Linear Algebra and its Applications 434 (2011) 2042–2080

Contents lists available at ScienceDirect

Linear Algebra and its Applications

journal homepage: www.elsevier.com/locate/laa

Numerical shadows: Measures and densities on the numerical range

Charles F. Dunkl [a], Piotr Gawron [b], John A. Holbrook [c,*], Zbigniew Puchała [b], Karol Życzkowski [d,e]

[a] *Department of Mathematics, University of Virginia, Charlottesville, VA 22904-4137, USA*
[b] *Institute of Theoretical and Applied Informatics, Polish Academy of Sciences, Bałtycka 5, 44-100 Gliwice, Poland*
[c] *Department of Mathematics and Statistics, University of Guelph, Guelph, Ontario, Canada N1G 2W1*
[d] *Instytut Fizyki im. Smoluchowskiego, Uniwersytet Jagielloński, Reymonta 4, 30-059 Kraków, Poland*
[e] *Centrum Fizyki Teoretycznej, Polska Akademia Nauk, Aleja Lotników 32/44, 02-668 Warszawa, Poland*

ARTICLE INFO

Article history:
Received 25 October 2010
Accepted 5 December 2010
Available online 21 January 2011

Submitted by C. Davis

AMS classification:
47A12
60B05
81P16
33C05
51M15

Keywords:
Numerical range
Probability measures
Non-normal matrices
Numerical shadow
Higher-rank numerical range

ABSTRACT

For any operator M acting on an N-dimensional Hilbert space $\mathcal{H}_N$ we introduce its *numerical shadow*, which is a probability measure on the complex plane supported by the numerical range of M. The shadow of M at point z is defined as the probability that the inner product (Mu, u) is equal to z, where u stands for a random complex vector from $\mathcal{H}_N$, satisfying $||u|| = 1$. In the case of $N = 2$ the numerical shadow of a non-normal operator can be interpreted as a shadow of a hollow sphere projected on a plane. A similar interpretation is provided also for higher dimensions. For a hermitian M its numerical shadow forms a probability distribution on the real axis which is shown to be a one dimensional B-spline. In the case of a normal M the numerical shadow corresponds to a shadow of a transparent solid simplex in $\mathbb{R}^{N-1}$ onto the complex plane. Numerical shadow is found explicitly for Jordan matrices J_N, direct sums of matrices and in all cases where the shadow is rotation invariant. Results concerning the moments of shadow measures play an important role. A general technique to study numerical shadow via the Cartesian decomposition is described, and a link of the numerical shadow of an operator to its higher-rank numerical range is emphasized.

* Corresponding author.
E-mail addresses: cfd5z@virginia.edu (C.F. Dunkl), gawron@iitis.pl (P. Gawron), jholbroo@uoguelph.ca (J.A. Holbrook), z.puchala@iitis.pl (Z. Puchała), karol@tatry.if.uj.edu.pl (K. Życzkowski).

0024-3795/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.laa.2010.12.003

1. Introduction

The classical numerical range $W(M)$ of a complex $N \times N$ matrix M is the subset of $\mathbb{C}$ defined by

$$W(M) = \{(Mu, u) : u \in \mathbb{C}^N, \|u\| = 1\}.$$

This concept has a long history and has proved a useful tool in operator theory and matrix analysis as well as in more applied areas; for a nice account of some of the lore of $W(M)$, see [15]. Here we are concerned with the measures or densities induced on $W(M)$ by various distributions of the unit vector u; we use the term **numerical shadow** to refer to such densities (see Section 2 for our motivation in using this terminology). Although it seems natural to study the numerical shadow, this subject does not seem to have received much attention until recently. The only earlier extended account (that we know of) is the thesis [26]. This apparent neglect is not the only reason for our attempts to understand the numerical shadow better; the numerical shadow also plays an interesting role in quantum information theory. Applications in this area are the focus of a companion paper [9, in preparation]. Very recently a preprint by Gallay and Serre [16] has appeared, dealing also with mathematical aspects of the numerical shadow (numerical measure, in their terminology). In several ways their development of the subject is parallel to our own (as presented, for example, in [30] and [19]). In the present paper we stress those aspects of our work which are complementary to [16], particularly those based on the **moments** of the numerical shadow.

In this paper we work mainly with the "uniform" distribution of u over the unit sphere Ω_N in $\mathbb{C}^N \equiv \mathbb{R}^{2N}$, i.e. the probability distribution on Ω_N that is invariant under all orthogonal transformations of $\mathbb{R}^{2N}$. In some cases, however, the internal structure of $W(M)$ is better revealed through the use of other distributions, and these are of special importance for the applications discussed in [9].

Although the numerical shadow is in general difficult to determine explicitly, this is possible in a number of interesting cases. Methods based on identifying the moments of the shadow measures are often effective. We also present, via the figures, the results of numerical simulations that display shadow densities in various other cases.

In Section 2 we treat the simple situation occurring when M is 2×2. Here we obtain a "real-life" shadow. In Section 3 we discuss the analogous treatment for $N \times N$ matrices M, i.e. we view the numerical shadow of M as the image of an appropriate measure on the pure states uu^* under a linear map φ_M. In Section 4 we show that in the case of $N \times N$ normal M the numerical shadow is the orthogonal projection of a well-placed model of the $(N-1)$-dimensional simplex (with the uniform density). Thus the density for the numerical shadow is a 2-dimensional B-spline (1-dimensional in the Hermitian case).

Section 5 studies the moments of numerical shadows, yielding a key technique for the identification and comparison of shadows. In Section 6, criteria for the equality of the numerical shadows of two matrices are obtained (in terms, for example, of traces of words in the matrices and their adjoints). Evidently equality occurs when the matrices are unitarily equivalent, but this is not necessary (if $N > 2$).

In Section 7, the numerical shadows are found explicitly for the Jordan nilpotents J_N. Section 8 extends the techniques developed in Section 7 to obtain explicit densities for all rotation invariant shadows.

Section 9 introduces a useful view of the numerical shadow in terms of the (Hermitian) components $\mathrm{Re}(M)$ and $\mathrm{Im}(M)$ in the Cartesian decomposition of M, and the unitary matrix linking those components. Several related aspects of the numerical shadow are treated in that section, including a connection with the Radon transform. Section 10 relates the numerical shadow of a direct sum to the shadows of its summands. Section 11 is concerned with numerical approximations of shadow densities in terms of moments and Zernike expansions.

Finally, in Section 12, we relate the numerical shadow of M to the so-called rank-k numerical ranges $\Lambda_k(M)$. The theory and applications of these ranges has been advanced vigorously since their introduction only a few years ago as a tool in quantum information theory (see for example [5–8,28, 24,23,14]). One way to describe $\Lambda_k(M)$ is that it consists of those points (Mu, u) in $W(M)$ where u may be chosen from the unit sphere in a whole k-dimensional subspace of $\mathbb{C}^N$. Thus it is natural to ask to what extent $\Lambda_k(M)$ may be identified as a region of greater density within the numerical shadow. Here the shadows corresponding to **real** unit vectors u play a role.

Let us fix some notation. The algebra of complex $N \times N$ matrices is here denoted by $M_N(\mathbb{C})$ or simply M_N. The adjoint or conjugate transpose of a matrix $M \in M_N$ is denoted by M^*; we consider vectors v in $\mathbb{C}^N$ as column vectors and v^* is the conjugate transpose. Our inner product (v, w) may be computed as w^*v. Recall that the unit sphere in $\mathbb{C}^N$ is denoted by Ω_N, i.e.

$$\Omega_N = \{u \in \mathbb{C}^N : \|u\| = 1\}.$$

The uniform probability measure on Ω_N is denoted by μ. Given $M \in M_N$, the notion of "numerical shadow of M" is captured formally as the probability measure P_M on $W(M)$ such that

$$P_M(S) = \mu\{u \in \Omega_N : (Mu, u) \in S\},$$

for each Borel subset S of $W(M)$. Equivalently, for any continuous function $g : W(M) \to \mathbb{C}$ we have

$$\int_{W(M)} g(z)\, dP_M(z) = \int_{\Omega_N} g((Mu, u))\, d\mu(u). \tag{1}$$

If P_M has a probability density (with respect to planar measure in $\mathbb{C}$) it is denoted by f_M. In those cases where f_M is rotation-invariant, i.e. $f_M(z) = f_M(|z|)$ we consider f_M as a function of $r \in (0, w(M))$, where $w(M)$ is the so-called numerical radius of M:

$$w(M) = \max\{|z| : z \in W(M)\}.$$

2. The 2 × 2 case; real-life shadows

Here we compute the shadow density for an arbitrary 2×2 matrix M. Our method is to view this density as a real-life shadow of the Bloch sphere model for Ω_2. An equivalent result was obtained by Ng (see [26]) by a rather different method.

The Bloch sphere model sees $W(M)$ as the genuine shadow of a hollow sphere made of infinitely thin semi-transparent uniform material, where in general the light would fall obliquely on the (complex) plane. Of course, this situation cannot quite be realized physically, but playing with a hollow plastic ball in bright sunlight may yield a good approximation. It is well-known that when $M \in M_2$ the numerical range $W(M)$ is a filled ellipse with the eigenvalues of M as foci. The following proposition, visualized in Fig. 1, supplies further information in the form of an explicit shadow density.

Proposition 2.1. *Let E be the filled ellipse formed by $W(M)$ and let a and b be the lengths of the semimajor and semiminor axes of E; then the shadow density is*

$$\frac{1}{2\pi ab\sqrt{1-r^2}} \tag{2}$$

at every point on the elliptical curve bounding rE ($0 \leqslant r \leqslant 1$).

Proof. Recall that we assume that u is chosen "uniformly" over $\{u \in \mathbb{C}^2 : \|u\| = 1\}$, i.e. according to the measure μ on Ω_2. It is known that $|u_1|^2$ will then be uniform in $[0, 1]$. This a special case of the fact that u uniform in Ω_N implies $(|u_1|^2, |u_2|^2, \ldots, |u_N|^2)$ has the uniform distribution in the $N-1$-dimensional simplex, see [29,3]. It is also important to note that for fixed u_1 the relative phase of u_2 is $e^{i\theta}$ where θ is uniform in $[0, 2\pi]$. We have

$$uu^* = \begin{bmatrix} |u_1|^2 & u_1\overline{u_2} \\ \overline{u_1}u_2 & |u_2|^2 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 2|u_1|^2 - 1 & 2u_1\overline{u_2} \\ 2\overline{u_1}u_2 & 1 - 2|u_1|^2 \end{bmatrix}.$$

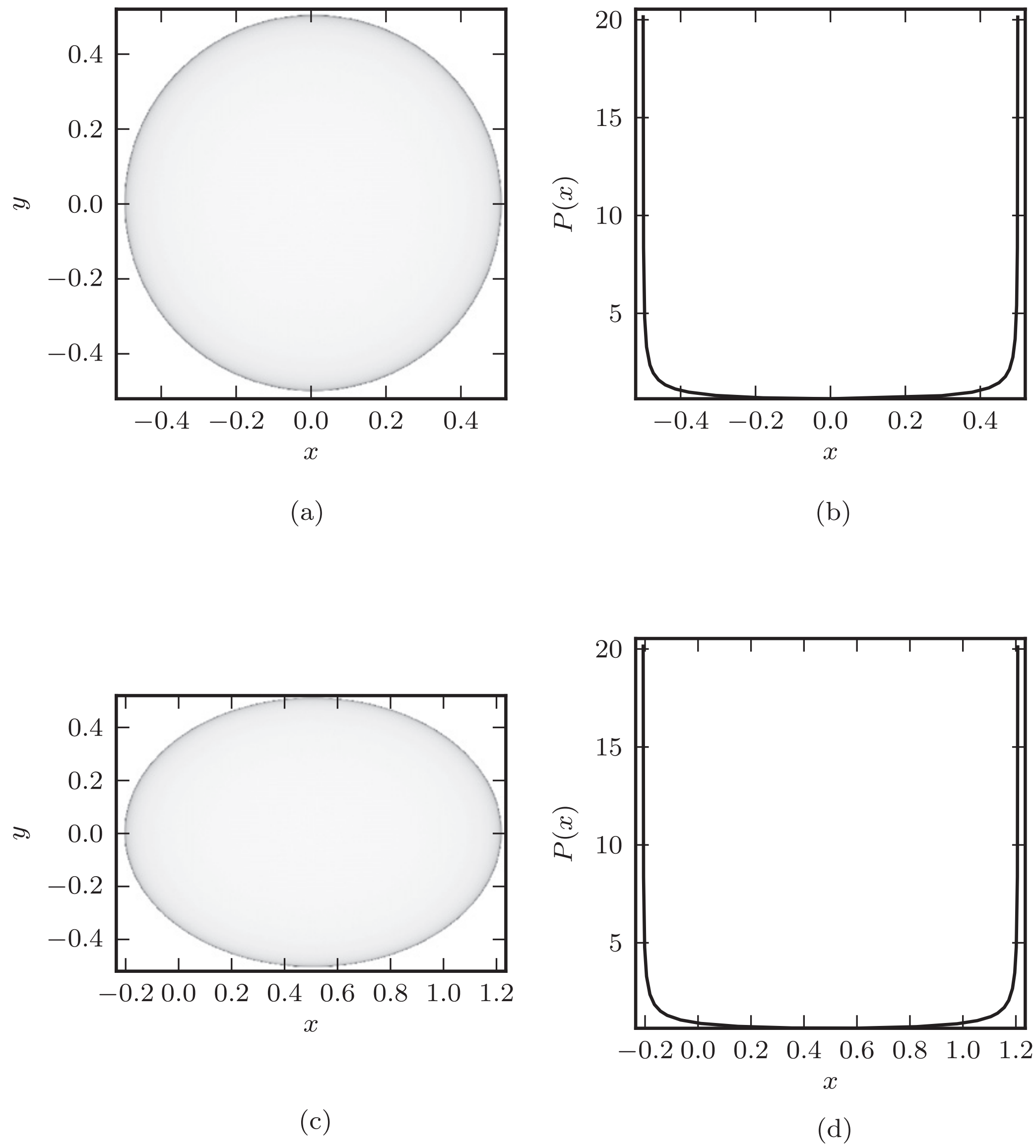

Fig. 1. (a) Numerical shadow of matrix nonnormal $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ resembles physical shadow cast by the hollow sphere made of transparent material when illuminated by a light source at infinity. (b) Cross-section of the shadow along the real axis. (c) Numerical shadow of matrix nonnormal $\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$ resembles physical shadow cast by the hollow sphere made of transparent material when illuminated by a light source at infinity, but with screen not perpendicular to the light rays. (d) Cross-section of the shadow along the real axis.

Let $z = 2|u_1|^2 - 1$; then z is uniform in $[-1, 1]$ and $2u_1\overline{u_2} = \sqrt{1 - z^2}e^{-i\theta} = x - iy$, so that (recalling Archimedes) (x, y, z) is uniform on the unit sphere. This is one way to see that the corresponding distribution on the Bloch sphere $\{uu^* : u \in \Omega_2\}$ is uniform.

Following Davis [12], we compute (Mu, u) as

$$\mathrm{tr}(u^*Mu) = \mathrm{tr}(Muu^*) = \mathrm{tr}\left(\frac{1}{2}M + \frac{1}{2}M\begin{bmatrix} z & x - iy \\ x + iy & -z \end{bmatrix}\right).$$

For convenience we take $M = \begin{bmatrix} 1 & q \\ 0 & -1 \end{bmatrix}$ with $q \geqslant 0$. This is a harmless normalization, achieved via translation and rotation of M (which respects the distribution of (Mu, u)) and unitary similarity (which leaves the distribution unchanged). Then $(Mu, u) = z + (q/2)(x + iy) = (z + bx, by)$ in the complex plane, with $b = q/2$. Consider the region E bounded by the ellipse centred at $(0, 0)$ with horizontal semiaxis of length $a = \sqrt{1 + b^2}$ and vertical semiaxis of length b. Given $r \in [0, 1]$, (Mu, u) lies in rE iff

$$\frac{(z + bx)^2}{1 + b^2} + \frac{(by)^2}{b^2} \leqslant r^2.$$

A calculation verifies that this is equivalent to $((x, y, z) \cdot (1/a, 0, -b/a))^2 \geqslant 1 - r^2$, saying that (x, y, z) lies on either of the spherical caps of the unit sphere that are symmetrical about the axis determined by $(1/a, 0, -b/a)$ and have radius r. According to Archimedes (or the related formulas found in calculus texts) the relative area of these caps is $1 - \sqrt{1 - r^2}$. Hence the probability

$$P((Mu, u) \in rE) = 1 - \sqrt{1 - r^2}.$$

To find the corresponding planar density, we first observe that the region $(r + \Delta r)E \setminus rE$ corresponds to symmetrical rings bordering the spherical caps mentioned above. Thus the planar density will be constant on the ellipse bounding rE. Its value there is then given by

$$\lim_{\Delta r \to 0} \frac{\left(1 - \sqrt{1 - (r + \Delta r)^2}\right) - \left(1 - \sqrt{1 - r^2}\right)}{\pi ab(r + \Delta r)^2 - \pi abr^2} = \frac{1}{2\pi ab\sqrt{1 - r^2}}. \qquad \square$$

This result is equivalent to the formula for the density obtained by Ng (see [26, p. 67]). He computes the density of (Mu, u) at (x, y) in the ellipse as

$$p(x, y) = \frac{1}{2\pi ab\sqrt{1 - (x^2/a^2 + y^2/b^2)}}.$$

His method seems unrelated to the Bloch sphere approach worked out above.

3. Numerical shadows as linear images of the pure quantum states

It is clear that the argument of Section 2 can be extended in part to cases where $N > 2$. The Bloch sphere is replaced by the set of density matrices representing pure quantum states:

$$PQS_N = \{uu^* : u \in \Omega_N\}.$$

Just as before, for any $u \in \Omega_n$ and $M \in M_N$ we have

$$(Mu, u) = \operatorname{tr}(u^*Mu) = \operatorname{tr}(Muu^*),$$

so that $W(M)$ is the image of PQS_N under the linear map $\varphi_M : M_N \to \mathbb{C}$ defined by

$$\varphi_M(X) = \operatorname{tr}(MX).$$

Since each $X \in PQS_N$ is Hermitian we may also write

$$\varphi_M(X) = \operatorname{tr}(MX^*) = (M, X)_F,$$

where $(\cdot, \cdot)_F$ is the Frobenius inner product on M_N.

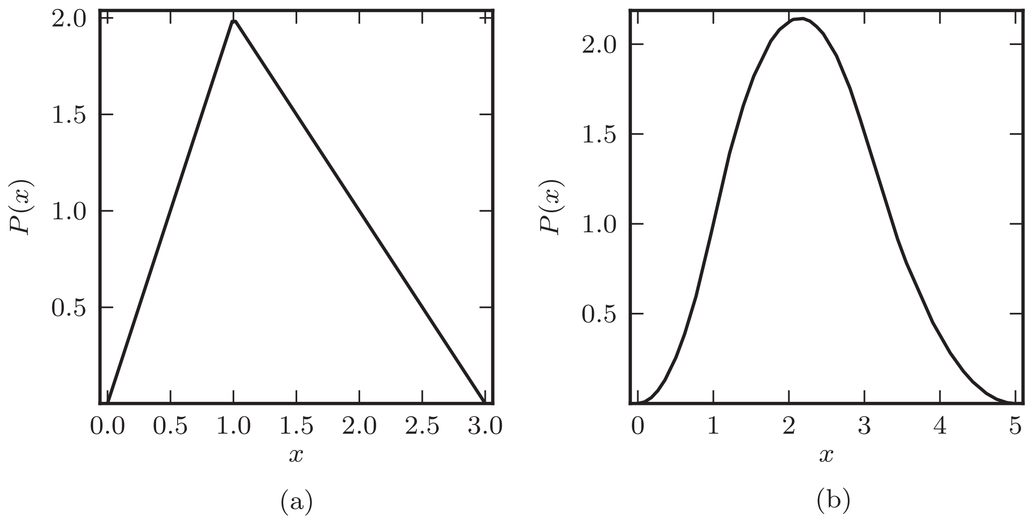

Fig. 2. Probability density function of hermitian matrices of dimensions 3 and 4. (a) Shadow of matrix diag$(0, 1, 3)$ is a spline function of degree one. (b) Shadow of matrix diag$(0, 1, 3, 5)$ is a spline function of degree two.

Thus the numerical shadow of M may be viewed as the measure on $W(M)$ induced by applying the linear map φ_M to the fixed measure ν on PQS_N that corresponds to the uniform μ on Ω_N. As M varies the resulting numerical shadows may be regarded as a tomographic study of the measure ν. Thus such detailed information as we have about numerical shadows (see Section 9, for example, or [16]) reveals much about the structure of ν on PQS_N.

4. The Hermitian and normal cases: B-splines

The standard N-simplex Δ_N is defined by

$$\Delta_N = \left\{ r \in \mathbb{R}^N : \text{all } r_k \geqslant 0 \text{ and } \sum_{k=1}^{N} r_k = 1 \right\}.$$

This is an $(N-1)$-dimensional convex subset of $\mathbb{R}^N$. We say r is "uniformly distributed" over Δ_N to mean that r is uniform with respect to normalized $(N-1)$-dimensional Lebesgue measure vol_{N-1} on Δ_N.

Lemma 4.1. *Let P_A be the shadow measure of a normal matrix $A \in M_N$. Then for any Borel subset B of the plane*

$$P_A(B) = Prob\{r^*\lambda \in B\},$$

where $\lambda = (\lambda_1, \ldots, \lambda_N)^t$ is the spectrum of A and r is uniformly distributed over the standard N-simplex Δ_N.

Proof. Since A is normal, it is unitarily similar to diag(λ). As μ is invariant under unitary transformations we have

$$P_A(B) = \text{Prob}\{(\text{diag}(\lambda)u, u) \in B : u \in \Omega_N\} = \text{Prob}\left\{ \sum_{k=1}^{N} \lambda_k |u_k|^2 \in B : u \in \Omega_N \right\}.$$

It is known (see e.g. [29,3]) that if u is uniform over Ω_N (i.e. distributed according to μ) then

$$r = (|u_1|^2, \ldots, |u_N|^2)$$

is uniform over the simplex Δ_N. □

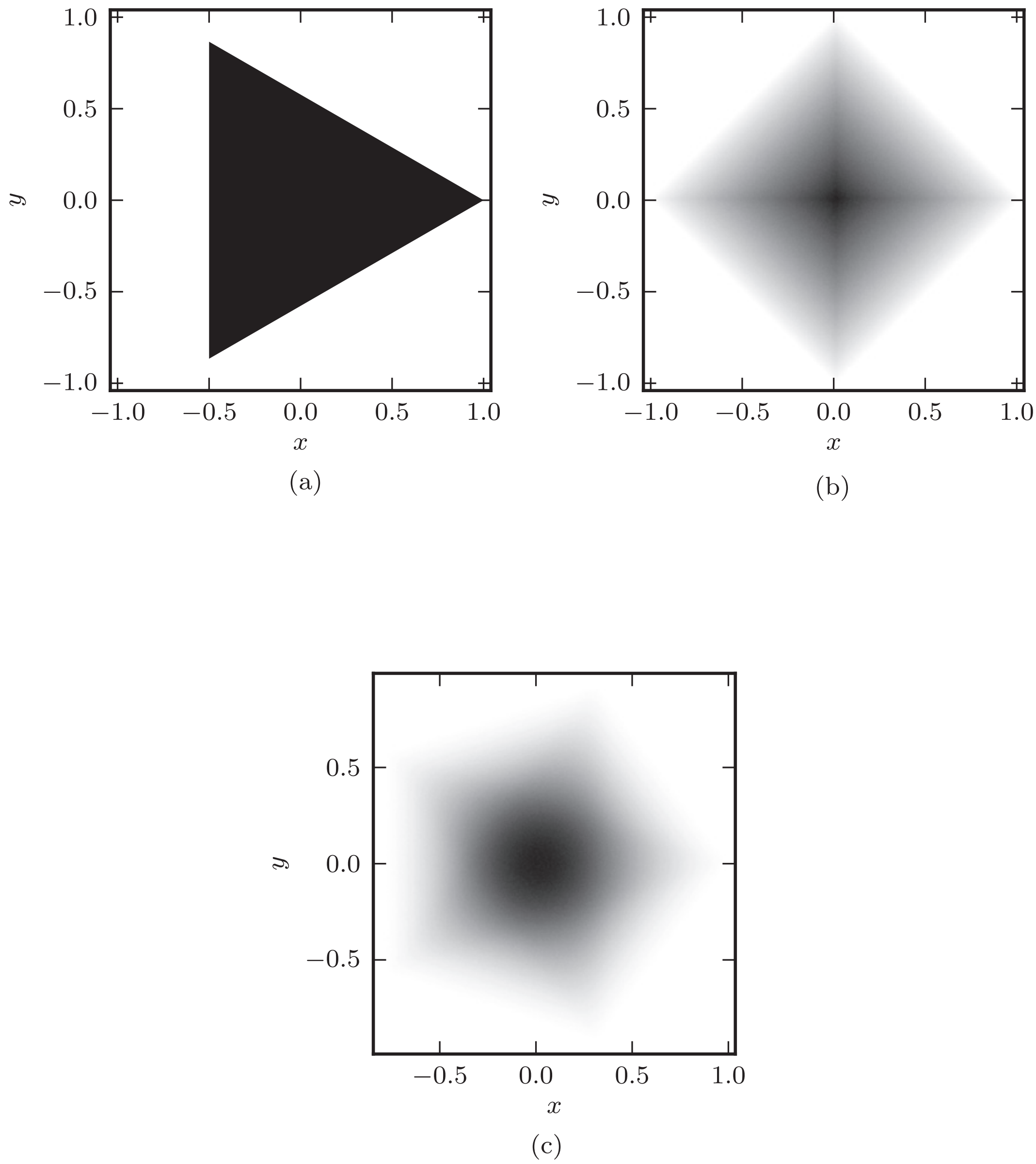

Fig. 3. Probability density function of unitary matrices of dimensions 3, 4 and 5. (a) Shadow of the matrix $\mathrm{diag}(e^{2/3i\pi}, e^{-2/3i\pi}, 1)$ forms an uniform distribution supported by the triangle spanning its eigenvalues. (b) Shadow of the matrix $\mathrm{diag}(1, i, -1, -i)$ forms a regular pyramid. (c) Shadow of the matrix $\mathrm{diag}(e^{2/5i\pi}, e^{4/5i\pi}, e^{6/5i\pi}, e^{8/5i\pi}, 1)$.

With λ as above, write $\lambda = x + iy$ where $x, y \in \mathbb{R}^N$. Assume first that x and y are independent vectors. Let W be a real matrix with x, y as its first two columns and with columns 3 to N forming an orthonormal basis for $\{x, y\}^{\perp}$. For any vector $v \in \mathbb{R}^N$,

$$v^*\lambda = a + ib \ \ (a, b \in \mathbb{R}) \quad \Leftrightarrow \quad (W^*v)_1 = a \text{ and } (W^*v)_2 = b.$$

Thus for Borel $B \subset \mathbb{C} \equiv \mathbb{R}^2$

$$\mathrm{Prob}\{r^*\lambda \in B\} = vol_{N-1}\{r \in \Delta_N : ((W^*r)_1, (W^*r)_2) \in B\},$$

and, in view of Lemma 4.1, the density for P_A at $(a, b) \in W(A)$ is

$$vol_{N-3}\{r \in \Delta_N : (W^*r)_1 = a, (W^*r)_2 = b\} = vol_{N-3}\{v \in W^*(\Delta_N) : v_1 = a, v_2 = b\}.$$

Let us recall the definition of an s-dimensional B-spline (from [13]).

Definition 4.2. Let σ be a nontrivial simplex in $\mathbb{R}^{s+k}$. On $\mathbb{R}^s$ we define the B-spline of order k from σ by

$$M_{k,\sigma}(x_1, \ldots, x_s) = vol(\sigma \cap \{v \in \mathbb{R}^{s+k} : v_j = x_j \ (j = 1, 2, \ldots, s)\}).$$

Using this terminology we may summarize our results as follows.

Proposition 4.3. *The numerical shadow of an $N \times N$ normal matrix with eigenvalues $\lambda \in \mathbb{C}^N$ having linearly independent real and imaginary parts has as density a 2-dimensional B-spline $M_{N-2,\sigma}(a, b)$, where the simplex $\sigma = W^*(\Delta_N)$ with some W chosen as above.*

Remark 4.4. In the case where the real and imaginary parts of λ are dependent (as when A is Hermitian), it is easy to see that the numerical shadow is 1-dimensional (a line segment, in fact) with density given by a 1-dimensional B-spline (compare [13, Lemma 9.1]). As examples, the shadows of Hermitian matrices of size $N = 3$ and $N = 4$ are shown in Fig. 2.

Remark 4.5. This observation for the Hermitian case was worked out in detail by Ng in [26]. He also made the right conjecture regarding the normal case. In a sense, the normal case was earlier understood by statisticians studying the distribution of quadratic forms; see for example [1, chapter 6]; Anderson points out that some of the relevant ideas go back to von Neumann in the 40's. Anderson seems to discuss only real quadratic forms; thus the normal case corresponds to roots of multiplicity 2.

Remark 4.6. In view of Proposition 4.3, the theory of B-splines may be applied to see that normal shadow densities are piecewise polynomial functions of two variables in the normal case – see Fig. 3 and of a single variable in the Hermitian case. The latter case is analyzed in some detail in Sec. 9. A thorough analysis of the B-spline shadow densities for normal matrices is also provided in [16].

5. Moments of the numerical shadow

We denote the moments of the numerical shadow of $A \in M_N$ by

$$\nu_{jk}(A) = \int_{W(A)} z^j \overline{z}^k \, dP_A(z).$$

Note that, since the polynomials in z and $\overline{z}$ are uniformly dense in the continuous functions on $W(A)$, these moments determine the numerical shadow uniquely. Moreover, in view of (1), we have

$$\nu_{jk}(A) = \int_{\Omega_N} (Au, u)^j (\overline{(Au, u)})^k \, d\mu(u) = \int_{\Omega_N} (Au, u)^j (A^* u, u)^k \, d\mu(u). \tag{3}$$

Given $\lambda \in \mathbb{C}^N$ and a multi-index $\alpha \in \mathbb{N}_0^N$ (where $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$), we use the following notation: $\lambda^\alpha = \lambda_1^{\alpha_1} \lambda_2^{\alpha_2} \ldots \lambda_N^{\alpha_N}$, $|\alpha| = \sum_1^N \alpha_k$, $\alpha! = \alpha_1! \alpha_2!, \ldots, \alpha_N!$. We also use the Pochhammer symbol or shifted factorial $(x)_n = \prod_{j=1}^n (x + j - 1)$; by convention $(x)_0 = 1$.

The effective evaluation of the moments $\nu_{jk}(A)$ depends on the following proposition.

Proposition 5.1. *Given $A \in M_N$, let $\lambda \in \mathbb{C}^N$ list the eigenvalues of A repeated according to multiplicity. Then*

$$\int_{\Omega_N} (Au, u)^n \, d\mu(u) = \frac{n!}{(N)_n} h_n(\lambda), \tag{4}$$

where $h_n(\lambda)$ is the complete symmetric polynomial of degree n, i.e.

$$h_n(\lambda) = \sum_{\alpha \in \mathbb{N}_0^N, \ |\alpha| = n} \lambda^\alpha.$$

Proof. Given multi-indices α, β, let

$$Q(\alpha, \beta) = \int_{\Omega_N} u^\alpha (\overline{u})^\beta \, d\mu(u), \tag{5}$$

where the conjugation $\overline{u}$ is applied entrywise. Since μ is invariant under the unitary map $u \to (e^{i\theta}u_1, u_2, \ldots, u_N)^t$, we have $Q(\alpha, \beta) = e^{i(\alpha_1-\beta_1)\theta}Q(\alpha, \beta)$ for each real θ. Hence $Q(\alpha, \beta) = 0$ unless $\alpha_1 = \beta_1$. Similarly for the other components, so that $Q(\alpha, \beta) = 0$ unless $\alpha = \beta$. More work is required to evaluate $Q(\alpha, \alpha)$:

$$Q(\alpha, \alpha) = \int_{\Omega_N} |u|^{2\alpha}\, d\mu(u) = \frac{\alpha!}{(N)_{|\alpha|}}, \tag{6}$$

where $|u| = (|u_1|, |u_2|, \ldots, |u_N|)^t$. A convenient trick here is to consider

$$I = \int_{\mathbb{R}^{2N}} e^{-\sum_1^N (x_k^2+y_k^2)} \prod_1^N (x_k^2 + y_k^2)^{\alpha_k}\; dx\, dy.$$

As a product of Gamma-integrals we obtain $I = \pi^N \alpha!$. Integrating first over $r\Omega_N$, with $r^2 = \sum_1^N (x_k^2 + y_k^2)$, then over $0 < r < \infty$, we find that $I = \frac{1}{2}|S^{2N-1}|(N + |\alpha| - 1)!Q(\alpha, \alpha)$, where $|S^{2N-1}|$ denotes the $(2N-1)$-dimensional area of Ω_N. Since $Q(\vec{0}, \vec{0}) = 1$, the formula (6) follows.

We may assume A is in the Schur upper-triangular form, since this is obtained via a unitary similarity and μ is invariant under unitary transformations on $\mathbb{C}^N$. Thus $A_{jj} = \lambda_j$ (some listing of the eigenvalues of A, with multiplicity) and

$$(Au, u) = \sum_{j=1}^N \lambda_j |u_j|^2 + \sum_{j>i} A_{ij} u_j \overline{u_i}.$$

Aside from $(\sum_j \lambda_j |u_j|^2)^n$, the terms of $(Au, u)^n$ are scalar multiples of expressions of the form

$$\prod_{k=1}^a |u_{\ell_k}|^2 \prod_{k=1}^b u_{j_k} \overline{u_{i_k}},$$

where $b \geqslant 1$ and each $j_k > i_k$. Such an expression has the form $u^\gamma (\overline{u})^\gamma u^\alpha (\overline{u})^\beta$ where, using $e(j)$ as a temporary notation for the multi-index with 1 in the jth position and 0's elsewhere,

$$\alpha = \sum_{k=1}^b e(j_k), \quad \beta = \sum_{k=1}^b e(i_k).$$

Clearly $\alpha_1 = 0$ and for some first $k_0 > 1$ we have $\alpha_{k_0} > 0$; hence $\beta_k > 0$ for some $k < k_0$, so that $\alpha \neq \beta$. Thus $Q(\gamma + \alpha, \gamma + \beta) = 0$ so that such terms make no contribution to the integral over Ω_N. It follows that

$$\int_{\Omega_N} (Au, u)^n\, d\mu(u) = \int_{\Omega_N} \left(\sum_{j=1}^N \lambda_j |u_j|^2 \right)^n d\mu(u).$$

Using the multinomial formula and (6), this integral is

$$\sum_{\alpha \in \mathbb{N}_0^N,\ |\alpha|=n} \frac{n!}{\alpha!} \lambda^\alpha Q(\alpha, \alpha) = \sum_{\alpha \in \mathbb{N}_0^N,\ |\alpha|=n} \frac{n!}{\alpha!} \lambda^\alpha \frac{\alpha!}{(N)_n} = \frac{n!}{(N)_n} \sum_{\alpha \in \mathbb{N}_0^N,\ |\alpha|=n} \lambda^\alpha = \frac{n!}{(N)_n} h_n(\lambda). \qquad \square$$

It will be convenient to use the notation $\lambda(A)$ to denote any listing of the eigenvalues of A, repeated according to multiplicity. Applying (4) with A replaced by $sA + tA^*$ (t, s real) and recalling (3) we obtain

$$\sum_{j=0}^n \binom{n}{j} s^j t^{n-j} \nu_{j,n-j}(A) = \frac{n!}{(N)_n} h_n(\lambda(sA + tA^*)). \tag{7}$$

Moreover, the RHS of (7) may be evaluated in terms of traces of words in A and A^*, using known relations [25] among $h_n(\lambda)$, the power sums

$$p_j(\lambda) = \sum_{k=1}^N \lambda_k^j$$

(these equal $\mathrm{tr}(A^j)$ if $\lambda = \lambda(A)$), and the elementary symmetric polynomials

$$e_j(\lambda) = \sum_{1 \leqslant i_1 < i_2 < \cdots < i_j \leqslant N} \lambda_{i_1} \lambda_{i_2} \cdots \lambda_{i_j}$$

(note that $e_j(\lambda) = 0$ if $j > N$; by convention $e_0(\lambda) = 1$). For $n \geqslant 1$ we have

$$e_0 h_n - e_1 h_{n-1} + e_2 h_{n-2} \cdots \pm e_n h_0 = 0 \tag{8}$$

(by convention $h_0(\lambda) = 1$) and

$$ne_n = p_1 e_{n-1} - p_2 e_{n-2} + \cdots \pm p_n e_0. \tag{9}$$

For example, $h_1 = p_1$ so that (7) implies

$$\begin{aligned} t\nu_{0,1}(A) + s\nu_{1,0}(A) &= \frac{1}{N} p_1(\lambda(sA + tA^*)) = \frac{1}{N}\mathrm{tr}(sA + tA^*) \\ &= \frac{1}{N}(t\ \mathrm{tr}(A^*) + s\ \mathrm{tr}(A)). \end{aligned}$$

Hence

$$\nu_{1,0}(A) = \frac{1}{N}\mathrm{tr}(A) \text{ and } \nu_{0,1}(A) = \frac{1}{N}\mathrm{tr}(A^*).$$

Likewise $1h_2 = e_1 h_1 - e_2 h_0 = e_1 h_1 - \frac{1}{2}(p_1 e_1 - p_2 e_0) = p_1^2 - \frac{1}{2}p_1^2 + \frac{1}{2}p_2 = \frac{1}{2}(p_2 + p_1^2)$, so that (7) implies

$$\begin{aligned} &t^2 \nu_{0,2}(A) + 2ts\nu_{1,1}(A) + s^2 \nu_{2,0}(A) \\ &\quad = \frac{2}{N(N+1)} \left(\frac{1}{2}\mathrm{tr}(sA + tA^*)^2 + \frac{1}{2}\mathrm{tr}^2(sA + tA^*)\right) \\ &\quad = \frac{1}{N(N+1)}\left(t^2 \left(\mathrm{tr}(A^*)^2 + \mathrm{tr}^2 A^*\right) + 2ts(\mathrm{tr}(AA^*) + \mathrm{tr}(A)\mathrm{tr}(A^*)) + s^2 \left(\mathrm{tr}(A^2) + \mathrm{tr}^2 A\right)\right) \end{aligned}$$

(using $\mathrm{tr}(A^*A) = \mathrm{tr}(AA^*)$). Thus we have

$$\nu_{2,0}(A) = \frac{\mathrm{tr}(A^2) + \mathrm{tr}^2 A}{N(N+1)}, \quad \nu_{1,1}(A) = \frac{\mathrm{tr}(AA^*) + \mathrm{tr}(A)\mathrm{tr}(A^*)}{N(N+1)},$$

and

$$\nu_{0,2}(A) = \frac{\mathrm{tr}(A^*)^2 + \mathrm{tr}^2 A^*}{N(N+1)}.$$

Similarly we find that $h_3 = \frac{1}{3}p_3 + \frac{1}{2}p_1 p_2 + \frac{1}{6}p_1^3$, so that (7) implies

$$\begin{aligned} &t^3 \nu_{0,3}(A) + 3t^2 s\nu_{1,2}(A) + 3ts^2 \nu_{2,1}(A) + s^3 \nu_{3,0}(A) \\ &\quad = \frac{6}{N(N+1)(N+2)} \left(\frac{1}{3}\mathrm{tr}(sA + tA^*)^3 + \frac{1}{2}\mathrm{tr}(sA + tA^*)\mathrm{tr}(sA + tA^*)^2 + \frac{1}{6}\mathrm{tr}^3(sA + tA^*)\right) \\ &\quad = \frac{1}{N(N+1)(N+2)}(t^3(2\mathrm{tr}(A^*)^3 + 3\mathrm{tr}(A^*)^2\mathrm{tr}(A^*) + \mathrm{tr}^3 A^*) + t^2 s(6\mathrm{tr}(A(A^*)^2) \\ &\qquad + 6\mathrm{tr}(A^*)\mathrm{tr}(AA^*) + 3\mathrm{tr}(A)\mathrm{tr}(A^*)^2 + 3\mathrm{tr}(A)\mathrm{tr}^2 A^*) + ts^2 \cdots), \end{aligned}$$

where again the cyclicity of the trace plays a role.

From these calculations we obtain

$$\nu_{3,0}(A) = \frac{1}{N(N+1)(N+2)}\left(2\mathrm{tr}(A^3) + 3\mathrm{tr}(A^2)\mathrm{tr}(A) + \mathrm{tr}^3 A\right),$$
$$\nu_{2,1}(A) = \frac{1}{N(N+1)(N+2)}\left(2\mathrm{tr}(A^2A^*) + 2\mathrm{tr}(A)\mathrm{tr}(AA^*) + \mathrm{tr}(A^2)\mathrm{tr}(A^*) + \mathrm{tr}^2 A\mathrm{tr}(A^*)\right),$$

and $\nu_{1,2}(A)$, $\nu_{0,3}(A)$ by interchanging the roles of A and A^*.

What is not clear from the approach above is the important fact that all the moments $\nu_{j,k}(A)$ are polynomials in the traces of A, A^* words of length at most N. One way to see this is to note that (8) and (9) allow us to express h_n for $n > N$ in terms of $p_1, p_2, \ldots, p_N$. For example, when $N = 3$ we find that $h_4 = \frac{2}{3}p_1p_3 + \frac{1}{12}p_1^4 + \frac{1}{4}p_2^2$.

In general then we need only compute $\mathrm{tr}(sA + tA^*)^k$ for $k \leqslant N$, and therefore (in view of the noncommutative binomial formula) we need only compute $\mathrm{tr}(P_{k,j}(A, A^*))$ for $k \leqslant N$, where

$$P_{k,j}(x, y) = \sum_{S\subseteq\{1,2,\ldots,k\},\ \#(S)=j} \{z_1z_2\ldots z_k \text{ where } z_i = x \text{ if } i \in S,\ z_i = y \text{ if } i \notin S\}.$$

We summarize this discussion in the following proposition.

Proposition 5.2. *Given $A \in M_N$, all the moments $\nu_{i,j}(A)$ of the shadow measure (and therefore the measure P_A itself) are determined by the traces of $(sA + tA^*)^k$ (as polynomials in s, t) for $k \leqslant N$. Thus they are determined by the values of*

$$tr\left(P_{k,j}(A, A^*)\right)$$

for $k \leqslant N$.

Remark. The cyclicity of the trace (i.e. $\mathrm{tr}(AB) = \mathrm{tr}(BA)$) reduces $\mathrm{tr}\left(P_{k,j}(A, A^*)\right)$ to a single term when $k \leqslant 3$ but this is not always the case. For example

$$\mathrm{tr}\left(P_{4,2}(A, A^*)\right) = 4\mathrm{tr}\left(A^2(A^*)^2\right) + 2\mathrm{tr}(AA^*)^2.$$

The information about moments $\nu_{j,k}(A)$ that is provided by (7) may also be encoded in the series

$$S(A, s, t, q) = \sum_{n=0}^{\infty} q^n \frac{(N)_n}{n!}\left(\sum_{j=0}^{n}\binom{n}{j}s^j t^{n-j}\nu_{j,n-j}(A)\right) = \sum_{n=0}^{\infty} q^n h_n(\lambda(sA + tA^*))$$

(absolutely convergent for small real s, t, q).

The following proposition provides powerful alternative forms for this series.

Proposition 5.3. *Given $A \in M_N$ we have (for sufficiently small s, t, q)*

$$S(A, s, t, q) = det^{-1}\left(I - q(sA + tA^*)\right) \tag{10}$$

and

$$S(A, s, t, q) = \left(\sum_{k=0}^{N}(-q)^k e_k\left(\lambda(sA + tA^*)\right)\right)^{-1}. \tag{11}$$

Proof. These follow from the identities

$$\sum_{n=0}^{\infty} q^n h_n(\lambda) = \prod_{j=1}^{N}(1 + q\lambda_j + q^2\lambda_j^2 + \cdots)$$
$$= \left(\prod_{j=1}^{N}(1 - q\lambda_j)\right)^{-1} = \left(\sum_{k=1}^{N}(-q)^k e_k(\lambda)\right)^{-1}. \qquad \square$$

Remark 5.4. In view of (10), the shadow measure P_A is completely determined by $\det(I-(sA+tA^*))$ as a polynomial in s, t (q may be absorbed into t, s).

Remark 5.5. In view of (9), the relation (11) provides another viewpoint on Proposition 5.2.

6. Criteria for equality of numerical shadows

Given $A, B \in M_N$, we have seen in the last section that $P_A = P_B$
iff

$$\operatorname{tr}(sA+tA^*)^k = \operatorname{tr}(sB+tB^*)^k \tag{12}$$

(as polynomials in s, t) for all $k \leqslant N$
iff

$$\operatorname{tr}\left(P_{k,j}(A,A^*)\right) = \operatorname{tr}\left(P_{k,j}(B,B^*)\right) \tag{13}$$

for all $j \leqslant k \leqslant N$
iff

$$\det(I-(sA+tA^*)) = \det\left(I-(sB+tB^*)\right) \tag{14}$$

(as polynomials in s, t).

Since the uniform measure μ on Ω_N is invariant under unitary transformations, $P_A = P_{U^*AU}$ for any unitary U. It is natural, therefore, to ask whether the numerical shadow P_A determines A up to unitary similarity. This is the case for $A \in M_2$, for example, since the ellipse $E = W(A)$, just as a set, determines an upper-triangular form for A: A is unitarily similar to $\begin{bmatrix} \alpha & s \\ 0 & \beta \end{bmatrix}$ where the eigenvalues α, β are the foci of E and s is the length of the minor axis of E. The answer is "yes" also for normal matrices A since the eigenvalues are determined by P_A in that case (see Section 4).

More generally, however, the answer is "no", on several levels. First of all, the measure μ is also invariant under any orthogonal transformation of $\mathbb{R}^{2N} \equiv \mathbb{C}^N$, so that, in particular, $d\mu(u) = d\mu(\overline{u})$. Thus A and its transpose A^t, though they are not usually unitarily similar, always have the same numerical shadow:

$$(Au, u) = u^*Au = (u^*Au)^t = u^tA^t(u^*)^t = (\overline{u})^*A^t\overline{u} = (A^t\overline{u}, \overline{u}).$$

In fact, the maps $A \mapsto U^*AU$ and $A \mapsto U^*A^tU$, are the only linear maps on M_N that preserve the numerical shadow, since they are the only linear maps that preserve the numerical range as a set (see Li's survey [22]).

Moreover, particular pairs A, B may have the same numerical shadow without being related by unitarily similarity or transpose. This phenomenon is somewhat clarified by comparing the trace criterion (13) for $P_A = P_B$ with the analogous criteria for unitary similarity. In a 1940 paper [27] Specht observed that A and B are unitarily similar iff

$$\operatorname{tr}\left(w(A,A^*)\right) = \operatorname{tr}\left(w(B,B^*)\right)$$

for all two-variable words $w(\cdot,\cdot)$. Since then much work has been done with the aim of limiting the set of words required in Specht's criterion when matrices of a given size are involved. In [11] Djoković and Johnson provide a welcome account of recent results in this direction. In particular, the following result (see Theorem 2.4 in [11]) may be compared with (13).

Proposition 6.1. *Given $A, B \in M_N$, there exists unitary U such that $B = U^*AU$ iff*

$$tr\left(w(A,A^*)\right) = tr\left(w(B,B^*)\right) \tag{15}$$

for all words $w(\cdot,\cdot)$ of length $\leqslant N^2$.

The disparity between (15) and (13) certainly suggests that A and B might have the same numerical shadow without being simply related by unitaries. Let us see how this does occur when $N = 3$. In [11], Djoković and Johnson refer to a result of Sibirskiĭ: the unitary equivalence class of $A \in M_3$ is determined by

$$\operatorname{tr}(A),\quad \operatorname{tr}(A^2),\quad \operatorname{tr}(AA^*),\quad \operatorname{tr}(A^3),\quad \operatorname{tr}(A^2A^*),\quad \operatorname{tr}\left(A^2(A^*)^2\right),\quad \text{and}\quad \operatorname{tr}\left(A^2(A^*)^2AA^*\right),$$

and this set is minimal. In contrast, (13) tells us that the numerical shadow P_A is determined (when $A \in M_3$) by

$$\mathrm{tr}(A), \mathrm{tr}(A^2), \mathrm{tr}(AA^*), \mathrm{tr}(A^3), \text{ and } \mathrm{tr}(A^2A^*).$$

Thus we expect to find $A, B \in M_3$ such that $P_A = P_B$ but A and B are not unitarily related.

A class of specific examples is provided by

$$A = \begin{bmatrix} 0 & x & 0 \\ y & 0 & 0 \\ z & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & y & 0 \\ x & 0 & 0 \\ z & 0 & 0 \end{bmatrix}.$$

Note that $\det(I - (sA + tA^*)) = 1 - st(|x|^2 + |y|^2 + |z|^2) - s^2xy - t^2\overline{xy}$. Since this expression is symmetric in x, y, (14) tells us that $P_A = P_B$. Consider the choice $x = 0, y = z = 1$: then A has rank 1 while B has rank 2. Clearly B is not unitarily similar to A or to A^t.

Remark. The common numerical shadow of these A, B is identified explicitly in Section 7, because B is unitarily equivalent to the Jordan nilpotent J_3.

7. Numerical shadows of Jordan nilpotents J_N

Here we compute explicit shadow densities for certain special matrices, focusing on the Jordan nilpotent J_N, i.e. $J_N \in M_N(\mathbb{C})$ with 1's on the superdiagonal and 0's elsewhere. Of course, the discussion of the 2×2 case in Section 2 applies to J_2 and shows that the planar density of the shadow P_{J_2} at $z \in \mathbb{C}$ is $f_2(|z|)$ where

$$f_2(r) = \frac{1}{2\pi(1/2)^2\sqrt{1-4r^2}} = \frac{2}{\pi\sqrt{1-4r^2}},$$

since $W(J_2)$ is a disc of radius $1/2$. The shadow density for J_3 can be computed by several methods but here we will do it as the simplest case of a general method that exploits the moment techniques from Section 5. We shall see that the shadow density for J_N is an alternating sum of densities supported on discs with centre at 0 and with various radii, the largest being $\cos(\pi/(N+1))$. This is a striking development beyond the well-known numerical radius formula: $w(J_N) = \cos(\pi/(N+1))$ (see [10] for information about the numerical radii of certain matrices with simple structure; for more, see [20]).

Observe first that the shadow measure P_{J_N} is certainly circularly symmetric about 0; in fact J_N and $e^{i\theta}J_N$ are unitarily similar (use $U = \mathrm{diag}(1, e^{i\theta}, e^{i2\theta}, \ldots)$). Thus, from Proposition 5.3, we have

$$\sum_{m=0}^{\infty} \frac{(N)_m}{m!m!} s^m t^m \nu_{mm}(J_N) = \det{}^{-1}(I_N - (sJ_n + tJ_N^*)). \tag{16}$$

We may take $s = t$ and identify $\nu_{mm}(J_N)$ via the coefficient of t^{2m} in $\det^{-1}\left(I_N - t(J_N + J_N^*)\right)$. Now the eigenvalues of $J_N + J_N^*$ are well-known:

$$2\cos\left(\frac{k\pi}{N+1}\right) \quad (k = 1, 2, \ldots, N).$$

[Some say that this was the first nontrivial eigenvalue problem ever solved, and that it goes all the way back to Cauchy.]

Thus the RHS of (16) (for $s = t$) can be calculated explicitly. The details appear in the proof of the following proposition.

Proposition 7.1. *For each $N \geqslant 2$ and $m = 0, 1, \ldots$*

$$\nu_{mm}(J_N) = \sum_{k=1}^{\lfloor N/2 \rfloor} c_k \left(\cos^2\left(\frac{k\pi}{N+1}\right)\right)^m \frac{m!}{\left(\frac{N}{2}\right)_m} \frac{m!}{\left(\frac{N+1}{2}\right)_m},$$

where

$$c_k = (-1)^{k-1}\frac{2^{N+1}}{N+1}\sin^2\left(\frac{k\pi}{N+1}\right)\left(\cos\left(\frac{k\pi}{N+1}\right)\right)^{N-1}.$$

Proof. By (16), with (small) $s = t$,

$$\sum_{m=0}^{\infty}\frac{(N)_m}{m!m!}t^{2m}\nu_{mm}(J_N) = \frac{1}{\prod_{k=1}^{N}\left(1-2t\cos\left(\frac{k\pi}{N+1}\right)\right)}.$$

Now the $\cos\left(\frac{k\pi}{N+1}\right)$ are the roots of the monic polynomial $C_N(x)$, where

$$C_N(\cos\theta) = \frac{1}{2^N}\frac{\sin(N+1)\theta}{\sin\theta}$$

(a version of the Chebyshev polynomials of the second kind). Thus

$$\frac{1}{C_N(x)} = \frac{1}{\prod_{k=1}^{N}\left(x-\cos\left(\frac{k\pi}{N+1}\right)\right)} = \sum_{k=1}^{N}\frac{a_k}{\left(x-\cos\left(\frac{k\pi}{N+1}\right)\right)},$$

where the coefficients a_k in the partial fraction decomposition are given by $a_k = 1/C_N'\left(\cos\left(\frac{k\pi}{N+1}\right)\right)$. Using the formula for $C_N(\cos\theta)$ we find that

$$a_k = (-1)^{k-1}\frac{2^N\sin^2\left(\frac{k\pi}{N+1}\right)}{N+1}.$$

We now have

$$\begin{aligned}\sum_{m=0}^{\infty}\frac{(N)_m}{m!m!}t^{2m}\nu_{mm}(J_N) &= \frac{1}{(2t)^N\prod_{k=1}^{N}\left((1/2t)-\cos\left(\frac{k\pi}{N+1}\right)\right)} = \frac{1}{(2t)^N}\frac{1}{C_N(1/2t)}\\ &= \frac{1}{t^N}\sum_{k=1}^{N}\frac{(-1)^{k-1}}{N+1}\sin^2\left(\frac{k\pi}{N+1}\right)\frac{1}{\left((1/2t)-\cos\left(\frac{k\pi}{N+1}\right)\right)}\\ &= \frac{1}{t^{N-1}}\frac{2}{N+1}\sum_{k=1}^{N}(-1)^{k-1}\sin^2\left(\frac{k\pi}{N+1}\right)\sum_{j=0}^{\infty}\left(2t\cos\left(\frac{k\pi}{N+1}\right)\right)^j.\end{aligned}$$

Evidently the summed coefficients for j odd and for $j < N-1$ are 0 [we need not worry about how this happens!] so that the term in t^{2m} for the final expression corresponds to $j = 2m+N-1$. Thus

$$\nu_{mm}(J_N) = \sum_{k=1}^{N}(-1)^{k-1}\frac{2^N}{N+1}\sin^2\left(\frac{k\pi}{N+1}\right)\left(\cos\left(\frac{k\pi}{N+1}\right)\right)^{N-1}\left(4\cos^2\left(\frac{k\pi}{N+1}\right)\right)^m\frac{m!m!}{(N)_{2m}}.$$

Note that the terms for k and $N-k+1$ are the same and that $(N)_m$ may be replaced by $\left(\frac{N}{2}\right)_m\left(\frac{N+1}{2}\right)_m 2^{2m}$. Then $\nu_{mm}(J_N)$ is given by

$$\sum_{k=1}^{\lfloor N/2\rfloor}(-1)^{k-1}\frac{2^{N+1}}{N+1}\sin^2\left(\frac{k\pi}{N+1}\right)\left(\cos\left(\frac{k\pi}{N+1}\right)\right)^{N-1}\left(\cos^2\left(\frac{k\pi}{N+1}\right)\right)^m\frac{m!m!}{\left(\frac{N}{2}\right)_m\left(\frac{N+1}{2}\right)_m},$$

(the additional factor of 2 is correct even if N is odd because then $\cos\left(\frac{k\pi}{N+1}\right) = 0$ for $k = (N+1)/2$). □

The value of Proposition 7.1 lies in the possibility of identifying explicitly those densities with moments

$$b^m\frac{m!}{\left(\frac{N}{2}\right)_m}\frac{m!}{\left(\frac{N+1}{2}\right)_m}.$$

Proposition 7.2. *Suppose $f(x)$ and $g(x)$ are probability densities on $[0, 1]$ with moments*

$$\int_0^1 x^m f(x)dx = a_m, \quad \int_0^1 x^m g(x)dx = b_m.$$

Then

(i) *for any $b > 0$, $(1/b)f(x/b)$ is a probability density on $[0, b]$ with moments $b^m a_m$, and*

(ii) *a probability density on $[0, 1]$ with moments $a_m b_m$ is given by*

$$h(x) = \int_x^1 f(s) g\left(\frac{x}{s}\right) \frac{ds}{s}.$$

Proof

(i) With the substitution $y = x/b$,

$$\int_0^b x^m \frac{1}{b} f\left(\frac{x}{b}\right) dx = \int_0^1 b^m y^m f(y)dy = b^m a_m.$$

(ii) Consider independent random variables X, Y with f, g as probability densities. Then XY has moments

$$E((XY)^m) = E(X^m)E(Y^m) = a_m b_m.$$

Since X, Y have joint density $f(x)g(y)$, Prob$\{XY \leqslant t\}$ is

$$\int_0^1 f(x)\left(\int_0^{t/x} g(y)dy\right)dx = \int_0^t f(x)\left(\int_0^1 g(y)dy\right)dx + \int_t^1 f(x)\left(\int_0^{t/x} g(y)dy\right)dx,$$

for $t \in [0, 1]$. Differentiate with respect to t to obtain the density $h(t)$ for XY:

$$f(t) - f(t) + \int_t^1 f(x) g\left(\frac{t}{x}\right)\frac{dx}{x}. \quad \square$$

We can now compute the density $F_N(x)$ on $[0, 1]$ having moments

$$\int_0^1 x^m F_N(x)dx = \frac{m!}{\left(\frac{N+1}{2}\right)_m} \frac{m!}{\left(\frac{N}{2}\right)_m},$$

for $m = 0, 1, 2, \ldots$ and $N \geqslant 2$. For any $\beta > 0$ we have the beta-integrals

$$\int_0^1 \beta x^m (1-x)^{\beta-1}\, dx = \frac{m!}{(\beta+1)_m},$$

for $m = 0, 1, 2, \ldots$ (use induction on m via integration by parts). For $N = 2$ take $\beta = 1/2$ to see that

$$F_2(x) = \frac{1}{2\sqrt{1-x}}.$$

For $N \geqslant 3$ we apply Proposition 7.2(ii): $F_N(x) = h(x)$ computed with

$$f(x) = \frac{N-1}{2}(1-x)^{\frac{N-3}{2}}, \quad g(x) = \frac{N-2}{2}(1-x)^{\frac{N-4}{2}}.$$

Consider even $N = 2\ell$:

$$\begin{aligned} F_{2\ell}(x) &= \int_x^1 \frac{2\ell-1}{2}(1-s)^{\frac{2\ell-3}{2}} \frac{2\ell-2}{2}\left(1-\frac{x}{s}\right)^{\frac{2\ell-4}{2}} \frac{ds}{s} \\ &= \frac{(\ell-1)(2\ell-1)}{2}\int_x^1 (1-s)^{\ell-\frac{3}{2}}(s-x)^{\ell-2} s^{1-\ell}\, ds. \end{aligned}$$

With the substitution $s = 1 - u^2$ we obtain

$$F_{2\ell}(x) = (\ell - 1)(2\ell - 1)\int_0^{\sqrt{1-x}} u^{2\ell-2}(1 - x - u^2)^{\ell-2}\,\frac{du}{(1-u^2)^{\ell-1}}.$$

For odd $N = 2\ell + 1$ we reverse the roles of f and g to obtain

$$\begin{aligned}F_{2\ell+1}(x) &= \int_x^1 \frac{2\ell-1}{2}(1-s)^{\ell-\frac{3}{2}}\,\ell\left(1-\frac{x}{s}\right)^{\ell-1}\,\frac{ds}{s}\\ &= \ell(2\ell-1)\int_0^{\sqrt{1-x}} u^{2\ell-2}(1-x-u^2)^{\ell-1}\,\frac{du}{(1-u^2)^{\ell}}.\end{aligned}$$

The integrals representing $F_N(x)$ are elementary in the sense that they may in principle be computed explicitly (using partial fractions, for example). In particular,

$$F_3(x) = \int_0^{\sqrt{1-x}} \frac{du}{1-u^2} = \log\frac{1+\sqrt{1-x}}{\sqrt{x}},$$

and

$$F_4(x) = 3\int_0^{\sqrt{1-x}} u^2\frac{du}{1-u^2} = 3\log\frac{1+\sqrt{1-x}}{\sqrt{x}} - 3\sqrt{1-x}.$$

In fact, there is a recurrence relation for the $F_N(x)$ that makes the calculation of F_N for $N > 4$ a simple task; such matters are discussed at the end of this section.

Returning to shadow densities, let the planar density of P_{J_N} at $z \in \mathbb{C}$ be denoted by $f_N(|z|)$ so that

$$\nu_{mm}(J_N) = \int_0^{2\pi}\int_o^{w(J_N)} r^{2m}f_N(r)r\,dr\,d\theta = 2\pi\int_0^{w(J_N)} r^{2m+1}f_N(r)dr.$$

With the substitution $x = r^2$ we have

$$\nu_{mm}(J_N) = \pi\int_0^{w^2(J_N)} x^m f_N\left(\sqrt{x}\right)dx.$$

In view of Proposition 7.1 and Proposition 7.2(i),

$$\pi\int_0^{w^2(J_N)} x^m f_N\left(\sqrt{x}\right)dx = \sum_{k=1}^{\llcorner N/2\lrcorner} c_k\int_0^{\cos^2\frac{k\pi}{N+1}} \frac{x^m}{\cos^2\frac{k\pi}{N+1}}F_N\left(\frac{x}{\cos^2\frac{k\pi}{N+1}}\right)dx,$$

where c_k are as in Proposition 7.1. Thus $\pi f_N(\sqrt{x})$ and

$$\sum_{k=1}^{\llcorner N/2\lrcorner} \frac{c_k}{\cos^2\frac{k\pi}{N+1}}F_N\left(\frac{x}{\cos^2\frac{k\pi}{N+1}}\right)$$

coincide (since they have the same moments). We obtain

$$f_N(x) = \frac{1}{\pi}\sum_{k=1}^{\llcorner N/2\lrcorner} \frac{c_k}{\cos^2\frac{k\pi}{N+1}}F_N\left(\frac{x^2}{\cos^2\frac{k\pi}{N+1}}\right),$$

an (alternating) sum of densities supported on $[0, \cos\frac{k\pi}{N+1}]$; in particular we have a greatly refined version of the result $w(J_N) = \cos\frac{\pi}{N+1}$ $(= \max_k \cos\frac{k\pi}{N+1})$.

In summary, we have proved

Proposition 7.3. *The radial density $f_N(r)$ of P_{J_N} is given by*

$$f_N(r) = \frac{1}{\pi}\sum_{k=1}^{\llcorner N/2\lrcorner} (-1)^{k-1}\frac{2^{N+1}}{N+1}\sin^2\frac{k\pi}{N+1}\left(\cos\frac{k\pi}{N+1}\right)^{N-3} F_N\left(\frac{r^2}{\cos^2\frac{k\pi}{N+1}}\right)$$

for any $N \geqslant 2$.

For $N = 2$ we see again that

$$f_2(r) \quad = \left(\frac{8}{3\pi}\sin^2\frac{\pi}{3}\cos^{-1}\frac{\pi}{3}\frac{1}{2\sqrt{1-r^2/\cos^2\frac{\pi}{3}}}\right) = \frac{2}{\pi}\frac{1}{\sqrt{1-4r^2}}.$$

Likewise

$$f_3(r) = \frac{4}{\pi}\sin^2\frac{\pi}{4}F_3\left(\frac{r^2}{\cos^2\frac{\pi}{4}}\right) = \frac{2}{\pi}\log\frac{1+\sqrt{1-2r^2}}{\sqrt{2}r}.$$

For $N > 3$ the radial density combines densities on discs of several different radii. For example,

$$\begin{aligned} f_4(r) &= \frac{32}{5\pi}\left(\sin^2\frac{\pi}{5}\ \cos^2\frac{\pi}{5}F_4\left(\frac{r^2}{\cos^2\frac{\pi}{5}}\right) - \sin^2\frac{2\pi}{5}\ \cos^2\frac{2\pi}{5}F_4\left(\frac{r^2}{\cos^2\frac{2\pi}{5}}\right)\right) \\ &= \frac{1}{5\pi}\left(\left(5+\sqrt{5}\right)F_4\left(\frac{8r^2}{3+\sqrt{5}}\right) - \left(5-\sqrt{5}\right)F_4\left(\frac{8r^2}{3-\sqrt{5}}\right)\right). \end{aligned}$$

We shall see that the functions F_N are the basic building blocks for many circularly symmetric numerical shadows. Hence it will be worthwhile to explore their properties more thoroughly. To this end, we introduce the following hypergeometric series:

$$H(a, b; c; t) = \sum_{j=0}^{\infty}\frac{(a)_j(b)_j}{(c)_j j!}t^j;$$

in this context H is often denoted by $_2F_1$. Here we may assume that the parameters a, b, c are real and that $c \neq 0, -1, -2, \ldots$. Note that the series converges absolutely for $|t| < 1$ since it has the form $\sum d_j t^j$ where

$$|d_{j+1}/d_j| = |(a+j)(b+j)/((c+j)(j+1))| \to_j 1.$$

Recall the Gauss summation formula, which tells us that the series converges also for $t = 1$ whenever $a, b \geqslant 0$ and $c - a - b > 0$ and that, in such a case,

$$H(a, b; c; 1) = \frac{\Gamma(c-a-b)\Gamma(c)}{\Gamma(c-a)\Gamma(c-b)}.$$

Given $\beta \geqslant 0$, $\delta > 0$ and $k = 1, 2, \ldots$, let

$$G(x) = \frac{\Gamma(\beta+1)\Gamma(\delta+k)}{\Gamma(\beta+\delta)(k-1)!}(1-x)^{\beta+\delta-1}H(\delta, \beta+1-k; \beta+\delta; 1-x), \tag{17}$$

for $0 < x < 1$.

Proposition 7.4. *The function defined by (17) is a probability density on* $(0, 1)$ *with moments*

$$\int_0^1 x^m G(x)dx = \frac{m!(k)_m}{(\beta+1)_m(\delta+k)_m} \quad (m = 0, 1, 2, \ldots).$$

Proof. Evaluating the beta-functions $\int_0^1 x^m(1-x)^{\beta+\delta-1+j}\,dx$ as

$$\frac{m!\Gamma(\beta+\delta+j)}{\Gamma(\beta+\delta+j+m+1)} = \frac{m!\ \Gamma(\beta+\delta)(\beta+\delta)_j}{\Gamma(\beta+\delta)(\beta+\delta)_{m+1+j}} = \frac{m!\ (\beta+\delta)_j}{(\beta+\delta)_{m+1}(\beta+\delta+m+1)_j},$$

we see that

$$\int_0^1 x^m G(x)\,dx = \frac{m!\ \Gamma(\beta+1)\Gamma(\delta+k)}{\Gamma(\beta+\delta)(k-1)!\ (\beta+\delta)_{m+1}}\sum_{j=0}^{\infty}\frac{(\delta)_j(\beta+1-k)_j}{(\beta+\delta+m+1)_j\ j!}.$$

Using the Gauss summation formula, we obtain

$$\frac{m!\ \Gamma(\beta+1)\Gamma(\delta+k)}{\Gamma(\beta+\delta)(k-1)!\ (\beta+\delta)_{m+1}} \cdot \frac{\Gamma(m+k)\Gamma(\beta+\delta+m+1)}{\Gamma(\beta+1+m)\Gamma(\delta+k+m)}$$
$$= \frac{m!\ \Gamma(\beta+1)\Gamma(\delta+k)}{\Gamma(\beta+\delta+m+1)(k-1)!} \cdot \frac{(k-1+m)!\ \Gamma(\beta+\delta+m+1)}{\Gamma(\beta+1)(\beta+1)_m\Gamma(\delta+k)(\delta+k)_m}$$
$$= \frac{m!\ (k)_m}{(\beta+1)_m(\delta+k)_m}. \quad \square$$

Taking $k = 1$, $\beta = (N-2)/2$, $\delta = (N-1)/2$ we see in particular that

$$F_N(x) = \frac{\Gamma\left(\frac{N}{2}\right)\Gamma\left(\frac{N+1}{2}\right)}{\Gamma(N-\frac{3}{2})}(1-x)^{N-\frac{5}{2}}H\left(\frac{N-1}{2}, \frac{N-2}{2}; N-\frac{3}{2}; 1-x\right). \tag{18}$$

Several useful recurrence relations will follow from the following general recurrence for H.

Lemma 7.5. *If a, b are real and $b > 1$ then*

$$H\left(a-\frac{1}{2}, a-1; b; t\right) - H\left(a-1, a-\frac{3}{2}; b-1; t\right)$$
$$= \frac{(2b+1-2a)(a-1)}{2b(b-1)} \cdot tH\left(a, a-\frac{1}{2}; b+1; t\right).$$

This lemma may be verified by a careful comparison of the terms involving t^{j+1} $(j = 0, 1, \ldots)$.

Using (18) and invoking the lemma with $t = 1-x$, $a = (N+1)/2$, $b = N-\frac{1}{2}$, we obtain the recurrence relation for $F_{N+2}(x)$ $(N \geqslant 2)$:

$$F_{N+2}(x) = \frac{N+1}{(N-1)^2}((2N-3)F_{N+1}(x) - (1-x)N\,F_N(x)). \tag{19}$$

In fact, then, each $F_N(X)$ has the form $a_N(x)F_2(x) + b_N(x)F_3(x)$ for certain polynomials a_N, b_N.

We extend the definition of $F_N(x)$ by setting it equal to 0 for $x \geqslant 1$; this is the natural continuous extension except that $F_2(x) = (2\sqrt{1-x})^{-1} \uparrow \infty$ as $x \uparrow 1$. From (18) it is clear that for $N \geqslant 3$ we have $F_N(x) \to 0$ as $x \uparrow 1$; hence $a_N(1) = 0$ for $N \geqslant 3$. We may also examine the behavior of the functions F_N at 0: $F_3(x) = \log(1+\sqrt{1-x}) - \frac{1}{2}\log x \uparrow \infty$ as $x \downarrow 0$, whereas $F_2(x) \to \frac{1}{2}$. On the other hand (18) shows that, for $N \geqslant 3$, $F_N(x)$ tends to a constant times $H((N-1)/2, (N-2)/2; N-3/2; 1)$ as $x \downarrow 0$; the Gauss summation formula tells us that this limit is $+\infty$ (since $c-a-b = 0$). Thus $b_N(0) > 0$ for $N \geqslant 3$, and $F_N(x)$ grows like $-\log x$ as $x \downarrow 0$

8. Rotation-invariant shadows

Here we shall see that the methods of Section 7 extend to determine explicit densities for all rotation-invariant numerical shadows. These are shadows of $A \in M_N$ such that A and $e^{i\theta}A$ have the same shadow for all real θ. Characterizing such A in terms of moments is easy: $\nu_{jk}(A) = 0$ whenever $j \neq k$. More elusive are characterizations directly in terms of A.

Simple examples are provided by the "superdiagonal" matrices: i.e. A such that $a_{ij} = 0$ unless $j = i+1$. For such A we actually have $e^{i\theta}A$ unitarily similar to A: let $U = \mathrm{diag}(1, e^{i\theta}, e^{i2\theta}, \ldots)$; then $U^*AU = e^{i\theta}A$. The Jordan nilpotents J_N are special cases of these superdiagonal matrices.

More generally, consider the incidence graph $G(A)$ of $A \in M_N$: vertices are $\{1, 2, \ldots, N\}$ and i, j are joined by an edge iff $a_{ij} \neq 0$. The interesting case in this context is when $G(A)$ consists of disjoint chains (no cycles are allowed; in particular, A has zero diagonal). One can see that this condition is equivalent to requiring that A have zero diagonal, have no more than two nonzero entries in each cross-shaped region formed by the kth row and the kth column, and that $G(A)$ have no cycles.

Proposition 8.1. *If $G(A)$ consists of disjoint chains, then A and $e^{i\theta}A$ are unitarily similar (so that A has rotation-invariant shadow).*

Proof. Consider the unitary $U = \operatorname{diag}(u)$ where $u_{j_k} = e^{ik\theta}$ for each chain

$$j_1 \to j_2 \to \cdots \to j_k \to \cdots \to j_K$$

of $G(A)$ (it does not matter which orientation of the chain is chosen). Set $u_j = 1$ for any j that does not occur in any of the chains that make up $G(A)$. Note that

$$(U^*AU)_{j_k,j_{k+1}} = e^{ik\theta} a_{j_k,j_{k+1}} e^{i(k+1)\theta} = e^{i\theta} a_{j_k,j_{k+1}}.$$

Since other entries of A are 0, we do have $U^*AU = e^{i\theta}A$. □

This proposition applies, for example, to superdiagonal A as well as to strictly upper-triangular A that are "subpermutation" matrices, i.e. have at most one nonzero entry in each row and in each column.

The next proposition notes that A with rotation-invariant shadow must be nilpotent, so that it is unitarily similar to a strictly upper-triangular matrix (Schur form).

Proposition 8.2. *If $A \in M_N$ has rotation-invariant numerical shadow, then all eigenvalues are 0.*

Proof. Putting $t = 0$ in (7), we see that $h_n(\lambda(A)) = 0$ $(n \geqslant 1)$; indeed, this is the case whenever $\nu_{n,0}(A) = 0$. From (8) and (9) we conclude that $p_n(\lambda(A)) = 0$ for $n \geqslant 1$. Thus $\sum_1^N p(\lambda_k) = 0$ for any polynomial $p(x)$ with $p(0) = 0$. Suppose λ_i occurs with multiplicity m. Let

$$p(x) = x \prod_{\lambda_j \neq \lambda_i} (x - \lambda_j);$$

then

$$0 = \sum_1^N p(\lambda_k) = m\lambda_i \prod_{\lambda_j \neq \lambda_i} (\lambda_i - \lambda_j),$$

so that $\lambda_i = 0$. □

Proposition 8.3. *The matrix $A \in M_N$ has rotation-invariant numerical shadow*

iff

(i) $\det\left(I - (sA + tA^*)\right)$ *is a function of* st;

iff

(ii) *A is nilpotent and* $tr\left(P_{k,j}(A, A^*)\right) = 0 \quad \left(\frac{k}{2} < j < k \leqslant N\right)$.

Proof. In view of (10), (i) is equivalent to $\nu_{k,j}(A) = 0$ for $k \neq j$. To see that (ii) follows from rotation-invariance, invoke Proposition 8.2 and apply (13) with $B = e^{i\theta}A$ to obtain

$$\operatorname{tr}\left(P_{k,j}(B, B^*)\right) = e^{i[j-(k-j)]\theta} \operatorname{tr}\left(P_{k,j}(A, A^*)\right). \tag{20}$$

When $j \neq k/2$, this cannot hold (for all θ) unless $\operatorname{tr}\left(P_{k,j}(A, A^*)\right) = 0$. For the converse, note that (20) is automatic when $j = k/2$ and that nilpotence ensures that both sides of (13) are zero also when $j = k$. For $j < k/2$, note that

$$\operatorname{tr}(P_{k,j}(A, A^*)) = \operatorname{tr}\left((P_{k,k-j}(A, A^*))^*\right) = \overline{\operatorname{tr}(P_{k,k-j}(A, A^*))}. \quad \square$$

When $N = 3$, either (i) or (ii) easily implies that the upper-triangular form of $A \in M_3$ with rotation-invariant shadow is

$$\begin{bmatrix} 0 & x & y \\ 0 & 0 & z \\ 0 & 0 & 0 \end{bmatrix},$$

where at least one of x, y, z is zero. For example, the only condition in (ii) is that $\mathrm{tr}(P_{3,2}(A, A^*)) = 0$, i.e. that $\mathrm{tr}(A^2A^*) = 0$, and one easily computes $\mathrm{tr}(A^2A^*) = x\bar{y}z$. Note that A and $B = e^{i\theta}A$ are unitarily similar. One can appeal to Proposition 8.1 to see this or use Sibirskiĭ's list of words (mentioned in Section 6): all the traces are automatically the same for nilpotent A and B except that $\mathrm{tr}(A^2A^*) = \mathrm{tr}(B^2B^*)$ requires $\mathrm{tr}(A^2A^*) = 0$.

For $N = 4$ it is perhaps more convenient to use (ii) to identify those A having rotation-invariant shadow. Let the upper-triangular form of A be

$$\begin{bmatrix} 0 & a & b & c \\ 0 & 0 & x & y \\ 0 & 0 & 0 & z \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

The only conditions in (ii) when $N = 4$ are $\mathrm{tr}(P_{3,2}(A, A^*)) = 0$ and $\mathrm{tr}(P_{4,3}(A, A^*)) = 0$, i.e. $\mathrm{tr}(A^2A^*) = 0$ and $\mathrm{tr}(A^3A^*) = 0$. Computing these traces we find that A has rotation-invariant shadow iff

$$ax\bar{b} + (ay + bz)\bar{c} + x\bar{y}z = 0 \text{ and } ax^2z\bar{c} = 0. \tag{21}$$

Remark 8.4. Although the earlier examples of A with rotation-invariant shadow were also unitarily similar to $e^{i\theta}A$, the analysis (above) of the 4×4 case shows that this is not necessary. If A and $e^{i\theta}A$ are unitarily similar we must have $\mathrm{tr}(A^3(A^*)^2) = 0$, i.e. $ax^2z\overline{(ay + bz)} = 0$, and this does not follow from (21) (e.g. take $c = 0$, $a = b = x = 1$, $y = -1/2$, and $z = 2$).

If $A \in M_N$ has rotation-invariant shadow, the relation (10) simplifies:

$$\sum_{m=0}^{\infty} t^{2m} \frac{(N)_{2m}}{m!\, m!} \nu_{mm}(A) = \det{}^{-1}(I - 2t\mathrm{Re}A) \tag{22}$$

(for all sufficiently small real t), where ReA is the Hermitian $(A + A^*)/2$. If $\lambda_1, \ldots, \lambda_K$ are the nonzero eigenvalues (real) of ReA, the RHS of (22) is $\left(\prod_{k=1}^{K}(1 - 2t\lambda_k)\right)^{-1}$; since the LHS is a function of t^2, these eigenvalues come in $\pm$ pairs. We may assume that $\lambda_1, \ldots, \lambda_p$ are the positive eigenvalues of ReA so that the spectrum of ReA is

$$(\lambda_1, \lambda_2, \ldots, \lambda_p, 0, -\lambda_1, -\lambda_2, \ldots, -\lambda_p),$$

where 0 has multiplicity $N - 2p$. Note that $p \geqslant 1$ unless $A = 0_N$, since Re$A = 0$ implies that A is skew-Hermitian and Proposition 8.2 then implies that $A = 0$. We may therefore write (22) in the following form:

$$\sum_{m=0}^{\infty} t^{2m} \frac{(N)_{2m}}{m!\, m!} \nu_{mm}(A) = \frac{1}{\prod_{j=1}^{p}(1 - 4t^2\lambda_j^2)}. \tag{23}$$

The methods of Section 7 extend most readily to the case where $\lambda_1, \ldots, \lambda_p$ are distinct (as they are for $A = J_N$, where $p = \llcorner \frac{N}{2} \lrcorner$ and $\lambda_j = \cos\left(\frac{j\pi}{N+1}\right)$). The following more general proposition replaces Proposition 7.3.

Proposition 8.5. *If $0 \neq A \in M_N$ has rotation-invariant shadow and the positive eigenvalues of ReA are the distinct $\lambda_1, \ldots, \lambda_p$ then the planar shadow density at each z with $|z| = r$ is given by*

$$f(r) = \frac{1}{\pi} \sum_{k=1}^{p} \frac{\lambda_k^{2(p-2)}}{\prod_{1 \leqslant j \leqslant p, j \neq k}(\lambda_k^2 - \lambda_j^2)} F_N\left(\frac{r^2}{\lambda_k^2}\right), \tag{24}$$

where the function F_N is computable as in Section 7.

Remark 8.6. To see that Proposition 7.3 is a special case of Proposition 8.5, recall from the proof of Proposition 7.1 that

$$(-1)^{k-1}\frac{2^N}{N+1}\sin^2\left(\frac{k\pi}{N+1}\right)=\frac{1}{C_N'\left(\cos\left(\frac{k\pi}{N+1}\right)\right)}$$

where $C_N(x)=\prod_{j=1}^N\left(x-\cos\left(\frac{j\pi}{N+1}\right)\right)$. Thus

$$(-1)^{k-1}\frac{2^N}{N+1}\sin^2\left(\frac{k\pi}{N+1}\right)=\frac{1}{\prod_{1\leqslant j\leqslant N, j\neq k}\left(\cos\left(\frac{k\pi}{N+1}\right)-\cos\left(\frac{j\pi}{N+1}\right)\right)}\ . \tag{25}$$

When $\lambda_j=\cos\left(\frac{j\pi}{N+1}\right), j=1,2,\ldots,N$, the positive values are $\lambda_1,\ldots,\lambda_p$ with $p=\llcorner\frac{N}{2}\lrcorner$.

Suppose first that N is odd; then $p=(N-1)/2$ and for $k\leqslant p$ we have

$$\frac{\lambda_k^{2(p-2)}}{\prod_{1\leqslant j\leqslant p,j\neq k}(\lambda_k^2-\lambda_j^2)}=\frac{\lambda_k^{N-5}}{\prod_{1\leqslant j\leqslant p,j\neq k}(\lambda_k-\lambda_j)(\lambda_k+\lambda_j)}$$

$$=\frac{\lambda_k^{N-5}}{\prod_{1\leqslant j\leqslant p,j\neq k}(\lambda_k-\lambda_j)(\lambda_k-\lambda_{N+1-j})}=\frac{\lambda_k^{N-5}\cdot(\lambda_k-0)\cdot(\lambda_k-\lambda_{N+1-k})}{\prod_{1\leqslant j\leqslant N,j\neq k}(\lambda_k-\lambda_j)}\ ,$$

since $\lambda_{\frac{N+1}{2}}=0$. Thus, in view of (25),

$$\frac{\lambda_k^{2(p-2)}}{\prod_{1\leqslant j\leqslant p,j\neq k}\left(\lambda_k^2-\lambda_j^2\right)}=\lambda_k^{N-5}\cdot\lambda_k\cdot 2\lambda_k(-1)^{k-1}\frac{2^N}{N+1}\sin^2\left(\frac{k\pi}{N+1}\right)$$
$$=(-1)^{k-1}\frac{2^{N+1}}{N+1}\sin^2\left(\frac{k\pi}{N+1}\right)\lambda_k^{N-3},$$

and Proposition 7.3 follows from Proposition 8.5. The argument for even N is similar.

Proof of Proposition 8.5. Since $\lambda_1^2,\ldots,\lambda_p^2$ are distinct,

$$\frac{1}{\prod_{j=1}^p(x-\lambda_j^2)}=\sum_{k=1}^p\frac{b_k}{(x-\lambda_k^2)}\ ,$$

where

$$b_k=\frac{1}{\prod_{j\neq k}(\lambda_k^2-\lambda_j^2)}\ .$$

With $x=1/4t^2$, the RHS of (23) becomes

$$\frac{1}{(4t^2)^p}\sum_{k=1}^p\frac{b_k}{(x-\lambda_k^2)},$$

which we may write as

$$\frac{1}{(4t^2)^{p-1}}\sum_{k=1}^p b_k(1+4t^2\lambda_k^2+(4t^2)^2\lambda_k^4+\cdots).$$

Comparing terms involving t^{2m} with (23) we see that

$$\nu_{mm}(A)=\frac{m!m!}{(N)_{2m}}\sum_{k=1}^p b_k4^m\lambda_k^{2(m+p-1)}.$$

In terms of the radial density f we have

$$\pi \int_0^\infty x^m f\left(\sqrt{x}\right) dx = \frac{m!m!}{\left(\frac{N}{2}\right)_m \left(\frac{N+1}{2}\right)_m 2^{2m}} \sum_{k=1}^{p} b_k 4^m \lambda_k^{2(m+p-1)},$$

so that (in view of the moments that F_N was designed to have)

$$\pi \int_0^\infty x^m f\left(\sqrt{x}\right) dx = \left(\int_0^1 x^m F_N(x)dx\right) \sum_{k=1}^{p} b_k \lambda_k^{2(m+p-1)}$$
$$= \sum_{k=1}^{p} b_k \lambda_k^{2(p-1)} \left(\left(\lambda_k^2\right)^m \int_0^1 x^m F_N(x)dx\right).$$

Applying Proposition 7.2(i),

$$\pi \int_0^\infty x^m f\left(\sqrt{x}\right) dx = \sum_{k=1}^{p} b_k \lambda_k^{2(p-1)} \int_0^{\lambda_k^2} x^m \frac{1}{\lambda_k^2} F_N\left(x/\lambda_k^2\right) dx$$
$$= \int_0^\infty x^m \left(\sum_{k=1}^{p} b_k \lambda_k^{2(p-2)} F_N\left(x/\lambda_k^2\right)\right) dx.$$

Since all moments coincide,

$$f(\sqrt{x}) = \frac{1}{\pi} \sum_{k=1}^{p} b_k \lambda_k^{2(p-2)} F_N(x/\lambda_k^2). \quad \square$$

Remark 8.7. Whether $\lambda_1, \ldots, \lambda_p$ are distinct or not, (23) shows that the shadow measure depends only on the λ_k. Thus $B = \oplus_{k=1}^{p} 2\lambda_k J_2 \oplus 0_{N-2p}$ has the same shadow as A, because the positive eigenvalues of ReB are also $\lambda_1, \ldots, \lambda_p$.

All rotation-invariant numerical shadows are obtained as shadows of the simple superdiagonal matrices

$$B = \oplus_{k=1}^{p} 2\lambda_k J_2 \oplus 0_{N-2p},$$

where $\lambda_1, \ldots, \lambda_p > 0$. For example, $A = J_3$ has the same numerical shadow as

$$B = \begin{bmatrix} 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Here we have another simple example of a pair of matrices with different ranks but the same shadow (compare the discussion at the end of Section 6).

One way to deal with the case of repetitions among $\lambda_1, \ldots, \lambda_p$ is to follow the method of Proposition 8.5 but with the necessarily more complicated partial fraction decomposition. Suppose the distinct values are $\mu_1, \ldots, \mu_n$ and that μ_i occurs with multiplicity k_i; then $p = \sum_1^n k_i$ and $\det^{-1}(I - t(A + A^*))$ is

$$\frac{1}{\prod_{i=1}^{n}(1 - 4t^2\mu_i^2)^{k_i}} = \sum_{i=1}^{n} \sum_{j=0}^{k_i - 1} \frac{\alpha_{ij}}{(1 - 4t^2\mu_i^2)^{k_i - j}}, \tag{26}$$

for certain constants α_{ij}.

Remark 8.8. The α_{ij} are functions of the eigenvalue data. Computationally effective expressions for these functions are available: see [18, pp. 553–562].

Let $R_{N,k}(y)$ be defined by

$$\frac{\Gamma(\frac{N}{2})\Gamma(\frac{N+1}{2})}{\Gamma\left(N-\frac{1}{2}-k\right)(k-1)!}(1-y)^{N-\frac{3}{2}-k}\,H\left(\frac{N+1}{2}-k,\frac{N}{2}-k;N-\frac{1}{2}-k;1-y\right), \tag{27}$$

with the understanding that $R_{N,k}(y)=0$ for $y\geqslant 1$. In view of Proposition 7.4,

$$\int_0^1 y^m R_{N,k}(y)dy=\frac{m!(k)_m}{\left(\frac{N}{2}\right)_m\left(\frac{N+1}{2}\right)_m}. \tag{28}$$

Proposition 8.9. *If $0\neq A\in M_n$ has rotation-invariant shadow and the positive eigenvalues of ReA are distinct $\mu_1,\ldots,\mu_n$ where μ_i has multiplicity k_i, then the planar shadow density at each z with $|z|=r$ is given by*

$$f(r)=\frac{1}{\pi}\sum_{i=1}^{n}\left(\sum_{j=0}^{k_i-1}\alpha_{ij}\frac{1}{\mu_i^2}R_{N,k_i-j}\left(\frac{r^2}{\mu_i^2}\right)\right), \tag{29}$$

where α_{ij} are the constants occurring in (26).

Proof. From (23) we obtain

$$\sum_{m=0}^{\infty}t^{2m}\frac{(N)_{2m}}{m!\,m!}\nu_{mm}(A)=\sum_{i=1}^{n}\left(\sum_{j=0}^{k_i-1}\frac{\alpha_{ij}}{\left(1-4t^2\mu_i^2\right)^{k_i-j}}\right)$$
$$=\sum_{i=1}^{n}\left(\sum_{j=0}^{k_i-1}\alpha_{ij}\sum_{m=0}^{\infty}\frac{(k_i-j)_m}{m!}\left(4t^2\mu_i^2\right)^m\right),$$

where we have used the binomial theorem to express $\left(1-4t^2\mu_i^2\right)^{k_i-j}$ (for small t). Comparing coefficients,

$$\nu_{mm}(A)=\frac{m!}{(N)_{2m}/2^{2m}}\sum_{i=1}^{n}\left(\sum_{j=0}^{k_i-1}\alpha_{ij}(k_i-j)_m(\mu_i^2)^m\right)=\sum_{i=1}^{n}\left(\sum_{j=0}^{k_i-1}\alpha_{ij}\left(\mu_i^2\right)^m\frac{m!(k_i-j)_m}{\left(\frac{N}{2}\right)_m\left(\frac{N+1}{2}\right)_m}\right).$$

In terms of the radial density $f=f_A$, we have

$$\pi\int_0^{\infty}x^m f(\sqrt{x})dx=\sum_{i=1}^{n}\left(\sum_{j=0}^{k_i-1}\alpha_{ij}(\mu_i^2)^m\int_0^1 y^m R_{N,k_i-j}(y)dy\right)$$

(recall (27) and (28)). With the substitutions $x=\mu_i^2 y$, the RHS becomes

$$\sum_{i=1}^{n}\left(\sum_{j=0}^{k_i-1}\alpha_{ij}\frac{1}{\mu_i^2}\int_0^{\mu_i^2}x^m R_{N,k_i-j}\left(\frac{x}{\mu_i^2}\right)dx\right).$$

Because all moments coincide,

$$f(\sqrt{x})=\frac{1}{\pi}\sum_{i=1}^{n}\left(\sum_{j=0}^{k_i-1}\alpha_{ij}\frac{1}{\mu_i^2}R_{N,k_i-j}\left(\frac{x}{\mu_i^2}\right)\right),$$

and (29) follows. □

Remark 8.10. For example, if all λ_k have the same value μ, i.e. $n = 1$, $\mu_1 = \mu$, $k_1 = p$, the model matrix is

$$A = \oplus_1^p 2\mu J_2 \oplus 0_{N-2p}$$

and the only nonzero α_{ij} in (26) is $\alpha_{1,0} = 1$. Thus

$$f(r) = \frac{1}{\pi}\frac{1}{\mu^2}R_{N,p}\left(\frac{r^2}{\mu^2}\right).$$

In particular, if $A = \oplus_1^p 2\mu J_2$ (i.e. $N = 2p$) we have radial density

$$f(r) = \frac{1}{\pi}\frac{1}{\mu^2}R_{2p,p}\left(\frac{r^2}{\mu^2}\right).$$

Recalling (27), we see that

$$\begin{aligned}R_{2p,p}(x) &= \frac{\Gamma(p)\Gamma(p+\frac{1}{2})}{\Gamma(p-\frac{1}{2})(p-1)!}(1-x)^{p-\frac{3}{2}}H\left(\frac{1}{2},0;p-\frac{1}{2};1-x\right)\\ &= \left(p-\frac{1}{2}\right)(1-x)^{p-\frac{3}{2}} = \left(p-\frac{1}{2}\right)\left(\sqrt{1-x}\right)^{2p-3}.\end{aligned}$$

Thus $\oplus_1^p 2\mu J_2$ has radial density

$$\frac{1}{\pi}\frac{1}{\mu^2}\left(p-\frac{1}{2}\right)\left(1-\frac{r^2}{\mu^2}\right)^{p-\frac{3}{2}}.$$

Remark 8.11. There is a useful recurrence relation for the functions $R_{N,k}$. Apply Lemma 7.5 with $a = \frac{N}{2} - k + 1$, $b = N - k - \frac{1}{2}$, $t = 1 - x$ to see that for $N > 2k \geqslant 2$ we have

$$R_{N+1,k}(x) = \frac{N}{(N-2)(N-2k)}\left((2N-2k-3)R_{N,k}(x) - (N-1)(1-x)R_{N-1,k}(x)\right).$$

In using this recurrence relation to compute $R_{N,k}$ one would start with $R_{2k,k}$ and $R_{2k+1,k}$. In Remark 8.10 we saw that $R_{2k,k}(x) = (k - \frac{1}{2})(\sqrt{1-x})^{2k-3}$; it may also be shown that

$$R_{2k+1,k}(x) = \frac{k\left(\frac{1}{2}\right)_k}{2(k-1)!}\left((-x)^{k-1}\log\left(\frac{1+\sqrt{1-x}}{\sqrt{x}}\right) + p_k(x)\sqrt{1-x}\right),$$

where $p_k(x)$ is an explicitly computable polynomial of degree $k-2$.

Further insight into the case of repetitions among the $\lambda_1, \ldots, \lambda_p$ may be gained by considering the limit of (24) from Proposition 8.5 as some of the (initially distinct) λ_k coalesce. This procedure is legitimate in view of the models $A = \oplus_1^p 2\lambda_k J_2$, which always have rotation-invariant shadows. In this approach the theory of divided differences plays an important role. Recall that, given a function $g : [a, b] \to \mathbb{R}$ and distinct $y_1, \ldots, y_p \in [a, b]$, the divided difference

$$g[y_1, \ldots, y_p] = \sum_{k=1}^{p}\frac{g(y_k)}{\prod_{j\neq k}(y_k - y_j)}.$$

We shall appeal to the following facts about such divided differences (compare chapter 4 of [4]):

$g[y_1, \ldots, y_p]$ is invariant under permutations of the y_k;

$$g[y_1, \ldots, y_p] = \frac{g[y_1, \ldots, y_{p-1}] - g[y_2, \ldots, y_p]}{y_1 - y_p}; \tag{30}$$

if g is $p-1$ times continuously differentiable on $[a, b]$ then

$$\lim\{g[y_1, \ldots, y_p] : \text{ all } y_k \to y_0\} = \frac{g^{(p-1)}(y_0)}{(p-1)!}. \tag{31}$$

Setting $y_j = 1/\lambda_j^2$ in Proposition 8.5, we find that

$$f(r) = \frac{1}{\pi}(-1)^{p-1}\left(\prod_1^p y_j\right)g[y_1, \ldots, y_p],$$

where $g(y) = F_N(r^2 y)$. In this approach we see that if the distinct positive eigenvalues of ReA are $\mu_1, \ldots, \mu_n$ with multiplicities $k_1, \ldots, k_n$, then the radial density $f(r)$ for the numerical shadow of A may be computed as

$$f(r) = \frac{(-1)^{p-1}}{\pi}\left(\prod_1^n \mu_i^{2k_i}\right)^{-1} L(k_1, \ldots, k_n),$$

where $p = \sum_1^n k_i$,

$$L(k_1, \ldots, k_n) = \lim\{g[y_1, \ldots, y_p] : y_j \to 1/\mu_i^2 \text{ for } j \in J_i\}, \tag{32}$$

and the J_i partition $\{1, 2, \ldots, p\}$ with $\#(J_i) = k_i$.

The relations (30) and (31) provide us with a sort of L-calculus; for example,

$$L(k, 0, \ldots, 0) = \frac{g^{(k-1)}(1/\mu_1^2)}{(k-1)!} = \frac{r^{2(k-1)}F_N^{(k-1)}(r^2/\mu_1^2)}{(k-1)!},$$

and

$$L(k_1, \ldots, k_n) = \frac{L(k_1, \ldots, k_n - 1) - L(k_1 - 1, \ldots, k_n)}{\frac{1}{\mu_1^2} - \frac{1}{\mu_n^2}}.$$

Using those relations repeatedly, we find that

$$L(k_1, \ldots, k_n) = \sum_{i=1}^n \left(\sum_{j=0}^{k_i - 1} \frac{\beta_{ij} r^{2j}}{j!} F_N^{(j)}(r^2/\mu_i^2)\right) \tag{33}$$

for certain constants β_{ij}. Again (compare Remark 8.8), the β_{ij} are functions of the eigenvalue data.

Summarizing, we have the following alternate method of computing $f_A(r)$.

Proposition 8.12. *If $0 \neq A \in M_n$ has rotation-invariant shadow and the positive eigenvalues of ReA are distinct $\mu_1, \ldots, \mu_n$ where μ_i has multiplicity k_i, then the planar shadow density at each z with $|z| = r$ is given by*

$$f(r) = \frac{(-1)^{p-1}}{\pi}\left(\prod_1^n \mu_i^{2k_i}\right)^{-1} \sum_{i=1}^n \left(\sum_{j=0}^{k_i - 1} \frac{\beta_{ij} r^{2j}}{j!} F_N^{(j)}(r^2/\mu_i^2)\right), \tag{34}$$

where $p = \sum_1^n k_i$ and β_{ij} are the constants found in (33).

Remark 8.13. A comparison of Propositions 8.9 and 8.12 suggests a relation between $R_{N,k}$ and the derivatives $F_N^{(j)}$. Indeed, if

$$A = \oplus_1^p 2J_2 \oplus 0_{N-2p}$$

we have $n = 1, k_1 = p, \mu_1 = 1, \alpha_{1,0} = 1$, and

$$L(p) = \frac{r^{2(p-1)}}{(p-1)!} F_N^{(p-1)}(r^2),$$

i.e. $\beta_{1,p-1} = 1$ and all other $\beta_{ij} = 0$. The two forms for the radial density f_A tell us that

$$R_{N,p}(r^2) = (-1)^{p-1} \frac{r^{2(p-1)}}{(p-1)!} F_N^{(p-1)}(r^2),$$

i.e.

$$R_{N,k}(x) = (-1)^{k-1} \frac{x^{k-1}}{(k-1)!} F_N^{(k-1)}(x). \tag{35}$$

This relation between the $R_{N,k}$ and the derivatives of F_N $(= R_{N,1})$ may also be obtained directly by using the identity

$$H(a, b; c; t) = (1-t)^{c-a-b} H(c-a, c-b; c; t).$$

Remark 8.14. A study of the behavior of the radial density $f_A(x)$ (when $0 \neq A \in M_N$ has rotation-invariant shadow) near $x = 0$ reveals that it has dominant singularity $x^{p-1} \log x$ there (recall that p is the number of positive eigenvalues of ReA, counted with multiplicity) unless $N = 2p$, in which case $f_A(x)$ is analytic near $x = 0$.

9. Numerical shadows via the Cartesian decomposition

In this section we discuss aspects of the numerical shadow related to the so-called Cartesian decomposition of a matrix A into its Hermitian components ReA and ImA = Re$(-iA)$. For example, we investigate the shadow of a possibly nonnormal matrix by means of projections onto lines in $\mathbb{C}$. These projections have interpretations as shadows of Hermitian matrices and can also be thought of as Radon transforms of the shadow. We discuss how the eigenvalues of the projections are involved in the analysis of the map $\Omega_N \to \mathbb{C}$ taking u to (Au, u). We remark that, in general, the shadow measure of a nonnormal matrix is absolutely continuous with respect to area measure on $\mathbb{C} \equiv \mathbb{R}^2$ (see [16]).

9.1. Marginal densities

Recall that the numerical range of a Hermitian matrix is real and the density of the shadow is a nonnegative spline function, straightforward to express in terms of the eigenvalues. This fact can be exploited by means of a type of Cartesian decomposition.

Recall that for $A \in M_N$ we define the real part of A by

$$\text{Re}A = \frac{1}{2}\left(A + A^*\right).$$

Thus ReA is Hermitian, and $A = \text{Re}A + i\text{Re}\,(-iA)$. We will be concerned with the more general $\text{Re}\left(e^{-i\theta}A\right)$ with $-\pi \leqslant \theta \leqslant \pi$; then A can be expressed as

$$A = e^{i\theta}\text{Re}\left(e^{-i\theta}A\right) + ie^{i\theta}\text{Re}\left(-ie^{-i\theta}A\right). \tag{36}$$

For $-\pi \leqslant \theta \leqslant \pi$ let $\lambda_k(\theta)$ be the eigenvalues of Re$(e^{-i\theta}A)$, labeled so that

$$\lambda_1(\theta) \leqslant \lambda_2(\theta) \leqslant \ldots \leqslant \lambda_N(\theta).$$

Then for each $u \in \Omega_N$ we have $\lambda_1(\theta) \leqslant \text{Re}(e^{-i\theta}(Au, u)) \leqslant \lambda_N(\theta)$ and

$$\lambda_j(\theta + \pi) = -\lambda_{N+1-j}(\theta) \quad (1 \leqslant j \leqslant N).$$

A matrix with the property that $\lambda_1(\theta) < \lambda_2(\theta) < \ldots < \lambda_N(\theta)$ for $0 \leqslant \theta \leqslant 2\pi$ is called *generic* in the paper of Jonckheere, Ahmad and Gutkin [21].

We relate the shadow of $\text{Re}\left(e^{-i\theta}A\right)$ to a marginal density of P_A. Write the shadow measure $dP_A(z) = p_A(z)\,dm_2(z)$ (where dm_2 is the Lebesgue measure on $\mathbb{C} \equiv \mathbb{R}^2$). Recall: if $f(x, y)$ is a density function on $\mathbb{R}^2$ with compact support, then the marginal density along the x-axis is $f_X(x) = \int_{-\infty}^{\infty} f(x, y)\,dy$. Suppose $g(x)$ is a continuous function; then

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)\,dx = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x) f(x, y)\,dxdy.$$

Thus the moments, $\int_{-\infty}^{\infty} x^n f_X(x)\, dx$ equal $E[X^n]$ with respect to the density f. Now replace x by $x\cos\theta + y\sin\theta = \mathrm{Re}\left(e^{-i\theta}z\right)$ for some fixed θ. The line orthogonal to $\mathrm{Re}\left(e^{-i\theta}z\right) = 0$ is $\mathrm{Re}\left(ie^{-i\theta}z\right) = 0$. Let $u = \mathrm{Re}\left(e^{-i\theta}z\right)$, $v = \mathrm{Re}\left(ie^{-i\theta}z\right)$; then $z = e^{i\theta}(u - iv)$ and so the density of the shadow of $\mathrm{Re}\left(e^{-i\theta}A\right)$ is the marginal density of P_A for u:

$$f_U(u) = \int_{-\infty}^{\infty} p_A\left(e^{i\theta}(u - iv)\right) dv.$$

This is exactly the (2-dimensional) Radon transform of p_A evaluated at (u, θ), $u \in \mathbb{R}$, $-\pi \leqslant \theta \leqslant \pi$. We may restate this as follows: suppose $g(u)$ is real and continuous for $u \in \mathbb{R}$; then $E[g(U)]$ with respect to $P_{\mathrm{Re}(e^{-i\theta}A)}$ is

$$\begin{aligned}\int_{\Omega_N} g\left(\mathrm{Re}\left(e^{-i\theta}(Aw, w)\right)\right) d\mu(w) &= \int_{W(A)} g\left(\frac{1}{2}\left(e^{-i\theta}z + e^{i\theta}\overline{z}\right)\right) p_A(z)\, dm_2(z)\\ &= \int_{-\infty}^{\infty} g(u) \int_{-\infty}^{\infty} p_A\left(e^{i\theta}(u - iv)\right) dv du = E\left[g\left(\mathrm{Re}\left(e^{-i\theta}Z\right)\right)\right],\end{aligned}$$

where the latter expectation is with respect to P_A.

Remark. The Radon transform can be inverted to recover the shadow of A from the shadows of $\mathrm{Re}\left(e^{-i\theta}A\right)$, $-\pi \leqslant \theta \leqslant \pi$. There are practical algorithms, used in X-ray tomography, which produce approximations to the inverse transform by using a finite number of angles θ (see also [17, Ch. 1, Sect. 2]). The Radon transform approach is worked out thoroughly in [16].

For $A \in M_N$ let $\xi_A(s, t) = \det(I - sA - tA^*)$ and for a Hermitian matrix H let $\xi_H(r) = \det(I - rH)$ ("ξ" suggests "characteristic"). For a power series h, $\left[s^j t^k\right] h(s, t)$ denotes the coefficient of $s^j t^k$ in $h(s, t)$, and $[r^n]\, h(r)$ denotes to coefficient of r^n in $h(r)$, $j, k, n = 0, 1, 2, \ldots$.

Recall from Proposition 5.3 that the moments of A can be obtained from ξ_A,

$$\nu_{jk}(A) := \int_{W(A)} z^j \overline{z}^k dP_A(z) = \frac{j!k!}{(N)_{j+k}} \left[s^j t^k\right] \xi_A(s, t)^{-1}.$$

The central moments of a probability distribution are also of interest. Let $m_A = \frac{1}{N}\mathrm{tr}A$, then $E[Z] = m_A$. The central moments can be computed by expanding the integrand in

$$\nu^0_{jk}(A) := \int_{W(A)} (z - m_A)^j \left(\overline{z} - \overline{m_A}\right)^k dP_A(z),$$

or by using the shifted matrix $A - m_A I$.

Lemma. *For $A \in M_N$ and $c \in \mathbb{C}$,*

$$\xi_{A-cI}(s, t) = (1 + sc + t\overline{c})^N \xi_A\left(\frac{s}{1 + sc + t\overline{c}}, \frac{t}{1 + sc + t\overline{c}}\right).$$

Proof. Indeed,

$$\begin{aligned}\xi_{A-cI}(s, t) &= \det\left(I - s(A - cI) - t\left(A^* - \overline{c}I\right)\right)\\ &= \det\left((1 + sc + t\overline{c})\, I - sA - tA^*\right) = (1 + sc + t\overline{c})^N \det\left(I - \frac{sA + tA^*}{1 + sc + t\overline{c}}\right). \qquad \square\end{aligned}$$

It is clear that the shadow of $A - cI$ is a translate of P_A. Thus

$$\nu^0_{jk}(A) = \frac{j!k!}{(N)_{j+k}} \left[s^j t^k\right] \xi_{A - m_A I}(s, t)^{-1}, j, k = 0, 1, 2, \ldots.$$

We consider the (one-dimensional) moments of $\mathrm{Re}\left(e^{-i\theta}A\right)$.

Proposition 9.1. *For $n = 0, 1, 2, \ldots$,*

$$\int_{-\infty}^{\infty} u^n dP_{\mathrm{Re}\left(e^{-i\theta}A\right)}(u) = \frac{n!}{(N)_n}\left[r^n\right]\xi_A\left(\frac{1}{2}re^{-i\theta}, \frac{1}{2}re^{i\theta}\right)^{-1}.$$

Proof. The integral

$$\begin{aligned}\int_{-\infty}^{\infty} u^n dP_{\mathrm{Re}\left(e^{-i\theta}A\right)}(u) &= \frac{1}{2^n}\int_{W(A)}\left(e^{-i\theta}z + e^{i\theta}\overline{z}\right)^n dP_A(z)\\ &= \frac{1}{2^n}\sum_{j=0}^{n}\binom{n}{j}e^{i\theta(n-2j)}\int_{W(A)} z^j\overline{z}^{n-j}dP_A(z)\\ &= \frac{1}{2^n}\sum_{j=0}^{n}\binom{n}{j}e^{i\theta(n-2j)}\frac{j!\,(n-j)!}{(N)_n}\left[s^jt^{n-j}\right]\xi_A(s,t)^{-1}\\ &= \frac{n!}{(N)_n}\left[r^n\right]\xi_A\left(\frac{1}{2}re^{-i\theta}, \frac{1}{2}re^{i\theta}\right)^{-1}. \qquad \square\end{aligned}$$

That is, the moments of $P_{\mathrm{Re}\left(e^{-i\theta}A\right)}$ can be obtained from $\xi_{\mathrm{Re}\left(e^{-i\theta}A\right)}(r)$. Furthermore, $\xi_{\mathrm{Re}\left(e^{-i\theta}A\right)}(r) = \det\left(I - r\mathrm{Re}\left(\left(e^{-i\theta}A\right)\right)\right) = \prod_{j=1}^{N}\left(1 - r\lambda_j(\theta)\right)$.

Here are the basic quantities associated to $P_{\mathrm{Re}\left(e^{-i\theta}A\right)}$:

The mean of $P_{\mathrm{Re}\left(e^{-i\theta}A\right)}$ is

$$\frac{1}{2N}\left(e^{-i\theta}\mathrm{tr}A + e^{i\theta}\mathrm{tr}A^*\right) = \mathrm{Re}\left(e^{-i\theta}m_A\right) = \frac{1}{N}\sum_{j=1}^{N}\lambda_j(\theta).$$

The variance of $P_{\mathrm{Re}\left(e^{-i\theta}A\right)}$ is

$$\begin{aligned}&\frac{2}{(N)_2}\left[r^2\right]\xi_{\mathrm{Re}\left(e^{-i\theta}A\right)}(r)^{-1} - \mathrm{Re}\left(e^{-i\theta}m_A\right)^2\\ &= \frac{1}{N(N+1)}\left(\sum_{j=1}^{N}\lambda_j(\theta)^2 - \frac{1}{N}\left(\sum_{j=1}^{N}\lambda_j(\theta)\right)^2\right)\\ &= \frac{1}{4N(N+1)}\left(e^{-2i\theta}\mathrm{tr}\left((A - m_AI)^2\right) + 2\mathrm{tr}\left((A - m_AI)(A^* - \overline{m_A}I)\right) + e^{2i\theta}\mathrm{tr}\left((A^* - \overline{m_A}I)^2\right)\right).\end{aligned}$$

Let $\mathrm{tr}\left((A - m_AI)^2\right) = ae^{i\phi}$ with $a \geqslant 0$; then the variance of $P_{\mathrm{Re}\left(e^{-i\theta}A\right)}$ is maximized at $\theta = \frac{\phi}{2}$ and minimized at $\theta = \frac{\phi}{2} \pm \frac{\pi}{2}$. There is a relation with the 2-dimensional variance of P_A, namely,

$$\int_{W(A)}|z - m_A|^2\, dP_A(z) = \frac{1}{N(N+1)}\mathrm{tr}\left((A - m_AI)\left(A^* - \overline{m_A}I\right)\right).$$

The central moments of $\mathrm{Re}\left(e^{-i\theta}A\right)$ can be obtained from

$$\xi_{\mathrm{Re}\left(e^{-i\theta}A\right) - \mathrm{Re}\left(e^{-i\theta}m_A\right)I}(r) = \left(1 + \mathrm{Re}\left(e^{-i\theta}m_A\right)r\right)^N\xi_{\mathrm{Re}\left(e^{-i\theta}A\right)}\left(\frac{r}{1 + r\mathrm{Re}\left(e^{-i\theta}m_A\right)}\right).$$

9.2. The shadow of a Hermitian matrix

The density function for P_H is simple to find, given the eigenvalues of a Hermitian matrix H. Suppose H is not scalar; then H has at least two different eigenvalues and the shadow is absolutely continuous on $\mathbb{R}$. Let $dP_H(x) = p_H(x)\,dx$. The following is the basic fact.

Lemma 9.2. *Suppose* $1 \leqslant m < N$ *and* $\lambda \neq 0$. *For* $n = 0, 1, 2, \ldots$

$$m\binom{N-1}{m}\lambda^{1-N}\int_0^{\lambda} x^n x^{m-1}(\lambda - x)^{N-m-1}\,dx = \frac{(m)_n}{(N)_n}\lambda^n \quad (\lambda > 0);$$

$$m\binom{N-1}{m}(-\lambda)^{1-N}\int_{\lambda}^{0} x^n(-x)^{m-1}(x-\lambda)^{N-m-1}\,dx = \frac{(m)_n}{(N)_n}\lambda^n \quad (\lambda < 0).$$

To express the density functions for all real arguments we use the notation

$$x_+ = \max(x, 0),$$

with the convention that $x_+^0 = 1$ *for* $x \geqslant 0$ *and* $= 0$ *for* $x < 0$. *Thus the density in the first part of the lemma equals* $m\binom{N-1}{m}\lambda^{1-N}x_+^{m-1}(\lambda - x)_+^{N-m-1}$ *for* $x \in \mathbb{R}$.

Suppose H is Hermitian, not a multiple of I, and $\xi_H(r) = \prod_{i=1}^{N}(1 - r\lambda_i) = \prod_{i=1}^{M}(1 - r\mu_i)^{m_i}$, where $\{\mu_1, \ldots, \mu_M\}$ is the set of distinct nonzero eigenvalues of H and $\sum_{i=1}^{M} m_i \leqslant N$. By hypothesis, H has at least two different eigenvalues (each $m_i < N$). There are unique real numbers β_{ij}, $1 \leqslant i \leqslant M$, $1 \leqslant j \leqslant m_i$ such that

$$\frac{1}{\xi_H(r)} = \sum_{i=1}^{M}\sum_{j=1}^{m_i}\frac{\beta_{ij}}{(1 - r\mu_i)^j} = \sum_{i=1}^{M}\sum_{j=1}^{m_i}\beta_{ij}\sum_{n=0}^{\infty}\frac{(j)_n}{n!}\mu_i^n r^n, \quad |r| < \frac{1}{\min_i|\mu_i|}.$$

Thus for $n = 0, 1, 2, \ldots$

$$\int_{-\infty}^{\infty} x^n dP_H(x) = \sum_{i=1}^{M}\sum_{j=1}^{m_i}\beta_{ij}\frac{(j)_n}{(N)_n}\mu_i^n.$$

By the lemma, the density function of P_H is

$$p_H(x) = \sum_{\mu_i > 0}\sum_{j=1}^{m_i}\mu_i^{1-N}\beta_{ij}j\binom{N-1}{j}x_+^{j-1}(\mu_i - x)_+^{N-j-1}$$
$$+ \sum_{\mu_i < 0}\sum_{j=1}^{m_i}(-\mu_i)^{1-N}\beta_{ij}j\binom{N-1}{j}(-x)_+^{j-1}(x - \mu_i)_+^{N-j-1}.$$

9.3. The critical curves in W(A)

Suppose that in some interval $\theta_1 < \theta < \theta_2$ the eigenvalues of $\mathrm{Re}\left(e^{-i\theta}A\right)$ are pairwise distinct, and there are eigenvectors $\psi^{(j)}(\theta)$ so that

$$\mathrm{Re}\left(e^{-i\theta}A\right)\psi^{(j)}(\theta) = \lambda_j(\theta)\,\psi^{(j)}(\theta),$$

where

$$\left|\psi^{(j)}(\theta)\right| = 1 \quad (1 \leqslant j \leqslant N).$$

The image $\left\{(A\psi^{(j)}(\theta), \psi^{(j)}(\theta)) : \theta_1 < \theta < \theta_2\right\}$ is called a critical curve (see [21, Theorem 5, p. 238]).

Lemma 9.3. *Suppose $H(\theta), \psi(\theta), \lambda(\theta)$ are differentiable functions on $\theta_1 < \theta < \theta_2$ such that $H(\theta)$ is Hermitian, $\psi(\theta) \in \Omega_N$, $\lambda(\theta) \in \mathbb{R}$, and $H(\theta)\psi(\theta) = \lambda(\theta)\psi(\theta)$; then $\frac{d}{d\theta}\lambda(\theta) = ((\frac{d}{d\theta}H(\theta))\psi(\theta), \psi(\theta))$.*

Proof. Write $\lambda(\theta) = \psi(\theta)^* H(\theta)\psi(\theta)$ and differentiate to obtain

$$\begin{aligned}\frac{d}{d\theta}\lambda(\theta) &= \psi(\theta)^*\left(\frac{d}{d\theta}H(\theta)\right)\psi(\theta) + \frac{d}{d\theta}(\psi(\theta)^*)H(\theta)\psi(\theta) + \psi(\theta)^*H(\theta)\frac{d}{d\theta}\psi(\theta)\\ &= \psi(\theta)^*\left(\frac{d}{d\theta}H(\theta)\right)\psi(\theta) + \lambda(\theta)\left(\frac{d}{d\theta}(\psi(\theta)^*)\psi(\theta) + \psi(\theta)^*\frac{d}{d\theta}\psi(\theta)\right)\\ &= \psi(\theta)^*\left(\frac{d}{d\theta}H(\theta)\right)\psi(\theta),\end{aligned}$$

because $\psi(\theta)^*\psi(\theta) = 1$. □

Proposition 9.4. *For $\theta_1 < \theta < \theta_2$ the critical curve satisfies $\left(A\psi^{(j)}(\theta), \psi^{(j)}(\theta)\right) = \psi^{(j)}(\theta)^* A\psi^{(j)}(\theta) = e^{i\theta}\left(\lambda_j(\theta) + i\lambda_j'(\theta)\right)$.*

Proof. Let $H(\theta) = \mathrm{Re}\left(e^{-i\theta}A\right) = (\cos\theta)A_1 + (\sin\theta)A_2$, where $A_1 = \frac{1}{2}(A + A^*)$ and $A_2 = \frac{1}{2i}(A - A^*)$. Thus

$$\psi^{(j)}(\theta)^* A\psi^{(j)}(\theta) = \psi^{(j)}(\theta)^* A_1\psi^{(j)}(\theta) + i\psi^{(j)}(\theta)^* A_2\psi^{(j)}(\theta).$$

By definition and the lemma

$$\psi^{(j)}(\theta)^*((\cos\theta)A_1 + (\sin\theta)A_2)\psi^{(j)}(\theta) = \lambda_j(\theta),$$

and

$$\psi^{(j)}(\theta)^*(-(\sin\theta)A_1 + (\cos\theta)A_2)\psi^{(j)}(\theta) = \lambda_j'(\theta).$$

These equations are easily solved to establish the formula. □

The analyticity of $\lambda_j(\theta)$ is shown in [21, Lemma 2, p. 240].

To get an idea of the structure of critical curves one has to distinguish the generic and non-generic cases. If A is generic then $\lambda_j(\theta + \pi) = -\lambda_{N+1-j}(\theta)$ for $1 \leqslant j \leqslant N$, the curve $C_j := \left\{e^{i\theta}\left(\lambda_j(\theta) + i\lambda_j'(\theta)\right) : 0 \leqslant \theta \leqslant 2\pi\right\}$ agrees with C_{N+1-j} (as a point-set) so there are $\left\lfloor\frac{N+1}{2}\right\rfloor$ critical curves (see [21, Theorem 13, p. 244]); the outside C_1 is the boundary of Λ_A. Example 9.5.3 below is a generic 3×3 matrix.

When using numerical techniques for solving the characteristic equation for some number of angles (for example $\theta = j\pi/m, j = 0, \ldots, m$) the value of $\lambda'(\theta)$ can be computed as follows: let $p(\theta, \lambda) = \det\left(\lambda I - \mathrm{Re}\left(e^{-i\theta}A\right)\right)$, differentiate the equation $p(\theta, \lambda(\theta)) = 0$ to obtain $\frac{\partial}{\partial\theta}p(\theta, \lambda(\theta)) + \lambda'(\theta)\frac{\partial}{\partial\lambda}p(\theta, \lambda(\theta)) = 0$ (so the value of $\lambda(\theta)$ determines $\lambda'(\theta)$ except possibly for isolated points where $\frac{\partial}{\partial\lambda}p(\theta, \lambda(\theta)) = 0$; this indicates repeated roots which do not occur in the generic case).

In the non-generic case the same critical curve can arise from different eigenvalues: let θ_0 be an angle for which the eigenvalues are all distinct and ordered by $\lambda_1(\theta_0) < \ldots < \lambda_N(\theta_0)$; consider each $\lambda_i(\theta)$ as a real-analytic function in θ and extend it to the interval $\theta_0 \leqslant \theta \leqslant \theta_0 + \pi$. Because this is the non-generic case the curves $\lambda_i(\theta)$ may cross in the open interval (a finite number of times by analyticity). Form a set-partition of $\{1, 2, \ldots, N\}$ by declaring i and j equivalent if $\lambda_i(\theta_0) = -\lambda_j(\theta_0 + \pi)$; the relation is extended by transitivity. The equivalence classes correspond to distinct critical curves. There may be only one class; consider Example 9.5.2. In this case the boundary of $W(A)$ is the convex hull of the outside critical curve (from the class containing 1).

In the situation of radially symmetric shadows (see Section 8) the critical curves are circles centered at the origin.

9.4. A geometric approach

Any matrix can be expressed as a sum of two matrices with orthogonal 1-dimensional numerical ranges. For a fixed θ with $-\frac{\pi}{2} \leqslant \theta \leqslant \frac{\pi}{2}$ we can write

$$A = e^{i\theta}\,\mathrm{Re}\left(e^{-i\theta}A\right) + ie^{i\theta}\,\mathrm{Re}\left(e^{-i(\theta+\pi/2)}A\right)$$

Let $U_1, U_2 \in \mathcal{U}(N)$ satisfy $U_1^*\,\mathrm{Re}\left(e^{-i\theta}A\right)U_1 = B_1$, $U_2^*\,\mathrm{Re}\left(e^{-i(\theta+\pi/2)}A\right)U_2 = B_2$ where B_1, B_2 are diagonal matrices, such that $(B_1)_{jj} = \lambda_j(\theta)$ and $(B_2)_{jj} = \lambda_j\left(\theta + \frac{\pi}{2}\right)$ for $1 \leqslant j \leqslant N$. Then for $\psi \in \Omega_N$

$$\psi^*A\psi = e^{i\theta}\psi^*A_1\psi + ie^{i\theta}\psi^*A_2\psi = e^{i\theta}\psi^*U_1B_1U_1^*\psi + ie^{i\theta}\psi^*U_2B_2U_2^*\psi.$$

By the unitary invariance of the range (and the shadow) we may replace (generic) ψ by (generic) $U_1\psi$. Thus

$$(U_1\psi)^*A(U_1\psi) = e^{i\theta}\sum_{j=1}^{N}\lambda_j(\theta)\,|\psi_j|^2 + ie^{i\theta}\sum_{j=1}^{N}\lambda_j\left(\theta+\frac{\pi}{2}\right)\left|\left(U_2^*U_1\psi\right)_j\right|^2. \tag{37}$$

The value remains unchanged if $U_2^*U_1$ is replaced by $U = D_2U_2^*U_1D_1$ where D_1, D_2 are arbitrary diagonal unitary matrices (in M_N). For example, choose D_1, D_2 so that $U_{1,j} \geqslant 0$ and $U_{j,1} \geqslant 0$ for $1 \leqslant j \leqslant N$. Thus the numerical range and shadow can be interpreted in terms of a mapping from

$$T_{N-1}[U] := \left\{\left(\left(|\psi_j|^2\right)_{j=1}^{N}, \left(|U\psi_j|^2\right)_{j=1}^{N}\right) : \psi \in \Omega_N\right\} \subset \Delta_N \times \Delta_N$$

to $\mathbb{C}$. Every vector $\left(|\psi_j|^2\right)_{j=1}^{N}$ appears as a first and as a second component of a point in $T_{N-1}[U]$. The unitarily invariant measure on Ω_N induces a measure on $T_{N-1}[U]$, and the shadow of A is the image of this measure under the map

$$\left(\left(|\psi_j|^2\right)_{j=1}^{N}, \left(|U\psi_j|^2\right)_{j=1}^{N}\right) \mapsto e^{i\theta}\sum_{j=1}^{N}\left(\lambda_j(\theta)\,|\psi_j|^2 + i\lambda_j\left(\theta+\frac{\pi}{2}\right)\left|(U\psi)_j\right|^2\right).$$

The case $N = 2$ can be explicitly described. Let

$$U = \begin{pmatrix}\cos\theta_0 & \sin\theta_0\\ \sin\theta_0 & -\cos\theta_0\end{pmatrix}, \quad \psi = \begin{bmatrix}e^{i\phi_1}\cos\theta_1\\ e^{i\phi_2}\sin\theta_1\end{bmatrix},$$

with $0 \leqslant \theta_0, \theta_1 \leqslant \frac{\pi}{2}$ and $-\pi \leqslant \phi_1, \phi_2 \leqslant \pi$. It suffices to consider U of this form (θ_0 is fixed). Then

$$|U\psi_1|^2 = \frac{1}{2} + \frac{1}{2}\cos 2\theta_1\cos 2\theta_0 + \frac{1}{2}\sin 2\theta_1\sin 2\theta_0\cos(\phi_1-\phi_2),$$

and

$$|U\psi_2|^2 = \frac{1}{2} - \frac{1}{2}\cos 2\theta_1\cos 2\theta_0 - \frac{1}{2}\sin 2\theta_1\sin 2\theta_0\cos(\phi_1-\phi_2).$$

Thus $T_1[U]$ is

$$\left\{\left(\frac{1}{2}+\frac{1}{2}x_1(\theta_1), \frac{1}{2}-\frac{1}{2}x_1(\theta_1)\right),\right.$$
$$\left.\left(\frac{1}{2}+\frac{1}{2}x_2(\theta_1,\theta_2,\phi), \frac{1}{2}-\frac{1}{2}x_2(\theta_1,\theta_2,\phi)\right) : 0 \leqslant \theta_1 \leqslant \frac{\pi}{2}, 0 \leqslant \phi \leqslant \pi\right\},$$

where $x_1(\theta_1) = \cos 2\theta_1$ and

$$x_2(\theta_1, \theta_0, \phi) = \cos 2\theta_1 \cos 2\theta_0 + \sin 2\theta_1 \sin 2\theta_0 \cos\phi.$$

As expected (compare Section 2) this forms an ellipse (including the interior). Changing coordinates we transform to the square $\{(x_1, x_2) : -1 \leqslant x_1, x_2 \leqslant 1\}$; then $T_1[U]$ maps to $\left\{(x_1, x_2) : x_1^2 - 2x_1x_2\cos 2\theta_0 + x_2^2 \leqslant \sin^2 2\theta_0\right\}$. In the degenerate normal case this reduces to the interval $\{(x_1, x_1) : -1 \leqslant x_1 \leqslant 1\}$.

The invariant measure on Ω_2 is $\frac{1}{2\pi^2}\sin\theta_1\cos\theta_1 d\theta_1 d\phi_1 d\phi_2$. This is mapped to the measure $\frac{1}{2\pi}\left(\sin^2 2\theta_0 - x_1^2 + 2x_1x_2\cos 2\theta_0 - x_2^2\right)^{-\frac{1}{2}} dx_1 dx_2$.

9.5. Examples

Example 9.5.1. Let

$$A_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Then $\xi_{A_1}(s,t) = 1 - 3st + s^2t^2$ and $\xi_{\mathrm{Re}(e^{-i\theta}A_1)}(r) = 1 - \frac{3}{4}r^2 + \frac{1}{16}r^4$. The eigenvalues of $\mathrm{Re}\left(e^{-i\theta}A_1\right)$ are $\frac{1}{4}\left(\pm 1 \pm \sqrt{5}\right)$, independent of θ, labeled so that $\lambda_1 < \lambda_2 < 0 < \lambda_3 < \lambda_4$. Thus the density for $\mathrm{Re}(e^{-i\theta}A_1)$ is given by

$$\begin{aligned} P_1(x) &= 6\left(1 - \frac{\sqrt{5}}{5}\right) \times \left((-x)_+^0 (x - \lambda_1)_+^2 + x_+^0 (\lambda_4 - x)_+^2\right) \\ &\quad - 6\left(1 + \frac{\sqrt{5}}{5}\right)\left((-x)_+^0 (x - \lambda_2)_+^2 + x_+^0 (\lambda_3 - x)_+^2\right). \end{aligned}$$

In fact the shadow has circular symmetry, as discussed in Sections 7 and 8. The critical curves are $z = \frac{1}{4}\left(1 + \sqrt{5}\right)e^{i\theta}$ and $z = \frac{1}{4}\left(\sqrt{5} - 1\right)e^{i\theta}$, $0 \leqslant \theta \leqslant 2\pi$.

Example 9.5.2. Let

$$A_2 = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Then

$$\xi_{A_2}(s,t) = 1 - 6st - 4st(s+t) - st\left(s^2 + st + t^2\right)$$

and

$$\xi_{\mathrm{Re}(e^{-i\theta}A_2)}(r) = \left(1 + r - r^2\left(\frac{1}{4} - \frac{1}{2}\cos\theta\right)\right)\left(1 - r - r^2\left(\frac{1}{4} + \frac{1}{2}\cos\theta\right)\right).$$

The eigenvalues of $\mathrm{Re}\left(e^{-i\theta}A_2\right)$ are $\frac{1}{2} \pm \cos\frac{\theta}{2},\ -\frac{1}{2} \pm \sin\frac{\theta}{2}$. The eigenvalues are $\left[\frac{3}{2}, -\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}\right]$ at $\theta = 0$ and $\pm\frac{1}{2} \pm \frac{\sqrt{2}}{2}$ at $\theta = \frac{\pi}{2}$. In the range $-\frac{\pi}{2} \leqslant \theta \leqslant \frac{\pi}{2}$ we have

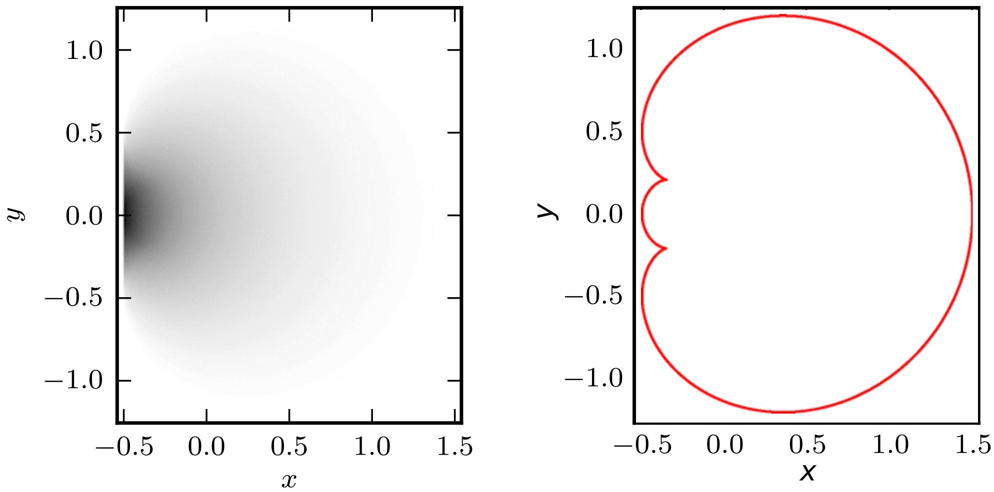

Fig. 4. Numerical shadow for matrix A_2 of size $N = 4$ and the corresponding critical lines.

$$\lambda_4(\theta) = \frac{1}{2} + \cos\frac{\theta}{2}, \quad \lambda_1(\theta) = -\frac{1}{2} - \left|\sin\frac{\theta}{2}\right|,$$

and so $-\frac{1}{2} - \left|\sin\frac{\theta}{2}\right| \leqslant \mathrm{Re}\left(e^{-i\theta}\psi^* A_2 \psi\right) \leqslant \frac{1}{2} + \cos\frac{\theta}{2}$ for $\psi \in \Omega_N$. The triple eigenvalue $-\frac{1}{2}$ at $\theta = 0$ results in a pronounced peak in the density for $\mathrm{Re}(A_2)$:

$$P_2(x) = 6\left(\frac{1}{2}+x\right)_+ (-x)_+^2 + 9\left(\frac{1}{2}+x\right)_+ (-x)_+ + \frac{27}{8}\left(\frac{1}{2}+x\right)_+^2 (-x)_+^0 + \frac{3}{8}x_+^0\left(\frac{3}{2}-x\right)_+^2.$$

The matrix A_2 is non-generic and there is only one critical curve which has two cusps as shown in Fig. 4. In this example the boundary of the shadow is the convex hull of the critical curve, so the line segment $[(-1-i)/2, (-1+i)/2]$ is a part of the boundary. One representation of critical lines is $z = e^{i\theta}\left(\frac{1}{2} + \cos\frac{\theta}{2} - \frac{i}{2}\sin\frac{\theta}{2}\right)$, $0 \leqslant \theta \leqslant 4\pi$. The cusps are at $\dfrac{-19 \pm 5\sqrt{5}}{54}$.

Example 9.5.3. Let

$$A_3 = \begin{pmatrix} 0 & 1 & 1 \\ 0 & i & 1 \\ 0 & 0 & -1 \end{pmatrix}.$$

Then $m_{A_3} = \frac{1}{3}(-1+i)$,

$$\xi_{A_3}(s,t) = 1 + (1-i)s + (1+i)t - i\left(s^2 - t^2\right) - 3st - st\left((2-i)s + (2+i)t\right),$$

$$\xi_{A_3 - m_{A_3}I}(s,t) = 1 - \frac{13}{3}st - \frac{i}{3}\left(s^2 - t^2\right) + \frac{5}{27}\left((1+i)s^3 + (1-i)t^3\right)$$
$$-\frac{st}{9}\left((8+i)s + (8-i)t\right),$$

and

$$\xi_{\mathrm{Re}(e^{-i\theta}A_3)}(r) = 1 + (\cos\theta - \sin\theta)r - \left(\frac{3}{4} + \sin\theta\cos\theta\right)r^2 + \left(\frac{1}{4}\sin\theta - \frac{1}{2}\cos\theta\right)r^3.$$

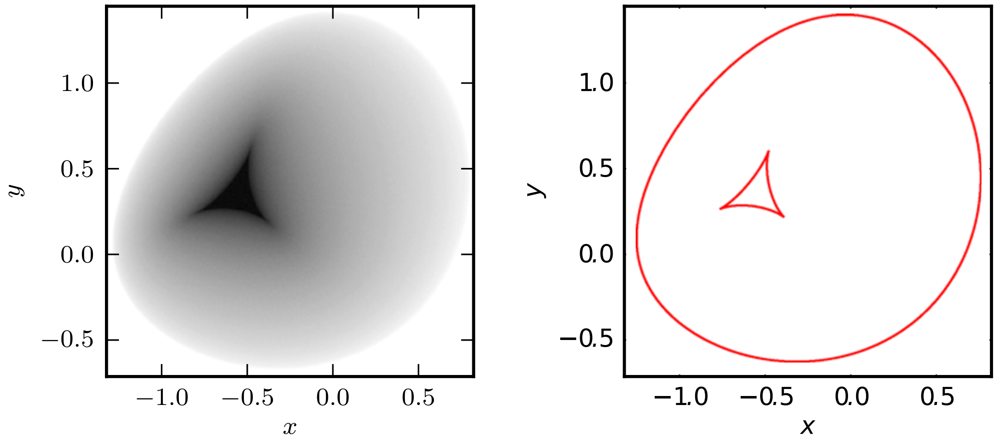

Fig. 5. Numerical shadow for matrix A_3 of size $N = 3$ and the corresponding critical lines.

The variance of $P_{\mathrm{Re}(e^{-i\theta}A_3)}$ equals $\frac{1}{72}(13 + 2\sin 2\theta)$. The eigenvalues of $\mathrm{Re}\left(e^{-i\theta}A_3\right)$ can be found approximately, or analytically by the classical formula. For $\theta = 0$ the eigenvalues are $\left[-\frac{1+\sqrt{17}}{4}, -\frac{1}{2}, \frac{-1+\sqrt{17}}{4}\right]$ for $\theta = 0$ and $[-0.6715, 0.2647, 1.407]$ (rounded) for $\theta = \frac{\pi}{2}$. The matrix is generic; there are two critical curves: one is the boundary of Λ_{A_3} and the other is a triangular curve with three cusps. A comparison of the numerical shadow for this matrix and its critical lines is presented in Fig. 5. Since $N = 3$ the density for $\mathrm{Re}\left(e^{-i\theta}A_3\right)$ is piecewise linear:

$$P_3(x) = (-x)_+^0\left(x + \frac{1+\sqrt{17}}{4}\right)_+ - 2(-x)_+^0\left(x + \frac{1}{2}\right)_+ + \left(1 - \frac{1}{\sqrt{17}}\right)x_+^0\left(\frac{\sqrt{17}-1}{4} - x\right)_+ .$$

10. Direct sums (block diagonal matrices)

Concerning the direct sum $A \oplus B$ (block diagonal matrix) of matrices A, B, it is well-known that

$$W(A \oplus B) = \mathrm{conv}\{W(A) \cup W(B)\}$$

(see, for example, [2, Exercise I.3.1]). In our context it is natural to ask how the numerical shadow of $A \oplus B$ is distributed over $\mathrm{conv}\{W(A) \cup W(B)\}$. Here A and B may be of different sizes – see an example presented in Fig. 6. We consider then $A \oplus B \in M_N$ with $A \in M_n$ and $B \in M_m$, so that $n + m = N$. Given $u \in \Omega_N$ (distributed according to μ, as usual), let $u = v_1 \oplus v_2$ where $v_1 \in \mathbb{C}^n$ and $v_2 \in \mathbb{C}^m$; then $\|v_1\|^2 + \|v_2\|^2 = 1$. It is known that $t = \|v_1\|^2$ has a beta-density given by

$$q(t) = \frac{(n+m-1)!}{(n-1)!\,(m-1)!}t^{n-1}(1-t)^{m-1} \quad (t \in [0,1]). \tag{38}$$

From this one can deduce that the shadow measure $P_{A\oplus B}$ is an "(n, m)-beta mixture" of the shadow measures P_A and P_B. Compare [16, Section 2.2].

Another version of this result relates the densities corresponding to A, B, and $A \oplus B$:

Proposition 10.1. *If $p_A(z)$, $p_B(z)$ are the shadow densities for A, B, then the corresponding density p for $A \oplus B$ is given by*

$$p(z) = \int_0^1 q(t)\Big(\int_{\mathbb{C}} t^{-2}p_A((z-w)/t)(1-t)^{-2}p_B\,(w/(1-t))\;dm_2(w)\Big)\,dt, \tag{39}$$

where $q(t)$ is as in (38).

Proof. For $u \in \Omega_N$ we have

$$((A \oplus B)u, u) = (Av_1, v_1) + (Bv_2, v_2) = \|v_1\|^2(Au_1, u_1) + \|v_2\|^2(Bu_2, u_2),$$

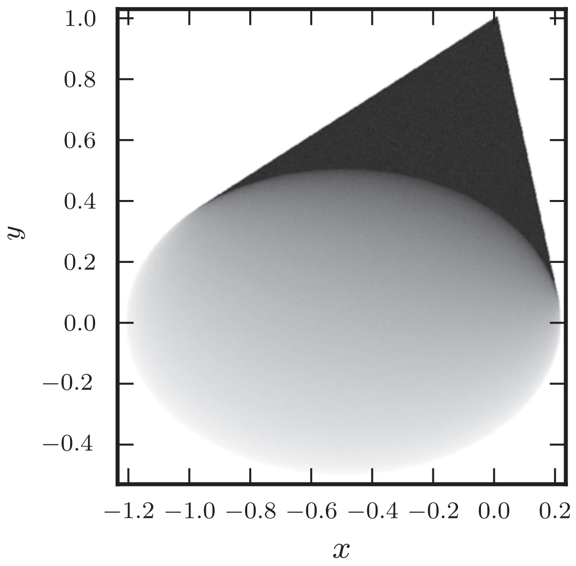

Fig. 6. Shadow of block diagonal matrix $\begin{pmatrix} -1 & 0 \\ 1 & 0 \end{pmatrix} \oplus (i)$.

where $u_j = v_j/\|v_j\|$. Note that $u_1 \in \Omega_n$, $u_2 \in \Omega_m$, u_1, u_2 are stochastically independent, and they have the corresponding uniform distributions over Ω_n, Ω_m.

Hence $((A \oplus B)u, u) = tZ_1 + (1-t)Z_2$ where Z_1 and Z_2 are independent complex random variables with densities $p_A(z)$ and $p_B(z)$. Thus

$$p(z) = \int_0^1 q(t)g(z,t)\,dt$$

where $g(z,t)$ is the density of the independent (for each fixed t) sum $tZ_1 + (1-t)Z_2$. This density is given by the usual convolution formula

$$g(z,t) = \int_{\mathbb{C}} g_1(z-w)g_2(w)\,dm_2(w),$$

where g_1 is the density of tZ_1 and g_2 is the density of $(1-t)Z_2$. If a complex random variable Z has density $h(z)$ with respect to area on $\mathbb{C}$, then tZ (where $t \in \mathbb{R}$) has density $t^{-2}h(z/t)$. Hence $g_1(z-w) = t^{-2}p_A((z-w)/t)$, $g_2(w) = (1-t)^{-2}p_B(w/(1-t))$ and (39) follows. $\square$

11. Zernike expansions

Given our methods for evaluating the moments of shadow measures (see Section 5), it is natural to construct orthogonal polynomial approximations using Zernike polynomials. These provide one way to generate pictures of specific numerical shadows.

The complex Zernike polynomials $Z_{mn}(z,\bar{z})$ are orthogonal for area measure on the unit disk $\{z \in \mathbb{C} : |z| \leqslant 1\}$. They can be defined by

$$Z_{mn}(z,\bar{z}) = z^{m-n}\sum_{j=0}^{n}\frac{(m+n-j)!}{(m-j)!\,(n-j)!j!}(-1)^j(z\bar{z})^{n-j} \quad (m \geqslant n)$$

$$Z_{mn}(z,\bar{z}) = \bar{z}^{n-m}\sum_{j=0}^{m}\frac{(m+n-j)!}{(m-j)!\,(n-j)!j!}(-1)^j(z\bar{z})^{m-j} \quad (m < n),$$

and satisfy the orthogonality relations

$$\frac{1}{\pi}\int\int_{|z|<1} Z_{mn}(z,\bar{z})\,\overline{Z_{kl}(z,\bar{z})}dm_2(z) = \frac{\delta_{mk}\delta_{nl}}{m+n+1}.$$

Suppose $f(z,\overline{z})$ is continuous on the disk and has coefficients

$$\widehat{f}_{mn} := \int\int_{|z|<1} f(z,\overline{z})\,\overline{Z_{mn}(z,\overline{z})}dm_2(z), \quad m,n=0,1,2,\ldots$$

then

$$f = \frac{1}{\pi}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}(m+n+1)\widehat{f}_{mn}Z_{mn},$$

with convergence at least in the L^2-sense. As is typical of Fourier expansions, the convergence behaviour is better for smoother functions f. If f is real then $\widehat{f}_{mn} = \overline{\widehat{f}_{nm}}$.

Suppose A is an $N \times N$ matrix whose numerical range is contained in the unit disk (otherwise work with $B = c_1A + c_0I$, with $c_1 > 0$ so that $\mathrm{tr}(B) = 0$ and the range of B satisfies the boundedness condition). We may use Proposition 5.3 to determine the moments of the shadow P_A (and we write $dP_A(z) = p_A(z)\,dm_2(z)$, so that p_A is the density). Thus

$$\int_{|z|\leqslant 1} z^m\overline{z}^n dP_A(z) = \frac{m!n!}{(N)_{m+n}}\left[s^mt^n\right]\xi_A(s,t)^{-1},$$

where $\xi_A(s,t) = \det(I - sA - tA^*)$ and $[s^mt^n]\,g(s,t)$ denotes the coefficient of s^mt^n in the power series expansion of g centered at $(s,t) = (0,0)$.

It is then straightforward to compute the Zernike coefficients of the density:

$$\begin{aligned}(p_A)\hat{}_{mn} &= \int_{|z|\leqslant 1}\overline{Z_{mn}(z,\overline{z})}p_A(z)\,dm_2(z)\\ &= \int_{|z|\leqslant 1}\sum_{j=0}^{n}\frac{(m+n-j)!}{(m-j)!\,(n-j)!j!}(-1)^j z^{n-j}\overline{z}^{m-j}dP_A(z)\\ &= \sum_{j=0}^{n}\frac{(m+n-j)!}{j!\,(N)_{m+n-2j}}(-1)^j\left[s^{n-j}t^{m-j}\right]\xi_A(s,t)^{-1},\end{aligned}$$

for $m \geqslant n$ (and $(p_A)\hat{}_{nm} = \overline{(p_A)\hat{}_{mn}}$). As an approximation, one may compute $(p_A)\hat{}_{mn}$ for all (m,n) with $m + n \leqslant M$ for some M (say 10 or 20). To write our formulas in real terms with $z = x + iy$ (and $x^2 + y^2 \leqslant 1$), let

$$Q_{mn}(u) = \sum_{j=0}^{\min(m,n)}\frac{(m+n-j)!}{(m-j)!\,(n-j)!j!}(-1)^j u^{\min(m,n)-j}.$$

Note the trivial identity $ab + \overline{a}\overline{b} = 2\mathrm{Re}a\mathrm{Re}b - 2\mathrm{Im}a\mathrm{Im}b$. Thus the partial sum for p_A can be written as:

$$\begin{aligned}&\sum_{j=0}^{\lfloor M/2\rfloor}(2j+1)\,(p_A)\hat{}_{jj}\,Q_{jj}\left(x^2+y^2\right)\\ &+2\sum_{j=1}^{M}\left(\mathrm{Re}\left((x+\mathrm{i}y)^j\right)\sum_{n=0}^{\lfloor (M-j)/2\rfloor}(2n+j+1)\,\mathrm{Re}\left((p_A)\hat{}_{n+j,n}\right)Q_{n+j,n}\left(x^2+y^2\right)\right)\\ &-2\sum_{j=1}^{M}\left(\mathrm{Im}\left((x+\mathrm{i}y)^j\right)\sum_{n=0}^{\lfloor (M-j)/2\rfloor}(2n+j+1)\,\mathrm{Im}\left((p_A)\hat{}_{n+j,n}\right)Q_{n+j,n}\left(x^2+y^2\right)\right).\end{aligned}$$

(The factor π has been ignored; it is merely a change of scale). It is a matter for experimentation to produce useful graphs for a given matrix. The polynomials tend to wiggle close to the edge of the disk; the graphs can not be expected to precisely show the boundary of the numerical range, but they do indicate the behaviour of p_A in the interior.

12. Numerical shadows and the higher-rank numerical ranges

The rank-k numerical ranges, denoted below by Λ_k, were introduced c. 2006 by Choi, Kribs, and Życzkowski as a tool to handle compression problems in quantum information theory. Since then their theory and applications have been advanced with remarkable enthusiasm. The sequence of papers [7,8,28,24], for example, led to a striking extension of the classical Toeplitz–Hausdorff theorem (convexity of $W(M)$): **all** the $\Lambda_k(M)$ are convex (though some may be empty), and they are intersections of conveniently computable half-planes in $\mathbb{C}$. Among the many more recent papers concerning the $\Lambda_k(M)$, let us mention [23,14].

Given a matrix $M \in M_N$ and $k \geqslant 1$, Choi, Kribs, and Życzkowski (see [5,6]) defined the rank-k numerical range of M as

$$\Lambda_k(M) = \{\lambda \in \mathbb{C} : \exists P \in P_k \text{ such that } PMP = \lambda P\},$$

where P_k denotes the set of rank-k orthogonal projections in M_N. It is not hard to verify that $\Lambda_k(M)$ can also be described as the set of complex λ such that there is some k-dimensional subspace S of $\mathbb{C}^N$ such that $(Mu, u) = \lambda$ for **all** unit vectors in S. In particular, we see that

$$W(M) = \Lambda_1(M) \supseteq \Lambda_2(M) \supseteq \Lambda_3(M) \supseteq \cdots \quad .$$

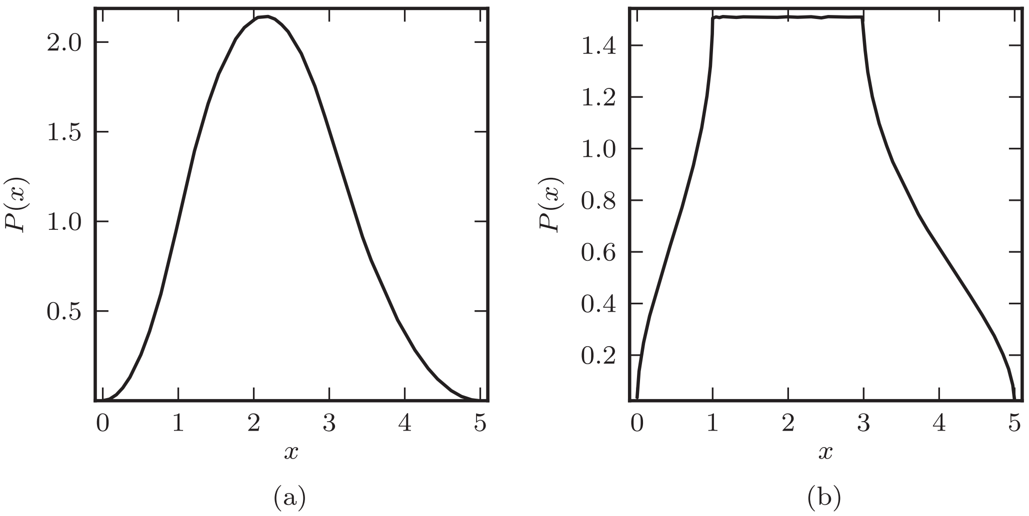

Fig. 7. (a) Shadow of diag(0, 1, 3, 5). (b) Shadow of diag(0, 1, 3, 5) with respect to real vectors.

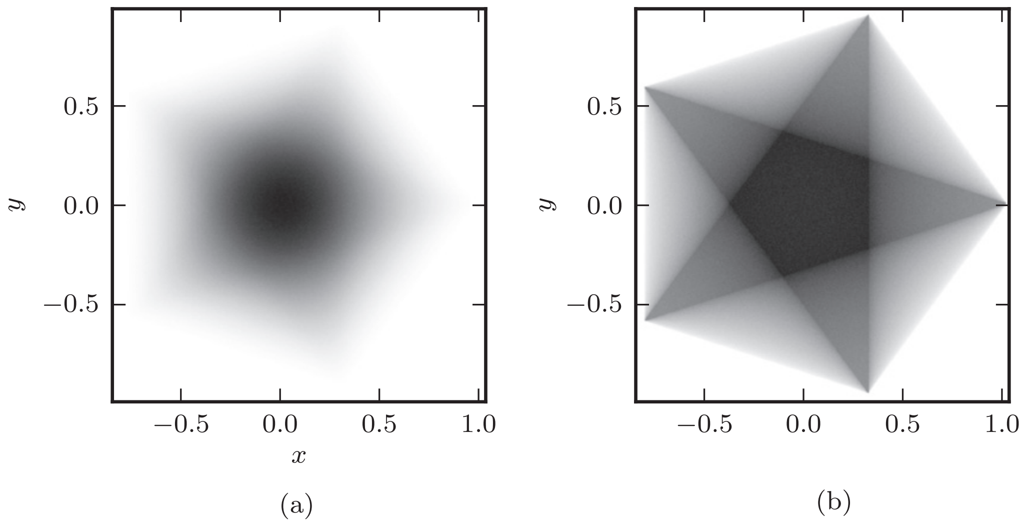

Fig. 8. (a) Shadow of unitary matrix in M_5. (b) Shadow of unitary matrix in M_5 with respect to real vectors.

This point of view also suggests that these higher-rank numerical ranges should be visible as regions of higher density within the numerical shadow of M. This idea is borne out, to some degree, by examining shadow densities of various matrices. In Fig. 7(a), for example, we see the (one-dimensional) shadow density of the Hermitian diag(0, 1, 3, 5) – a spline of degree 2. Here it is known that $\Lambda_2(M) = [1, 3]$; while the density is unimodal, there are values in [3,5] that are greater than some in [1,3]. If **real** unit vectors are used in such experiments, the higher-rank numerical ranges often seem to be revealed more clearly; compare Fig. 7(b).

A similar phenomenon is seen in Fig. 8. Here M is a unitary matrix in M_5 and it is known that $\Lambda_2(M)$ is the inner pentagon (with interior) formed by lines joining the non-adjacent eigenvalues in pairs. The shadow density is unimodal, but $\Lambda_2(M)$ is only seen distinctly in Fig. 8(b), where only real unit vectors in $u \in \Omega_5$ are use to generate the values (Mu, u). The distinction between shadows based on complex vs real unit vectors is a reflection of the fact that the latter follow a Dirichlet distribution (with parameter 1/2) rather than our usual measure μ. See [3].

Acknowledgements

Work by J. Holbrook was supported in part by an NSERC of Canada research grant. Work by P. Gawron and Z. Puchała was supported by the Polish Ministry of Science and Higher Education under the Grant Number N519 442339, while K. Życzkowski acknowledges support by the Grant Number N202 090239. Numerical calculations presented in this work were performed on the `Leming` server of The Institute of Theoretical and Applied Informatics, Polish Academy of Sciences.

References

[1] T.W. Anderson, The Statistical Analysis of Time Series, Wiley, New York, 1971.
[2] R. Bhatia, Matrix Analysis, Springer-Verlag, New York, 1997.
[3] I. Bengtsson, K. Zÿ yczkowski, Geometry of Quantum States, Cambridge UP, Cambridge, 2006.
[4] E.W. Cheney, D.R. Kincaid, Numerical Mathematics and Computing, Brooks/Cole, 1985.
[5] M.-D. Choi, D. Kribs, K. Życzkowski, Quantum error correcting codes from the compression formalism, Rep. Math. Phys. 58 (2006) 77–86.
[6] M.-D. Choi, D. Kribs, K. Życzkowski, Higher-rank numerical ranges and compression problems, Linear Algebra Appl. 418 (2006) 828–839.
[7] M.-D. Choi, J. Holbrook, D. Kribs, K. Życzkowski, Higher-rank numerical ranges of unitary and normal matrices, Oper. Matrices 1 (2007) 409–426.
[8] M.–D. Choi, M. Giesinger, J. Holbrook, D. Kribs, Geometry of higher-rank numerical ranges, Linear and Multilinear Algebra 56 (2008) 53–64.
[9] C.F. Dunkl, P. Gawron, J.A. Holbrook, J.A. Miszczak, Z. Puchała, K. Życzkowski, Numerical shadow as a tool to study geometry of quantum states, in preparation.
[10] K.R. Davidson, J. Holbrook, Numerical radii of zero-one matrices, Michigan Math. J. 35 (1988) 261–267.
[11] D.Ž. Djoković, C.R. Johnson, Unitarily achievable zero patterns and traces of words in A and A^*, Linear Algebra Appl. 421 (2007) 63–68.
[12] C. Davis, The Toeplitz–Hausdorff theorem explained, Canad. Math. Bull. 14 (1971) 245–246.
[13] C. de Boor, Splines as linear combinations of *B*-splines, in: G.G. Lorentz, C.K. Chui, L.L. Schumaker (Eds.), Approximation Theory II, Academic Press, New York, 1976, pp. 1–47.
[14] H.-L. Gau, C.-K. Li, P.Y. Wu, Higher-rank numerical ranges and dilations, J. Operator Theory 63 (2010) 181–189.
[15] K.E. Gustafson, D.K.M. Rao, Numerical Range, Springer, 1997.
[16] T. Gallay, D. Serre, The numerical measure of a complex matrix, arXiv:1009.1522v1[math.FA], 8 September 2010.
[17] S. Helgason, Groups and Geometric Analysis, Academic Press, New York, 1984.
[18] P. Henrici, Applied and Complex Analysis, vol. 1, Wiley-Interscience, New York, 1974.
[19] J. Holbrook, Diagonal compressions of matrices and numerical shadows, colloquium, February 19, U. of Hawaii, 2010.
[20] J. Holbrook, J.-P. Schoch, Theory vs. experiment: multiplicative inequalities for the numerical radius of commuting matrices, Oper. Theory Adv. Appl. 202 (2010) 273–284.
[21] E. Jonckheere, F. Ahmad, E. Gutkin, Differential topology of numerical range, Linear Algebra Appl. 279 (1998) 227–254.
[22] C.-K. Li, A survey on linear preservers of numerical ranges and radii, Taiwanese J. Math. 5 (2001) 477–496.
[23] C.-K. Li, Y.-T. Poon, N.-S. Sze, Condition for the higher-rank numerical range to be non-empty, Linear and Multilinear Algebra 57 (2009) 365–368.
[24] C.-K. Li, N.-S. Sze, Canonical forms, higher rank numerical ranges, totally isotropic subspaces, and matrix equations, Proc. Amer. Math. Soc. 136 (2008) 3013–3023.
[25] I.G. Macdonald, Symmetric Functions and Hall Polynomials, second ed., Clarendon Press, Oxford, 1995.
[26] K.-C. Ng, Some properties of doubly-stochastic matrices and distribution of density on a numerical range, MPhil Thesis, University of Hong Kong, 1982.
[27] W. Specht, Zur Theorie der Matrixen II, Jahresber. Deutsche Math. 50 (1940) 19–23.

[28] H. Woerdeman, The higher rank numerical range is convex, Linear and Multilinear Algebra 56 (2008) 65–67.
[29] K. Życzkowski, Volume of the set of separable states II, Phys. Rev. A60 (1999) 3496–3507.
[30] K. Życzkowski (with M.-D. Choi, C. Dunkl, J. Holbrook, P. Gawron, J.Miszczak, Z. Puchala, L. Skowronek), Generalized numerical range as a versatile tool to study quantum entanglement, Oberwolfach, December 2009 (see Oberwolfach Report No. 59/2009, 34-37), 2009.