國科會研究計畫成果自評

吳培元

本年研究計畫係三年期計畫「算子及矩陣數值域之研究」中的第二年。所提的研究內容係對一類算子$S(\phi)$的數值域性質作更深入的探討。在過去這一年來，本人和中央大學的高華隆每週聚會一個下午作腦力激盪，確實獲致不少具體的成果，包括：（一）將$S(\phi)$算子推廣到更一般類型的算子。設$a_1, a_2, \cdots, a_{n+1}$是平面上n+1個不同點且任何都是其他點所形成凸包的頂點。設$x = [x_1, \cdots, x_{n+1}]^t$是$C^{n+1}$中的單位向量，其中$x_j$都不等於零。設K表示$C^{n+1}$中所有和x正交的向量所形成的n維子空間，且A是矩陣$diag(a_1, \cdots, a_{n+1})$在K上的壓縮算子。如果$a_j$都在單位圓上，則A就是$S(\phi)$型的算子。故A可以看成是$S(\phi)$的適當推廣。很多$S(\phi)$數值域的性質也都可以推廣到A的數值域，包括了其龐斯利性質；（二）考慮無窮維的$S(\phi)$，此時$S(\phi)$數值域的性質有部份也都成立。只是這一方面的結果並不完整，還有很大的發展空間。以上兩方面的發展，我們都在這一次的成果報告中有較詳細的描述。該報告也將在2003年6月號的「台灣數學期刊」上印行發表。目前這一方面的研究，高華隆和我有很多進展。這顯然是一個相當豐富的研究領域，結合了古典幾何學和近世算子理論兩個源頭。後續的發展還很有可觀。

數值域與龐斯利性質高華隆、吳培元

摘要:在本篇論文中，我們對一類算子$S(\phi)$的數值域的龐斯利性質的近期研究發展作了一個統覽性的回顧。這可以視爲是本文第二個作者在「美國數學月報」上發表的統覽性文章的一個更及時和更深入的版本。本文中的新增資料有：(1)這個領域中的主要結果(龐斯利定理等之推廣)的一個更簡單的證法，(2)Mirman最近的發現，否定掉一個關於有限維$S(\phi)$的數值域邊界必爲龐斯利曲線的臆測，和(3)上述結果的兩個方向的部分推廣：由酉壓縮到正規壓縮及由有限維$S(\phi)$到無窮維。

Numerical Range and Poncelet Property

Hwa-Long Gau* and **Pei Yuan Wu***

Abstract

In this survey article, we give an expository account of the recent developments on the Poncelet property for numerical ranges of the operators $S(\phi)$. It can be considered as an updated and more advanced edition of the recent expository article published in the American Mathematical Monthly by the second author on this topic. The new information includes: (1) a simplified approach to the main results (generalizations of Poncelet, Brianchon–Ceva and Lucas–Siebeck theorems) in this area, (2) the recent discovery of Mirman refuting a previous conjecture on the coincidence of Poncelet curves and boundaries of the numerical ranges of finite-dimensional $S(\phi)$, and (3) some partial generalizations by the present authors of the above-mentioned results from the unitary-dilation context to the normal-dilation one and also from the finite-dimensional $S(\phi)$ to the infinite-dimensional.

2000 *Mathematics Subject Classification*: Primary 15A60, 47A12; Secondary 14H99.

Key words and phrases: Numerical range, Poncelet property, unitary dilation.

*This work was partially supported by the National Science Council of the Republic of China under research projects NSC-91-2115-M-008-011 and NSC-91-2115-M-009-009 of the respective author.

1. Introduction

In recent years, the research on the numerical ranges of finite matrices and bounded operators has been very active, thanks to the biennial convening of the WONRA (Workshop on Numerical Ranges and Numerical Radii). (For more information on this, check the webpage http://www.resnet.wm.edu/∼cklixx/wonra02.html.) One area of investigations concerns the numerical ranges of the finite-dimensional compressions of the shift. It was discovered that the boundaries of their numerical ranges possess the Poncelet property, meaning that there exist infinitely many polygons with the property that each has all its vertices on the unit circle and all its sides tangent to the asserted boundary. This yields an unexpected link between the twentieth-century subject of numerical range and some nineteen-century gems of projective geometry. An expository account of this development was given in [35], which explains the pertinent results in a historical context. The purpose of this survey is to update this previous account by providing a simplified approach and expounding the recent discoveries. Chief among the latter is the one by Mirman that not every algebraic convex curve in the open unit disc which has the Poncelet property arises as the boundary of the numerical range of the asserted operator, thus refuting a previous conjecture on identifying such numerical ranges by the Poncelet property. We will also elaborate on our recent attempts in generalizing the main results in this area to more general contexts such as general convex polygons instead of polygons with vertices on the unit circle and general compressions of the shift instead of mere the finite-dimensional ones.

In Section 2 below, we start with a brief review of the definition and basic properties of numerical ranges of operators on a Hilbert space. We also discuss the notion of dilation and its connection with numerical ranges. Section 3 then treats numerical ranges of finite matrices. Here the extra tool of Kippenhahn curve proves very useful. It involves the point-line duality of the projective plane. Section 4 considers the compressions of the shift, whose numerical ranges will be the main focus of this paper.

Several different representations of such operators, one analytic and two matricial, are presented, each of which has its merit in exposing certain properties of their numerical ranges. Section 5 gives the main results on the Poncelet property for the numerical ranges of the compressions of the shift on finite-dimensional spaces. There are three of them: generalizations of the Poncelet porism (on the existence of infinitely many interscribing polygons between two ellipses), Brianchon-Ceva theorem (on the condition for the tangent points of an inscribing ellipse of a triangle), and Lucas-Siebeck theorem (on the relation between zeros of a polynomial and its derivative). We then move on to the partial generalizations of these results in Section 6.

2. Numerical Range

Let A be a (bounded linear) operator on a complex Hilbert space H. The *numerical range* of A is the set $W(A) \equiv \{\langle Ax, x\rangle : x \in H, \|x\| = 1\}$ in the complex plane, where $\langle \cdot, \cdot \rangle$ denotes the inner product in H. In other words, $W(A)$ is the image of the unit sphere $\{x \in H : \|x\| = 1\}$ of H under the (bounded) quadratic form $x \mapsto \langle Ax, x\rangle$. Some properties of the numerical range follow easily from the definition. For one thing, the numerical range is unchanged under the unitary equivalence of operators: $W(A) = W(U^*AU)$ for any unitary U. It also behaves nicely under the operation of taking the adjoint of an operator: $W(A^*) = \{\overline{z} : z \in W(A)\}$. More generally, this is even the case when taking the affine transformation: if

$$f(x + iy) = (a_1x + b_1y + c_1) + i(a_2x + b_2y + c_2)$$

is an affine transformation of the complex plane $\mathbb{C}$, where x, y and a_j, b_j and $c_j, j = 1, 2$, are all real and the latter satisfy $a_1b_2 \neq a_2b_1$, and if we define $f(A)$ to be

$$(a_1\text{Re } A + b_1\text{Im } A + c_1I) + i(a_2\text{Re } A + b_2\text{Im } A + c_2I),$$

where $\text{Re } A = (A + A^*)/2$ and $\text{Im } A = (A - A^*)/(2i)$ are the *real* and *imaginary parts* of A, respectively, then $W(f(A)) = f(W(A)) \equiv \{f(z) : z \in W(A)\}$. Thus the numerical range can be considered as an affine property of the operator. In this study of numerical ranges, the reduction through affine transformations is a handy tool in many situations.

The most important property of the numerical range is that $W(A)$ *is always convex.* This is the celebrated Toeplitz–Hausdorff Theorem from 1918-19 [31, 15]. Over the years, there are numerous proofs and generalizations of this fact. The usual proof is to first reduce it to the case of 2-by-2 matrices (since the definition of convexity involves only two points at a time) and show that the numerical range of the latter is a closed elliptic disc or one of its degenerate forms (circular disc, line segment or a single point). Indeed, if $A = \begin{bmatrix} a & b \\ 0 & c \end{bmatrix}$, then $W(A)$ is the elliptic disc with foci a and c and minor axis of length $|b|$. An easy proof of this is to reduce A to $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ via some affine transformation and check directly that the latter has numerical range $\{z \in \mathbb{C} : |z| \leq 1/2\}$ (cf. [19]).

The numerical range is a bounded set, but is not closed in general. For example, if S is the (*simple*) *unilateral shift* on l^2:

$$S(x_0, x_1, \cdots) = (0, x_0, x_1, \cdots),$$

then $W(S)$ equals the open unit disc $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$. However, if the operator A acts on a finite-dimensional space, then $W(A)$ is obviously closed and hence compact. For an arbitrary operator A, the closure of its numerical range $\overline{W(A)}$ always contains the spectrum $\sigma(A)$. Hence the numerical range gives a rough estimate of the location of the spectrum. This is one of the reasons to study the numerical range and provides its main applications. If A is normal, then $\overline{W(A)}$ equals $\sigma(A)^\wedge$, the convex hull of $\sigma(A)$. Thus, in particular, if A is a normal (finite) matrix, then its numerical range is a polygonal region whose vertices are some of the eigenvalues of A.

A natural question in the study of numerical ranges is to determine which nonempty bounded convex set is the numerical range of some operator on a separable Hilbert space. (Note that if nonseparable Hilbert spaces are allowed, then every such set is the numerical range of some normal operator; compare [26].) Even more intricate is to determine, for each positive integer n, the numerical ranges of operators on an n-dimensional space. Although many necessary/sufficient conditions are known, a complete characterization is beyond reach at this moment. One condition on the boundary of the numerical range is worth noting. If $\triangle$ is a closed convex subset of the plane, then every nondifferentiable point of the boundary $\partial\triangle$ of $\triangle$ has two distinct supporting lines of $\triangle$ with angle less then π such that the closed section formed by them contains $\triangle$. Such a point is called a *corner* of $\triangle$. According to this definition, the endpoints of a line segment are corners. A result of Donoghue [8, Theorem 1] says that *a corner* λ *of* $\overline{W(A)}$ *which also belongs to* $W(A)$ *is a reducing eigenvalue of* A. The latter means that there is a nonzero vector x such that $Ax = \lambda x$ and $A^*x = \overline{\lambda}x$. The proof of this makes use of the geometric fact that an elliptic disc which is containing λ and contained in $\overline{W(A)}$ must be reduced to a line segment. It follows that if A is an n-dimensional operator, then $W(A)$ can have at most n corners. This gives a certain constraint on the shape of the numerical range of a finite-dimensional operator. Using the condition for the equality case of the Cauchy–Schwarz inequality, we may prove the analogous result that *any point* λ *in* $W(A)$ *satisfying* $|\lambda| = \|A\|$ *is a reducing eigenvalue of* A.

Associated with the numerical range $W(A)$ is the quantity $w(A)$, the *numerical radius* of A, defined by $\sup\{|z| : z \in W(A)\}$. For example, if S is the unilateral shift, then $w(S) = 1$, and if A is normal, then $w(A) = \sup\{|z| : z \in \sigma(A)\}$.

We say that the operator A on space H *dilates* to B on K or B *compresses* to A if there is an isometry V from H to K such that $A = V^*BV$. It is easily seen that this is equivalent to B being unitarily equivalent to a 2-by-2 operator matrix of the form $\begin{bmatrix} A & * \\ * & * \end{bmatrix}$. The notion of dilation and compression is closely related to that of numerical range. For one thing, the numerical range itself can be described in terms of dilation. Namely, for any operator A, the numerical range of A is the same as the set of complex numbers λ for which the 1-by-1 matrix $[\lambda]$ dilates to A. On the other hand, if A is an operator which dilates to B, then $W(A)$ is contained in $W(B)$. Hence a judicious choice of a nicely behaved B can yield useful information on the numerical range of A. One type of dilation which will be fully exploited in our derivations in Sections 5 and 6 is the unitary dilation of contractions. The classical result in this respect is Halmos's dilation: *every contraction A ($\|A\| \leq 1$) can be dilated to the unitary operator*

$$\begin{bmatrix} A & (I - AA^*)^{1/2} \\ (I - A^*A)^{1/2} & -A^* \end{bmatrix}$$

(cf. [14, Problem 222 (a)). With more care, the unitary dilation can be achieved in a most economical way: *if A is a contraction on H, then A can be dilated to a unitary operator U from $H \oplus K_1$ to $H \oplus K_2$ with K_1 and K_2 of dimensions $d_{A^*} \equiv \dim \overline{\text{ran } (I - AA^*)^{1/2}}$ and $d_A \equiv \dim \overline{\text{ran } (I - A^*A)^{1/2}}$, respectively, and, moreover, in this case d_{A^*} and d_A are the smallest dimensions of such spaces K_1 and K_2.* Here d_A and d_{A^*} are called the *defect indices* of the contraction A. They provide a measure on how far A deviates from the unitary operators and play a prominent role in the unitary dilation theory.

Properties of numerical ranges of operators are discussed in [14, Chapter 22]; those for finite matrices are in [16, Chapter 1]. The two classic monographs [4] and [5] treat the numerical ranges of elements of normed algebras; the more recent [13] emphasizes applications to numerical analysis.

3. Numerical Range of Finite Matrix

For the study of numerical ranges of finite matrices, the matrix-theoretic properties can be exploited to yield special tools which are not available for general operators. One such tool is the characteristic polynomial of the pencil $x\mathrm{Re}\ A+y\mathrm{Im}\ A$ associated with any matrix A. This can be utilized in two different ways to yield $W(A)$ or its boundary. One is via Kippenhahn's result that the numerical range of A coincides with the convex hull of the real part of the dual curve of $\det(x\mathrm{Re}\ A+y\mathrm{Im}\ A+zI)=0$. In this way, the classical algebraic curve theory can be brought to bear on the study here. On the other hand, a parametric representation of the boundary $\partial W(A)$ can also be obtained from the largest eigenvalue of $\cos\theta\,\mathrm{Re}\ A+\sin\theta\,\mathrm{Im}\ A$ yielding useful information on $W(A)$. Here we give a brief account of both approaches.

Let $\mathbb{CP}^2$ be the complex projective plane consisting of all equivalence classes $[x,y,z]$ of ordered triples of complex numbers x,y and z which are not all zero. Two such triples $[x,y,z]$ and $[x',y',z']$ are equivalent if $x=\lambda x', y=\lambda y'$ and $z=\lambda z'$ for some nonzero λ. The point $[x,y,z]$ $(z\neq 0)$ in homogeneous coordinates can be identified with $(x/z,y/z)$ in nonhomogeneous coordinates. On the other hand, the point (u,v) becomes $[u,v,1]$ in homogeneous coordinates. In this way, $\mathbb{C}^2$ is embedded in $\mathbb{CP}^2$. If $p(x,y,z)$ is a homogeneous polynomial of degree d in x,y and z, then the set of points $[x,y,z]$ in $\mathbb{CP}^2$ satisfying the equation $p(x,y,z)=0$ is an *algebraic curve* of *order* d. If C is such a curve, then its *dual* C^* is defined by

$$C^*=\{[u,v,w]\in\mathbb{CP}^2 : ux+vy+wz=0 \ \text{ is a tangent line of } C\}.$$

In this case, C^* is also an algebraic curve of order at most $d(d-1)$ and d is called the *class* of C^*. It is known that the dual of C^* is C itself. The point $[x_0,y_0,z_0]$ is a *focus* of C if it is not equal to $[1,\pm i,0]$ and the lines through $[x_0,y_0,z_0]$ and $[1,\pm i,0]$ are tangent to C at points other than $[1,\pm i,0]$. In general, if a curve is of class d and is defined by an equation with real coefficients, then it has d real foci and d^2-d complex ones, counting multiplicity.

For an n-by-n matrix A, let

$$p_A(x, y, z) = \det(x\text{Re } A + y\text{Im } A + zI_n)$$

and let $C(A)$ denote the dual curve of $p_A(x, y, z) = 0$. Since p_A is a real homogeneous polynomial of degree n, the curve $C(A)$ is given by a real polynomial of degree at most $n(n-1)$, is of class n, and has n real foci $[a_j, b_j, 1]$, $j = 1, \cdots, n$, which correspond exactly to the n eigenvalues $a_j + ib_j$ of A. The connection of $C(A)$ with the numerical range $W(A)$ is provided by a result of Kippenhahn [18]: *$W(A)$ is the convex hull of the real points of the curve $C(A)$*, namely, the convex hull of the set $\{a + ib \in \mathbb{C} : a, b \in \mathbb{R}, ax + by + z = 0 \text{ is tangent to } p_A(x, y, z) = 0\}$. Kippenhahn's result can be easily verified by noting that $x = \max \sigma(\text{Re } (e^{-i\theta}A))$ is a supporting line of $W(\text{Re } (e^{-i\theta}A))$ for any real θ. Since it can be shown that $\partial W(A)$ contains only finitely many line segments, the above result implies that $\partial W(A)$ is *piecewise algebraic*, that is, it is the union of finitely many algebraic curves.

There is another way to make the above to be more revealing. For any nonempty compact convex subset $\triangle$ of the plane, there is a *natural parametrization* of its boundary $\partial\triangle$. For any $\theta, 0 \leq \theta \leq 2\pi$, let L_θ be the ray from the origin which has inclination θ from the positive x-axis, and let M_θ be the supporting line of $\triangle$ which is perpendicular to L_θ. If $d(\theta)$ is the signed distance from the origin to M_θ, then $\partial\triangle$ can be "parametrized" by $\alpha(\theta) = (x(\theta), y(\theta))$, where

$$\begin{aligned} x(\theta) &= d(\theta)\cos\theta - d'(\theta)\sin\theta, \\ y(\theta) &= d(\theta)\sin\theta + d'(\theta)\cos\theta. \end{aligned}$$

It can be shown that $d(\theta)$ is differentiable for almost all θ and is equal to $\max\{\text{Re } (e^{-i\theta}z) : z \in \triangle\}$. In particular, if $\triangle = \overline{W(A)}$ for some operator A, then

$$\begin{aligned} d(\theta) &= \max \sigma(\text{Re } (e^{-i\theta}A)) \\ &= \max \sigma(\cos\theta \text{ Re } A + \sin\theta \text{ Im } A) \end{aligned}$$

for all θ. This shows that, for a finite matrix A, the degree-n polynomial $p_A(\cos\theta, \sin\theta, z)$ in z has $-d(\theta)$ as a zero. As an example, if A is the 3-by-3 matrix diag $(1, i, 0)$, then

$$d(\theta) = \begin{cases} \cos\theta & \text{if } 0 \le \theta \le \frac{\pi}{4} \text{ or } \frac{3}{2}\pi \le \theta \le 2\pi, \\ \sin\theta & \text{if } \frac{\pi}{4} \le \theta \le \pi, \\ 0 & \text{if } \pi \le \theta \le \frac{3}{2}\pi, \end{cases}$$

and the natural parametrization of $\partial W(A)$ (=the triangular region with vertices $1, i$ and 0) is given by

$$\alpha(\theta) = \begin{cases} 1 & \text{if } 0 < \theta < \frac{\pi}{4} \text{ or } \frac{3}{2}\pi < \theta < 2\pi, \\ i & \text{if } \frac{\pi}{4} < \theta < \pi, \\ 0 & \text{if } \pi < \theta < \frac{3}{2}\pi. \end{cases}$$

In particular, this shows that the natural parametrization is not a parametrization in the usual sense: it does not traverse the line segments on the boundary; but the convex hull of its image equals $\partial\triangle$.

4. Compression of the Shift

Compressions of the shift are a class of operators studied intensively in the 1960s and '70s. Playing a role analogous to the companion matrix in the rational form for finite matrices, they are the building blocks in the "Jordan form" (under quasisimilarity) for the class of C_0 contractions. The whole theory is subsumed under the dilation theory for contractions on Hilbert spaces developed by Sz.-Nagy and Foiaş. The standard reference is the monograph [30]; a more complete account of the theory of C_0 contractions is given in [3].

We start by noting that the unilateral shift S has another representation as $(Sf)(z) = zf(z)$ for f in H^2, the Hardy space of square-summable analytic functions on $\mathbb{D}$. This analytic model of S facilitates a complete description of its invariant subspaces. Indeed, according to the celebrated theorem of Beurling (1949), all nonzero invariant subspaces of S are of the form ϕH^2 for some inner function ϕ (ϕ is *inner* if it is bounded and analytic on $\mathbb{D}$ with $|\phi(e^{i\theta})| = 1$ for almost all real θ). The *compression of the shift* $S(\phi)$ is the operator on $H(\phi) \equiv H^2 \ominus \phi H^2$ defined by

$$S(\phi)f = P(zf(z)),$$

where P denotes the (orthogonal) projection from H^2 onto $H(\phi)$. Thus $S(\phi)$ is the operator in the lower-right corner of the 2-by-2 operator matrix representation of S as

$$\begin{bmatrix} * & * \\ 0 & S(\phi) \end{bmatrix} \quad \text{on } H^2 = \phi H^2 \oplus H(\phi).$$

This class of operators was first studied by Sarason [29] and has been under intensive investigation over the past 35 years. In particular, it is known that $\|S(\phi)\| = 1, S(\phi)$ is *cyclic* (there is a vector $f\ (= 1 - \overline{\phi(0)}\phi)$ in $H(\phi)$ such that $\bigvee\{S(\phi)^n f : n \geq 0\} = H(\phi)$), and its *commutant* $\{S(\phi)\}'$ ($\equiv \{X$ on $H(\phi) : XS(\phi) = S(\phi)X\}$) and *double commutant* $\{S(\phi)\}''$ ($\equiv \{Y$ on $H(\phi) : YX = XY$ for every X in $\{S(\phi)\}'\}$) are both equal to $\{f(S(\phi)) : f \in H^\infty\}$. The inner function ϕ is the *minimal function* of $S(\phi)$ in a sense similar to the minimal polynomial of a finite matrix, that is, it is such that (a) $\phi(S(\phi)) = 0$, and (b) ϕ is a factor of any function f in H^∞ for which $f(S(\phi)) = 0$. An operator A is (unitarily equivalent to) a compression of the shift if and only if it is a contraction, both A^n and A^{*n} converge to 0 in the strong operator topology, and the defect indices d_A and d_{A^*} are both equal to one. It follows from these conditions that the compression of the shift is *irreducible*, that is, it can have no nontrivial reducing subspace.

For finite matrices, the characterization of compressions of the shift in even easier: A is such an operator if and only if it is a contraction, it has no eigenvalue of modulus one and $d_A = 1$. In this case, A is unitarily equivalent to $S(\phi)$ with ϕ is the finite

Blaschke product

$$\phi(z) = \prod_{j=1}^{n} \frac{z - a_j}{1 - \overline{a_j} z},$$

where a_j's are the eigenvalues of A in $\mathbb{D}$. We let $\mathcal{S}_n$ denote the class of such matrices. An example in $\mathcal{S}_n$ is J_n, the n-by-n nilpotent Jordan block

$$\begin{bmatrix} 0 & 1 & & & \\ & \cdot & \cdot & & \\ & & \cdot & \cdot & \\ & & & \cdot & 1 \\ & & & & 0 \end{bmatrix},$$

with the corresponding inner function $\phi(z) = z^n$. By the results in Section 2, matrices in $\mathcal{S}_n$ admit unitary dilations on an $(n+1)$-dimensional space. For this reason, $\mathcal{S}_n$-matrices are called matrices admitting unitary bordering or UB-matrices by Mirman (cf. [21, 22, 23]). Since $S(\phi)$ is defined by its minimal function ϕ, we infer that for any n points $a_1, \cdots, a_n$ in $\mathbb{D}$ (not necessarily distinct) there is a matrix in $\mathcal{S}_n$, unique up to unitary equivalence, with the a_j's as its eigenvalues. A more specific description of a matrix in $\mathcal{S}_n$ with eigenvalues the a_j's is given by $[a_{ij}]_{i,j=1}^n$, where

$$a_{ij} = \begin{cases} a_j & \text{if } i = j, \\ [\prod_{k=i+1}^{j-1}(-\overline{a_k})](1 - |a_i|^2)^{1/2}(1 - |a_j|^2)^{1/2} & \text{if } i < j, \\ 0 & \text{if } i > j. \end{cases} \tag{1}$$

This matricial representation was first discovered by Young [36, p. 235] (cf. also [28, p. 201], [21, Theorem 4] and [11, Corollary 1.3]). In particular, it follows that $\mathcal{S}_2$ consists of 2-by-2 matrices which are unitarily equivalent to a matrix of the form

$$\begin{bmatrix} a & (1 - |a|^2)^{1/2}(1 - |b|^2)^{1/2} \\ 0 & b \end{bmatrix}$$

with a and b in $\mathbb{D}$. There is another representation for matrices in $\mathcal{S}_n$ which is useful for our discussions in Section 5. If A is in $\mathcal{S}_n$, then it has the singular value decomposition $A = UDW$, where U and W are unitary and D is a diagonal matrix diag $(1, \cdots, 1, a)$ with $0 \leq a < 1$. The equality $WAW^* = (WU)D$ shows that A is unitarily equivalent to $(WU)D$, a matrix of the form

$$[f_1 \cdots f_n] \tag{2}$$

whose columns f_j satisfy $\|f_j\| = 1$ for $1 \leq j \leq n-1$, $\|f_n\| < 1$ and $f_j \perp f_k$ for $1 \leq j \neq k \leq n$. Conversely, a matrix of the above form with no eigenvalue of modulus one is in $\mathcal{S}_n$.

¿From the information we have so far on the compressions of the shift, we can already deduce certain properties of their numerical ranges. Let A be a matrix in $\mathcal{S}_n$. Then $W(A)$ must be contained in the open unit disc $\mathbb{D}$. This is because if λ in $W(A)$ is such that $|\lambda| = 1$ $(= \|A\|)$, then it will be a reducing eigenvalue of A, which contradicts the irreducibility of A. On the other hand, by Donoghue's result and the irreducibility of A, we may deduce that the boundary of $W(A)$ is a differentiable curve. In the subsequent sections, we will discuss other finer properties of $W(A)$.

5. Poncelet Property

The recent establishment of a link between the numerical ranges of matrices in $\mathcal{S}_n$ and some classical geometric results from the 19th century was achieved by Mirman [21, 23, 22, 24] and the present authors [9, 10, 11, 12]. Here we give a brief account of this development.

Our first result has to do with a geometric theorem of Poncelet. In this treatise [27] of 1822, there is contained the following result, called *Poncelet's porism* or *Poncelet's closure theorem*: is C and D are ellipses in the plane with C inside D, and if there is one n-gon circumscribed about and inscribed in D, that is, the n-gon has n sides all tangent to C and n vertices on D, then for any point λ on D there is one such circumscribing-inscribing n-gon with λ as a vertex. This is a porism because the assertion says that some property (the existence of a circumscribing-inscribing n-gon) either fails or, if it holds for one instance, succeeds infinitely many times. It is a closure theorem since, from any point λ on D, we draw a tangent line to C, which intersects D at another point, then repeat this process by drawing tangent lines from successive points obtained in this fashion, and obtain the resulting closed n-gon when the nth tangent line reaches back to λ. Viewed dynamically, this gives a configuration of rotating n-gons with different shapes but all sharing this circumscribing-inscribing property. Since the appearance of this result, a huge literature has been developed to its explanation, exposition and generalization. A comprehensive survey of this topic can be found in [6]. We may normalize the outer ellipse D as the unit circle $\partial\mathbb{D}$ via some affine transformation and the inner ellipse C is transformed into one in $\mathbb{D}$ with the n-Poncelet property. More precisely, for $n \geq 3$, we say that a curve Γ in $\mathbb{D}$ has the *n-Poncelet property* if for every point λ on $\partial\mathbb{D}$ there is an n-gon which circumscribes about Γ, inscribes in $\partial\mathbb{D}$ and has λ as a vertex. It is natural to ask whether there are curves other than ellipses in $\mathbb{D}$ which also have the n-Poncelet property. The next theorem provides more examples.

Theorem 5.1. *For any matrix A in $\mathcal{S}_n$ and point λ on $\partial\mathbb{D}$, there is a unique $(n+1)$-gon which circumscribes about $\partial W(A)$, inscribes in $\partial\mathbb{D}$ and has λ as a vertex. In fact, such $(n+1)$-gons P are in one-to-one correspondence with (unitary-equivalence classes of) unitary dilations U of A on an $(n+1)$-dimensional space, under which the $n+1$ vertices of P are exactly the eigenvalues of the corresponding U.*

This theorem appeared in [21, Theorem 1] and [9, Theorem 2.1]. The easy part of the proof is to show that every $(n+1)$-dimensional unitary dilation of a matrix A in $\mathcal{S}_n$ has distinct eigenvalues which form an $(n+1)$-gon inscribed in $\partial\mathbb{D}$ and circumscribed about $W(A)$ with each side tangent to $\partial W(A)$ at exactly one point. To show that such $(n+1)$-gons run over every point of $\partial\mathbb{D}$ takes more work. Instead of outlining its details, we resort to the matrix representations (1) and (2) for A to give the specific $(n+1)$-dimensional unitary dilations U. If A is represented as in (1), then U can be taken as $[b_{ij}]_{i,j=1}^{n+1}$, where

$$(3) \qquad b_{ij} = \begin{cases} a_{ij} & \text{if } 1 \le i, j \le n, \\ \lambda[\prod_{k=1}^{j-1}(-\overline{a_k})](1-|a_j|^2)^{1/2} & \text{if } i = n+1 \text{ and } 1 \le j \le n, \\ [\prod_{k=i+1}^{n}(-\overline{a_k})](1-|a_i|^2)^{1/2} & \text{if } j = n+1 \text{ and } 1 \le i \le n, \\ \lambda\prod_{k=1}^{n}(-\overline{a_k}) & \text{if } i = j = n+1, \end{cases}$$

for some λ in $\partial\mathbb{D}$. Here λ acts as a parameter for the unitary dilations U. On the other hand, if A is as (2), then U can be

$$(4) \qquad \begin{bmatrix} f_1 & \cdots & f_{n-1} & f_n & g \\ 0 & \cdots & 0 & \lambda a & \lambda\|f_n\| \end{bmatrix},$$

where $|\lambda| = 1, a = (1 - \|f_n\|^2)^{1/2} > 0$ and

$$g = \begin{cases} -(a/\|f_n\|)f_n & \text{if } f_n \neq 0, \\ \text{any unit vector orthogonal to } f_1, \cdots, f_{n-1} & \text{if } f_n = 0. \end{cases}$$

Both (3) and (4) can be used to prove that the $(n+1)$-gons with vertices the eigenvalues of U cover all points of $\partial\mathbb{D}$ (the latter is in [9, Theorem 2.1]).

Theorem 5.1 yields additional properties for the numerical ranges of matrices in $\mathcal{S}_n$.

Corollary 5.2. *Let A be a matrix in $\mathcal{S}_n$. Then*

(a) $W(A)$ *is contained in no m-gon inscribed in $\partial\mathbb{D}$ for $m \le n$,*

(b) $w(A) > \cos(\pi/n)$,

(c) Re A *and* Im A *have simple eigenvalues, and*

(d) *the boundary of $W(A)$ contains no line segment and is an algebraic curve.*

Here (a) is an easy consequence of the $(n+1)$-Poncelet property of $\partial W(A)$, (b) follows from (a), (c) is a consequence of, besides the Poncelet property, the interlacing of the eigenvalues of Re A and Re U for $(n+1)$-dimensional unitary dilation U of A, and finally (d) follows from (c) by way of Kippenhahn's result. All assertions except (d) are in [9].

If A is in $\mathcal{S}_n$, so is $e^{-i\theta}A$ for any real θ. Hence the eigenvalues of Re $(e^{-i\theta}A)$ are all distinct by Corollary 5.2 (c). The curves $\Gamma_j, j = 1, \cdots, n$, described by $\alpha_j(\theta) = (x_j(\theta), y_j(\theta))$ with

$$\begin{aligned} x_j(\theta) &= \lambda_j(\theta)\cos\theta - \lambda_j'(\theta)\sin\theta, \\ y_j(\theta) &= \lambda_j(\theta)\sin\theta + \lambda_j'(\theta)\cos\theta, \end{aligned}$$

where $\lambda_j(\theta)$ is the jth largest eigenvalue of Re $(e^{-i\theta}A)$, are expected to have