}\right) = \frac{\prod_{j=1}^{n+1} w_j}{\prod_{j=1}^{n+1} z_j} f(z) = -f(z) \tag{10.15}$$

by (3.6).

Thus f is imaginary on the circle away from the poles which implies, by a partial fraction expansion that

$$f(z) = c_0 + \sum_{j=1}^{n+1} c_j \frac{z_j + z}{z_j - z}; \quad c_j \in \mathbb{R} \setminus \{0\}; c_0 \in i\mathbb{R}$$

Since $f(0) = 1$, we see that $c_0 + \sum_{j=1}^{n+1} c_j = 1$ which is real and so implies that $c_0 = 0$. Thus f is a quasi-Carathéodory function. \square

Proof of Theorem 9. If z_j and w_j are the zero of a POPUC and its associated second kind POPUC, then $\prod_{j=1}^{n+1} z_j = (-1)^n \bar{\lambda}$ while $\prod_{j=1}^{n+1} w_j = (-1)^n (-\bar{\lambda})$ proving (3.6). By (8.7) and Proposition 10.3(c), there are an odd number of zeros in each of the $n + 1$ intervals, so one per interval and the zeros interlace. That proves the direct result.

For the converse, given interlacing zeros obeying (3.6), by Proposition 10.4, f given by (10.14) is a quasi-Carathéodory function. Since there are an odd number of zeros in each interval, the signs of the c_j 's are all the same. Since they sum to 1, they are all positive, so f is the Carathéodory function of a point measure whose POPUC has zeros at the poles of f , so the z_j 's. Moreover, the second kind POPUC has zeros at the zeros of f , so the w 's. \square

11. Derivatives of POPUC

There is a classical connection relating zeros of a polynomial, P , and zeros of its derivative, P' to sums of the form (1.1) with all m_j 's equal because of the following basic relation.

Proposition 11.1. *Let $P(z)$ be a complex polynomial of degree k with distinct zeros $\{z_j\}_{j=1}^k$. Then the zeros of $P'(z)$ are the same as the zeros of*

$$M(z) \equiv \frac{1}{k} \sum_{j=1}^k \frac{1}{z - z_j} \quad (11.1)$$

Proof. Using $P(z) = a \prod_{j=1}^k (z - z_j)$ and taking logarithmic derivatives, one sees that

$$M(z) = \frac{P'(z)}{kP(z)} \quad (11.2)$$

Since the zeros of P are distinct, the zeros of M are precisely the zeros of $P'(z)$. \square

Remark. Of course, (11.2) holds if some of the zeros of P are not distinct, at least away from zeros of P . But at multiple zeros of P , M has a pole even though P' vanishes.

Given the relation of sums like (11.1) and Poncelet type polygons that we studied in Section 7 and what Theorem 7.2 tells us when there are equal m 's, we are interested in polygons tangent at midpoints. In this regard, there are two classical results in the case of triangles

Theorem F (Steiner [67], 1829). *Given a triangle, T in the complex plane with vertices $\{z_j\}_{j=1}^3$, there is a unique ellipse tangent to T at the midpoints of its sides and the foci of the ellipse are*

$$\frac{1}{3}(z_1 + z_2 + z_3) \pm \sqrt{\left(\frac{1}{3}(z_1 + z_2 + z_3)\right)^2 - \frac{1}{3}(z_1 z_2 + z_1 z_3 + z_2 z_3)} \tag{11.3}$$

Theorem G (Siebeck [57], 1864). *The foci of the Steiner ellipse are the zeros of P' where P is the cubic polynomial with zeros at $\{z_j\}_{j=1}^3$*

Remark. Of course, since $P(z) = (z - z_1)(z - z_2)(z - z_3)$, we have $P'(z) = az^2 + bz + c$; $a = 3, b = -2(z_1 + z_2 + z_3), c = (z_1 z_2 + z_1 z_3 + z_2 z_3)$ so this result is immediate. It was Siebeck who realized the special role of sums like (11.1) in studying the zeros of P' .

There are two papers in the series discussed in Section 1 that deal with extending these results to k zeros on $\partial\mathbb{D}$ (for the case of three zeros, they always lie on a circle so there is no loss in supposing that $z_1, z_2, z_3 \in \partial\mathbb{D}$), namely Gau–Wu [25] and Gorkin–Shubak [37]. Even though all the results later in this section are more or less in these papers, we decided to include this subject for several reasons: first, we believe the results are interesting and deserve to be better known; second, we want to emphasize how simple the basic proofs are; finally, and most importantly, while Gorkin–Shubak [37] state the basic equation for determining when n points in \mathbb{D} are the zeros of a derivative of a polynomial, P , whose zeros lie in $\partial\mathbb{D}$ (Theorem 11.3), they don’t analyze them nor note the different nature in the case n is odd vs. n even (Theorem 11.4 below is new).

The following is a variant of the main result in Gau–Wu [25]. Since the numerical range of an operator A in S_2 is an ellipse (indeed, all operators on \mathbb{C}^2 have ellipses as their numerical range), the $n = 2$ case of this theorem provides a proof of Theorems F and G.

Theorem 11.2. *Let $\{z_j\}_{j=1}^{n+1}$ be $n+1$ distinct points in $\partial\mathbb{D}$. Then up to unitary equivalence, there is a unique operator $A \in S_n$ so that $N(A)$ is tangent to the midpoints of the edges of the convex polygon with vertices $\{z_j\}_{j=1}^{n+1}$. If $P(z) = \prod_{j=1}^{n+1}(z - z_j)$, then $(n + 1)^{-1}P'(z)$ is the characteristic polynomial for A and, in particular, the zeros of P' are the eigenvalues of A .*

Proof. By Theorem 7.2, there is a unique operator $A \in S_n$ with $N(A)$ tangent at the midpoints of the polygon and its M –function has the form (11.1) with $k = n + 1$. By Theorem 6 and (11.2), we have that

$$\frac{P'(z)}{P(z)} = \frac{\Phi_n(z)}{z\Phi_n(z) - \lambda\Phi_n^*(z)} \tag{11.4}$$

so $P'(z) = \Phi_n(z)$. By Theorems 2.10 and 2, Φ_n is the characteristic polynomial of A . \square

Gorkin–Shubak [37] study the question of when $\{a_j\}_{j=1}^n$ in \mathbb{D} are the zeros of the derivative of a polynomial, P , all of whose zeros lie in $\partial\mathbb{D}$. While they focus on the case $n = 2$, they state the following (in a version that doesn’t mention POPUC or OPUC!).

Theorem 11.3. Let $\{a_j\}_{j=1}^n$ be n not necessarily distinct points in \mathbb{D} . Then there is a polynomial, P , with all of its zeros on $\partial\mathbb{D}$ so that P' has its zeros at $\{a_j\}_{j=1}^n$ if and only if there is a $\lambda \in \partial\mathbb{D}$ so that

$$\Phi'_{n+1}(z; \lambda) = (n + 1)\Phi_n(z) \tag{11.5}$$

where $\Phi_n(z) = \prod_{j=1}^n (z - a_j)$ and $\Phi_{n+1}(z; \lambda)$ is given by (2.19).

Proof. If (11.5) holds, then, clearly, $\{a_j\}_{j=1}^n$ is the set of critical points of a polynomial whose zeros all lie on $\partial\mathbb{D}$.

Conversely, let P be a monic polynomial whose zeros $\{z_j\}_{j=1}^{n+1}$ all lie in $\partial\mathbb{D}$ and so that P' vanishes at $\{a_j\}_{j=1}^n$. By Theorem 6, we know that $M(z)$ given by (11.1) (with $k = n + 1$) is the M -function associated to some $n + 1$ -point measure $d\mu$ on $\partial\mathbb{D}$ with OPUC obeying (3.4) with $d\nu = |\varphi_n|^2 d\mu = (n + 1)^{-1} \sum_{j=1}^{n+1} \delta_{z_j}$. Moreover, for some λ , z_j are the zeros of $\Phi_{n+1}(z; \lambda)$. Thus $P(z) = \Phi_{n+1}(z; \lambda)$ and, by (3.4) and (11.2),

$$\frac{1}{n + 1} \frac{P'(z)}{\Phi_{n+1}(z; \lambda)} = \frac{\Phi_n(z)}{\Phi_{n+1}(z; \lambda)} \tag{11.6}$$

so (11.5) holds. \square

Now, given $\{a_j\}_{j=1}^n$, a collection of points in \mathbb{D} , let $s_k = \sum a_{i_1} a_{i_2} \dots a_{i_k}$, with the sum over all $\binom{n}{k}$ sets of distinct $i_1 i_2 \dots i_k$ in $\{1, \dots, n\}$, be the elementary symmetric functions, $k = 1, \dots, n$ so that

$$\Phi_n(z) \equiv \prod_{j=1}^n (z - a_j) = z^n + \sum_{k=0}^{n-1} (-1)^{n-k} s_{n-k} z^k \tag{11.7}$$

Theorem 11.4. (11.5) is equivalent to the equations

$$(n - j)s_{n-j} + (-1)^{n-1} \bar{\lambda}(j + 1)\bar{s}_{j+1} = 0, \quad j = 0, \dots, n - 1 \tag{11.8}$$

For $n = 2k$, even, there are k independent equations and in the $2n$ real dimensional manifold of $(a_1, \dots, a_n) \in \mathbb{D}^n$, the set obeying (11.5) has real dimension $n + 1$. For $n = 2k + 1$, odd, there are $k + 1$ equations and if $s_{k+1} \neq 0$, we have that

$$\lambda = (-1)^{n-1} \bar{s}_k / s_k \tag{11.9}$$

In the $2n$ real dimensional manifold of $(a_1, \dots, a_n) \in \mathbb{D}^n$, the set obeying (11.5) has real dimension $n + 1$.

Remarks. 1. It may be that the set of allowed a 's has some singular points but except for a lower dimensional set, it is a real analytic manifold.

2. The case $n = 2$ is treated in detail in Gorkin–Shubak [37]. In that case, there is one equation where λ can be used to adjust the phase so that we get

$$|a_1 + a_2| = |a_1 a_2| \tag{11.10}$$

3. In terms of the operator, $A \in S_n$ associated to Φ_n , the a_j are eigenvalues and $s_k = \text{Tr}(\wedge^k(A))$ (see Simon [62] for discussion of $\wedge^k(A)$). In particular, (11.8) for $j = 0$ is

$$n \det(a) + (-1)^{n-1} \bar{\lambda} \text{Tr}(A^*) = 0 \tag{11.11}$$

4. The proof shows that the only possible singularities occur at points where two a 's are equal and, if $n = 2k + 1$ is odd, where $s_{k+1} = 0$. It is an interesting question whether these singularities actually occur.

Proof. (11.5) is equivalent to

$$n\Phi_n + \bar{\lambda}(\Phi_n^*)' - z\Phi_n = 0 \tag{11.12}$$

Write $\Phi_n(z) = z^n + \sum_{j=0}^{n-1} A_j z^j$. Then $(\Phi_n^*)'(z) = \sum_{k=1}^n k A_{n-k} z^k$ and (11.12) becomes (11.8).

It is well known that away from coincident points, the map $(a_1, \dots, a_n) \mapsto (s_1, \dots, s_n)$ has a non-singular inverse. The equations (11.9) are the same for j and for $n - 1 - j$ but are clearly independent linear equations for $j = 1, 2, \dots, \lfloor \frac{n}{2} \rfloor + 1$. The dimension counting results follow. \square

12. Poncelet's theorem

That completes most of we want to say about using OPUC ideas to understand the topics discussed in Section 1. These final two sections make some brief comments about the relation of algebraic geometric ideas in the area where we have not succeeded in leveraging OPUC methods.

We begin with Poncelet's Theorem. For the triangle case, the ideas of Section 7 are ideal. Given two points $z_1, z_2 \in \mathbb{D}$, there is a unique monic polynomial, $\Phi_2(z) = (z - z_1)(z - z_2)$ with simple zeros at those two points (assuming $z_1 \neq z_2$). As with all 2×2 matrices, the numerical range of the associated compressed multiplication operator, A , is an ellipse with foci at the eigenvalues z_1, z_2 . We can form the unitaries associated to the various POPUC, $\Phi_3(z; \lambda)$ as λ runs through $\partial\mathbb{D}$. Their numerical ranges provide an infinity of Poncelet triangles with vertices on the unit circle and tangent to $N(A)$.

Moreover, if the ellipse is made larger (with the same foci), it is easy to see that starting at $w_0 \in \partial\mathbb{D}$ and forming three successive tangents to the large ellipse ending at successive points w_1, w_2, w_3 don't go as far around the circle as for the critical ellipse and so one gets less than a closed triangle. Similarly, for a smaller ellipse, the w_j are further

along and so don't give a closed triangle. Thus we only have a triangle for the critical ellipse which has an infinity of triangles. This proves Poncelet's Theorem for triangles.

As we emphasized, the ideas of Section 7 give an analog of Poncelet's theorem from triangles to other polygons, but not the analog that Poncelet considered, since in that case, in general, the numerical range is not an ellipse. That said, Mirman [50] and Gau–Wu [26] did use some of these ideas for a proof of the Poncelet ellipse theorem for higher degree polygons but they rely on a result of Kippenhahn [47] who considers the dual of the boundary, $\partial N(A)$, of the numerical range of an $n \times n$ matrix. He considers the projective dual of this boundary, essentially, the set of tangents to the curve and proves that in projective coordinates, this dual is a degree n real variety (or more properly the outer part of the variety).

The above argument for triangles extends to n -gons and shows that if we fix the foci of an ellipse in \mathbb{D} and $w_0 \in \partial\mathbb{D}$, then there is a single eccentricity where the n touching tangents starting at w_0 lead to a closed convex n -gon. This is perhaps best understood using the idea of a billiard trajectory as in [22, Chapter 15]. Indeed, given $w_0 \in \partial\mathbb{D}$ and an ellipse $E \subset \mathbb{D}$, there are two tangents to E that pass through w_0 . Let ℓ be the tangent that intersects $\partial\mathbb{D}$ at w_0 and a point w_1 such that the path from w_0 to w_1 on $\partial\mathbb{D}$ in the counterclockwise direction is along the arc that is separated from E by ℓ . Using this notation, let us define $\mathcal{B}(w_0) = w_1$ and similarly define \mathcal{B} on all of $\partial\mathbb{D}$. We want to show that if the foci of an ellipse are given in \mathbb{D} , then there is a unique eccentricity of the ellipse E with those foci such that the corresponding map \mathcal{B} satisfies $\mathcal{B}^n(w_0) = w_0$. This last condition can be rephrased as

$$\sum_{j=0}^{n-1} (\arg(\mathcal{B}^{j+1}(w_0)) - \arg(\mathcal{B}^j(w_0))) = 2\pi \quad (12.1)$$

where each difference of arguments in this sum is taken in $(0, 2\pi)$. Our above argument for triangles proves monotonicity in the left-hand side of (12.1) as a function of the eccentricity. We also saw above that there is a unique eccentricity e such that if the eccentricity of E is e , then $\mathcal{B}^3(w_0) = w_0$ and Poncelet's Theorem implies that this value of e is independent of $w_0 \in \partial\mathbb{D}$. Monotonicity implies that if the eccentricity of E is smaller than e , then the left-hand side of (12.1) is strictly larger than 2π . Similarly, as the ellipse approaches touching the unit circle, it takes more than n iterations of \mathcal{B} to get past the point on E closest to $\partial\mathbb{D}$, so the left-hand side of (12.1) converges to something less than 2π . By continuity there is a unique eccentricity where the left-hand side of (12.1) is 2π . A priori, that eccentricity could be w_0 -dependent.

One can show there is an $A \in S_n$ whose numerical range is tangent to this n -gon at the same tangent points. Given Kippenhahn's result, the n fold agreement of the ellipse and $\partial N(A)$ and a use of Bezout's theorem proves that the two curves are the same, proving Poncelet's theorem in this case.

It would be interesting to have an OPUC understanding of Kippenhahn's Theorem.

13. From the numerical range to the eigenvalues

In this final section, we want to discuss how for $A \in S_n$, $N(A)$ determines the eigenvalues of A (and so, up to unitary equivalence, A). The results in Sections 1 and 3 provide two ways that we want to discuss first. Then we'll turn to a potential method via algebraic geometry.

Here are the two methods:

(1) Given n and $\partial N(A)$, draw two n gons with vertices in $\partial \mathbb{D}$ that circumscribe $N(A)$ and apply Theorem C/Theorem 8.

(2) Given n and $\partial N(A)$, draw one such polygon. Use Theorem 4 to compute the m_j and then Theorem B/Theorem 6 to determine the zeros of Φ_n which are the eigenvalues.

There is another connection between eigenvalues and a curve related to $N(A)$ that is related to the work of Kippenhahn [47] mentioned earlier and extended by Singer [66] and Langer–Singer [48]. There is a real projective curve, Γ , so that $N(A)$ is the convex hull of Γ and the eigenvalues are exactly the (real) foci of Γ (foci are an involved construction from algebraic geometry that, for an ellipse, are the usual foci).

When we first learned of this result, we assumed that it must imply for the case of elliptical $N(A)$, that the eigenvalues must be its two usual foci, obviously not both simple if $n \geq 3$. This was wrong! Indeed, there has been a detailed study of ellipses in the case $n = 4$ by Fujimura [23] and Gorkin–Wagner [38] that shows in this case there is a third usually distinct eigenvalue.

The key point is that while $\partial N(A)$ is included in Γ , Γ can have additional pieces inside $N(A)$ (and if $n \geq 3$, there are always such additional pieces). As we've seen, $\partial N(A)$ is the envelope of the lines obtained by joining successive eigenvalues of the various rank one unitary dilations, U_λ . As proven by Singer [66, Prop. 3.1], Γ is the envelope of the complete graph on the eigenvalues of U_λ , i.e. all lines joining pairs of distinct eigenvalues. We illustrated this with a simple illuminating example. Once one picks the eigenvalues $a, b, c \in \mathbb{D}$ for $A \in S_3$, one can form the various $\Phi_4(z; \lambda)$ and the polygons associated to their zeros. We took examples with $b = 0.8e^{34i}$ and $c = 0.57e^{4i}$ (chosen by experimenting to get clean output). For one example, we picked $a = 0.7i$ and for the second we chose $a = -0.74949 \dots + i0.164697 \dots$. The later was chosen using [38, Prop. 3.7] to give an elliptical $N(A)$.

Fig. 1 shows the outer polygons when $a = 0.7i$ (left) or for the elliptical case (right).

In conformance with the discussion in Section 7, one can see the convex $\partial N(A)$ formed in each case. Fig. 2 shows the complete graphs of the eigenvalues of the U_λ when $a = 0.7i$ (left) and in the elliptical case (right).

One sees that there is an extra piece of Γ in the resulting envelope.

As shown in [38], in the elliptical case, the extra piece of Γ is the single point $\{a\}$. As can be seen by looking at the case where all eigenvalues of A are zero and the extra pieces are circles when $n \geq 4$, this single point phenomena is only true for $n = 3$. In any event, one cannot use these ideas to go directly from $N(A)$ as a set to the eigenvalues.

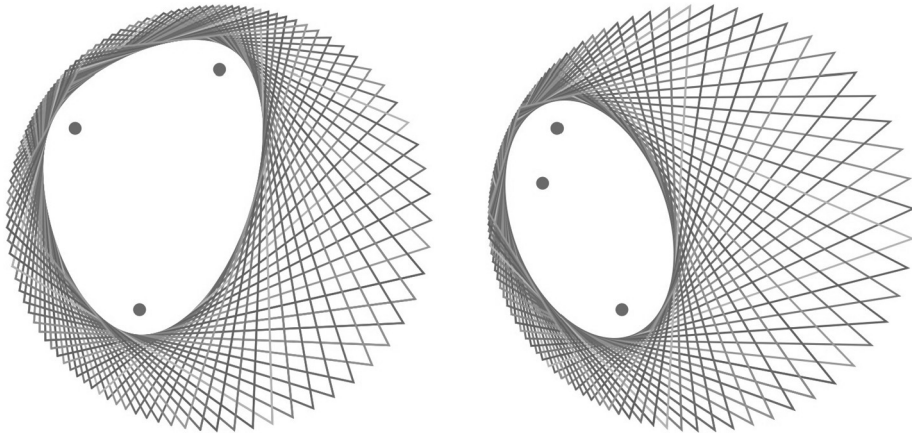


Fig. 1. The outer polygons and the numerical range $N(A)$ for $A \in S_3$ determined by its two eigenvalues at $0.8e^{34i}$ and $0.57e^{4i}$, and the third one either at $0.7i$ (left) or at $-0.74949 \dots + i0.164697 \dots$ (right). The eigenvalues are indicated by the fat dots in the interior of $N(A)$.

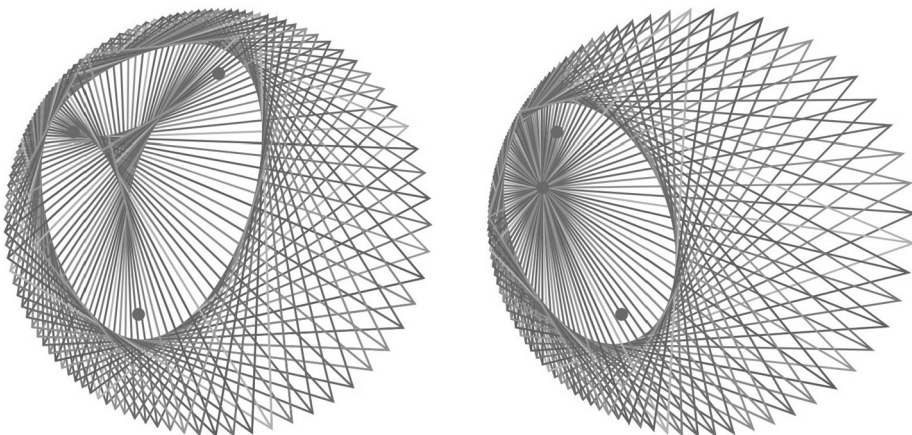


Fig. 2. The complete graphs of the eigenvalues of the U_λ and the numerical range $N(A)$ of two different $A \in S_3$ from Fig. 1.

References

- [1] Y. Arlinskii, Conservative discrete time-invariant systems and block operator CMV matrices, *Methods Funct. Anal. Topology* 15 (2009) 201–236.
- [2] Y. Arlinskii, L. Golinskii, E. Tsekanovskii, Contractions with rank one defect operators and truncated CMV matrices, *J. Funct. Anal.* 254 (2008) 154–195.
- [3] Y. Arlinskii, Schur parameters, Toeplitz matrices, and Kren shorted operators, *Integral Equations Operator Theory* 71 (2011) 417–453.
- [4] Y. Arlinskii, The Schur problem and block operator CMV matrices, *Complex Anal. Oper. Theory* 8 (2014) 875–923.
- [5] F.F. Bonsall, J. Duncan, *Numerical Ranges of Operators on Normed Spaces and of Elements of Normed Algebras*, Cambridge University Press, 1971.
- [6] J. Breuer, B. Simon, O. Zeitouni, Large deviations and the Lukic conjecture, *Duke Math. J.* 167 (2018) 2857–2902.

- [7] M.J. Cantero, L. Moral, L. Velázquez, Measures and para-orthogonal polynomials on the unit circle, *East J. Approx.* 8 (2002) 447–464.
- [8] 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.
- [9] M.J. Cantero, L. Moral, L. Velázquez, Measures on the unit circle and unitary truncations of unitary operators, *J. Approx. Theory* 139 (2006) 430–468.
- [10] K. Castillo, F. Marcellán, M.N. Rebocho, Zeros of para-orthogonal polynomials and linear spectral transformations on the unit circle, *Numer. Algorithms* 71 (2016) 699–714.
- [11] K. Castillo, R. Cruz-Barroso, F. Perdomo-Pío, On a spectral theorem in paraorthogonality theory, *Pacific J. Math.* 280 (2016) 327–347.
- [12] K. Castillo, J. Petronilho, Refined interlacing properties for zeros of paraorthogonal polynomials on the unit circle, *Proc. Amer. Math. Soc.* 146 (2018) 3285–3294.
- [13] M.D. Choi, C.K. Li, Constrained unitary dilations and numerical ranges, *J. Operator Theory* 46 (2001) 435–447.
- [14] U. Daepf, P. Gorkin, R. Mortini, Ellipses and finite Blaschke products, *Amer. Math. Monthly* 109 (2002) 785–795.
- [15] U. Daepf, P. Gorkin, A. Shaffer, K. Voss, Möbius transformations and Blaschke products: the geometric connection, *Linear Algebra Appl.* 516 (2017) 186–211.
- [16] U. Daepf, P. Gorkin, A. Shaffer, K. Voss, Finding Ellipses: What Blaschke Products, Poncelet’s Theorem, and the Numerical Range Know about Each Other, *Amer Mathematical Society*, 2018.
- [17] U. Daepf, P. Gorkin, K. Voss, Poncelet’s theorem, Sendov’s conjecture, and Blaschke products, *J. Math. Anal. Appl.* 365 (2010) 93–102.
- [18] L. Darius, J. Hernández, F. Marcellán, Spectral transformations for Hermitian Toeplitz matrices, *J. Comput. Appl. Math.* 202 (2007) 155–176.
- [19] P. Delsarte, Y. Genin, Tridiagonal approach to the algebraic environment of Toeplitz matrices II. Zero and eigenvalue problems, *SIAM J. Matrix Anal. Appl.* 12 (1991) 220–238.
- [20] L. Fejér, Über die Lage der Nullstellen von Polynomen, die aus Minimumforderungen gewisser Art entspringen, *Math. Ann.* 85 (1922) 41–48.
- [21] G. Freud, *Orthogonal Polynomials*, Pergamon Press, Oxford-New York, 1971.
- [22] L. Flatto, Poncelet’s Theorem, *American Mathematical Society*, Providence, RI, 2009.
- [23] M. Fujimura, Inscribed ellipses and Blaschke products, *Comput. Methods Funct. Theory* 13 (2013) 557–573.
- [24] H.L. Gau, P.Y. Wu, Numerical range of $S(\phi)$, *Linear Multilinear Algebra* 45 (1998) 49–73.
- [25] H.L. Gau, P.Y. Wu, Lucas’ theorem refined, *Linear Multilinear Algebra* 45 (1999) 359–373.
- [26] H.L. Gau, P.Y. Wu, Numerical range and Poncelet property, *Taiwanese J. Math.* 7 (2003) 173–193.
- [27] H.L. Gau, P.Y. Wu, Numerical range circumscribed by two polygons, *Linear Algebra Appl.* 382 (2004) 155–170.
- [28] H.L. Gau, P.Y. Wu, Numerical range of a normal compression, *Linear Multilinear Algebra* 52 (2004) 195–201.
- [29] H.L. Gau, P.Y. Wu, Numerical range of a normal compression, II, *Linear Algebra Appl.* 390 (2004) 121–136.
- [30] H.L. Gau, P.Y. Wu, Numerical ranges and compressions of S_n -matrices, *Oper. Matrices* 7 (2013) 465–476.
- [31] 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.
- [32] Ya.L. Geronimus, On the trigonometric moment problem, *Ann. of Math.* 47 (1946) 742–761.
- [33] Ya.L. Geronimus, *Orthogonal Polynomials: Estimates, Asymptotic Formulas, and Series of Polynomials Orthogonal on the Unit Circle and on an Interval*, Consultants Bureau, New York, 1961.
- [34] L. Golinskii, Quadrature formula and zeros of para-orthogonal polynomials on the unit circle, *Acta Math. Hungar.* 96 (2002) 169–186.
- [35] L. Golinskii, M. Kudryavtsev, An inverse spectral theory for finite CMV matrices, *Inverse Probl. Imaging* 4 (1) (2010) 93–110.
- [36] P. Gorkin, R.C. Rhoades, Boundary interpolation by finite Blaschke products, *Constr. Approx.* 27 (2008) 75–98.
- [37] P. Gorkin, E. Skubak, Polynomials, ellipses, and matrices: two questions, one answer, *Amer. Math. Monthly* 118 (2011) 522–533.
- [38] P. Gorkin, N. Wagner, Ellipses and compositions of finite Blaschke products, *J. Math. Anal. Appl.* 445 (2017) 1354–1366.
- [39] P. Gorkin, Four theorems with their foci on ellipses, *Amer. Math. Monthly* 126 (2019) 99–111.

- [40] 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.
- [41] K. Gustafson, D.K.M. Rao, *Numerical Range. The Field of Values of Linear Operators and Matrices*, Universitext, Springer-Verlag, New York, 1997.
- [42] P.R. Halmos, Numerical ranges and normal dilations, *Acta Sci. Math.* 25 (1964) 1–5.
- [43] R.A. Horn, C.R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, 1991.
- [44] L. Halbeisen, N. Hungerbühler, A simple proof of Poncelet's theorem (on the occasion of its bicentennial), *Amer. Math. Monthly* 122 (2015) 537–551.
- [45] W.B. Jones, O. Njåstad, W.I. Thron, Moment theory, orthogonal polynomials, quadrature, and continued fractions associated with the unit circle, *Bull. Lond. Math. Soc.* 21 (1989) 113–152.
- [46] S. Khrushchev, Schur's algorithm, orthogonal polynomials, and convergence of Wall's continued fractions in $L^2(\mathbb{T})$, *J. Approx. Theory* 108 (2001) 161–248.
- [47] R. Kippenhahn, Über den Wertevorrat einer Matrix, *Math. Nachr.* 6 (1951) 193–228.
- [48] J.C. Langer, D.A. Singer, Foci and foliations of real algebraic curves, *Milan J. Math.* 75 (2007) 225–271.
- [49] M. Marden, *Geometry of Polynomials*, second edition, *Mathematical Surveys*, vol. 3, American Mathematical Society, Providence, R.I., 1966; first edition 1947.
- [50] B. Mirman, Numerical ranges and Poncelet curves, *Linear Algebra Appl.* 281 (1998) 59–85.
- [51] B. Mirman, UB-matrices and conditions for Poncelet polygon to be closed, *Linear Algebra Appl.* 360 (2003) 123–150.
- [52] B. Mirman, Poncelet's porism in the finite real plane, *Linear Multilinear Algebra* 57 (2009) 439–458.
- [53] B. Mirman, V. Borovikov, L. Ladyzhensky, R. Vinograd, Numerical ranges, Poncelet curves, invariant measures, *Linear Algebra Appl.* 329 (2001) 61–75.
- [54] B. Mirman, P. Shukla, A characterization of complex plane Poncelet curves, *Linear Algebra Appl.* 408 (2005) 86–119.
- [55] J.V. Poncelet, *Traité des propriétés projectives des figures*, Mett-Paris, 1822.
- [56] I. Schur, Über Potenzreihen, die im Innern des Einheitskreises beschränkt sind, I, II, *J. Reine Angew. Math.* 147 (1917) 205–232, *J. Reine Angew. Math.* 148 (1918) 122–145. English translation in: I. Gohberg (Ed.), *I. Schur Methods in Operator Theory and Signal Processing*, in: *Operator Theory: Advances and Applications*, vol. 18, Birkhäuser, Basel, 1986, pp. 31–59, 66–88.
- [57] J. Siebeck, Ueber eine neue analytische Behandlungsweise der Brennpunkte, *J. Reine Angew. Math.* 64 (1864) 175–182.
- [58] B. Simanek, An electrostatic interpretation of the zeros of paraorthogonal polynomials on the unit circle, *SIAM J. Math. Anal.* 48 (2016) 2250–2268.
- [59] B. Simon, OPUC on one foot, *Bull. Amer. Math. Soc.* 42 (2005) 431–460.
- [60] B. Simon, *Orthogonal Polynomials on the Unit Circle, Part 1: Classical Theory*, *AMS Colloquium Series*, American Mathematical Society, Providence, RI, 2005.
- [61] B. Simon, *Orthogonal Polynomials on the Unit Circle, Part 2: Spectral Theory*, *AMS Colloquium Series*, American Mathematical Society, Providence, RI, 2005.
- [62] B. Simon, *Trace Ideals and Their Applications*, second edition, *Mathematical Surveys and Monographs*, vol. 120, American Mathematical Society, 2005.
- [63] B. Simon, Rank one perturbations and the zeros of paraorthogonal polynomials on the unit circle, *J. Math. Anal. Appl.* 329 (2007) 376–382.
- [64] B. Simon, CMV matrices: five years after, *J. Comput. Appl. Math.* 208 (2007) 120–154.
- [65] B. Simon, *A Comprehensive Course in Analysis, Part 4: Operator Theory*, American Mathematical Society, Providence, RI, 2015.
- [66] D.A. Singer, The location of critical points of finite Blaschke products, *Conform. Geom. Dyn.* 10 (2006) 117–124.
- [67] J. Steiner, Géométrie pure. Développement d'une série de théorèmes relatifs aux sections coniques, *Ann. Math. Pures Appl.* 19 (1828/1829) 37–64.
- [68] G. Szegő, Beiträge zur Theorie der Toeplitzschen Formen, I, II, *Math. Z.* 6 (1920) 167–202; *Math. Z.* 9 (1921) 167–190.
- [69] G. Szegő, *Orthogonal Polynomials*, *Amer. Math. Soc. Colloq. Publ.*, vol. 23, American Mathematical Society, Providence, RI, 1939; 3rd edition, 1967.
- [70] B. Sz.-Nagy, Unitary dilations of Hilbert space operators and related topics, in: *Conference Board of the Mathematical Sciences Regional Conference Series in Mathematics*, No. 19, American Mathematical Society, 1974.
- [71] B. Sz.-Nagy, C. Foias, H. Bercovici, L. Kérchy, *Harmonic Analysis of Operators on Hilbert Space*, second edition, Universitext, Springer, New York, 2010.

- [72] A.V. Teplyaev, Properties of polynomials that are orthogonal on the circle with random parameters, *J. Sov. Math.* 61 (1992) 1931–1935; Russian original in: *Problemy Teorii Veroyatnost. Raspred. XI, Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)* 177 (1989) 157–162, 191–192.
- [73] S. Verblunsky, On positive harmonic functions, I, II, *Proc. Lond. Math. Soc.* 38 (1935) 125–157; *Proc. Lond. Math. Soc.* 40 (1936) 290–320.
- [74] B. Wendroff, On orthogonal polynomials, *Proc. Amer. Math. Soc.* 12 (1961) 554–555.
- [75] M–W. Wong, First and second kind paraorthogonal polynomials and their zeros, *J. Approx. Theory* 146 (2007) 282–293.
- [76] P.Y. Wu, Polygons and numerical ranges, *Amer. Math. Monthly* 107 (6) (2000) 528–540.