The numerical measure of a complex matrix

Thierry Gallay* Denis Serre†

September 8, 2010

Abstract

We introduce and carefully study a natural probability measure over the numerical range of a complex matrix $A \in \mathbf{M}_n(\mathbb{C})$. This *numerical measure* μ_A can be defined as the law of the random variable $\langle AX, X\rangle \in \mathbb{C}$ when the vector $X \in \mathbb{C}^n$ is uniformly distributed on the unit sphere. If the matrix A is normal, we show that μ_A has a piecewise polynomial density f_A, which can be identified with a multivariate B-spline. In the general (nonnormal) case, we relate the Radon transform of μ_A to the spectrum of a family of Hermitian matrices, and we deduce an explicit representation formula for the numerical density which is appropriate for theoretical and computational purposes. As an application, we show that the density f_A is polynomial in some regions of the complex plane which can be characterized geometrically, and we recover some known results about lacunas of symmetric hyperbolic systems in $2+1$ dimensions. Finally, we prove under general assumptions that the numerical measure of a matrix $A \in \mathbf{M}_n(\mathbb{C})$ concentrates to a Dirac mass as the size n goes to infinity.

1 Introduction

If $A \in \mathbf{M}_n(\mathbb{C})$ is a complex square matrix of size $n \in \mathbb{N}^*$, the *numerical range* of A is the compact subset of the complex plane defined by

$$W(A) \,=\, \Big\{\langle Ax, x\rangle \in \mathbb{C} \,\Big|\, x \in \mathbb{C}^n\,,\ \|x\| = 1\Big\}\,,$$

where $\langle x, y\rangle = y^* x$ is the usual scalar product in $\mathbb{C}^n$ and $\|x\| = \langle x, x\rangle^{1/2}$. It is quite obvious that $W(A) \supset \sigma(A)$, where $\sigma(A)$ (the *spectrum* of A) is the collection of all eigenvalues of A, and that $W(A) = W(U^*AU)$ for any unitary matrix $U \in \mathbf{U}_n(\mathbb{C})$. Moreover, a celebrated result due to Toeplitz [24] and Hausdorff [13] asserts that $W(A)$ is always a convex subset of the complex plane. In particular, $W(A)$ contains the convex hull of $\sigma(A)$, and it is easy to verify that $W(A) = \mathrm{conv}(\sigma(A))$ if the matrix A is normal, namely $AA^* = A^*A$. The interested reader is referred to Chapter 1 of [15] for a detailed discussion of the various properties of the numerical range, including complete proofs.

Let $\partial\mathbb{B}^n = \{x \in \mathbb{C}^n \,|\, \|x\| = 1\}$ be the unit sphere in $\mathbb{C}^n$, considered as a real manifold of dimension $2n-1$. By definition, the numerical range $W(A)$ is the image of the *numerical map* $\Phi_A : \partial\mathbb{B}^n \to \mathbb{C}$ defined by

$$\Phi_A(x) \,=\, \langle Ax, x\rangle\,, \qquad x \in \partial\mathbb{B}^n\,.$$

*Université Joseph Fourier, Institut Fourier (UMR CNRS 5582), BP 74, 38402 St-Martin-d'Hères, FRANCE.
†École Normale Supérieure de Lyon, UMPA (UMR CNRS 5669), 46, allée d'Italie, 69364 Lyon, FRANCE.

The algebraic and geometric properties of the map Φ_A have been extensively studied, see [19, 26, 4, 9, 16, 17]. In particular, the set of all critical values of Φ_A, which we denote by $\Sigma_A \subset \mathbb{C}$, has received a lot of attention, because this is an interesting object which contains a lot of information on the matrix A. For instance, it is known that $\partial W(A) \subset \Sigma_A$ and $W(A) = \mathrm{conv}(\Sigma_A)$. In addition, there exists a real algebraic curve $C_A \subset \mathbb{C} \simeq \mathbb{R}^2$ with the property that $\Sigma_A = C_A \cup C'_A$, where C'_A denotes the set of all line segments joining pairs of points of C_A at which C_A has the same tangent line [17]. Under generic assumptions on A, the bitangent set C'_A is empty, and the critical set Σ_A is therefore the union of a finite number of closed curves, one of which is the boundary of the numerical range $W(A)$. This distinguished curve is smooth, and encloses all the other ones in its interior. We refer to Section 5 below for more details on the geometry of the singular set, and to Section 7 for a few concrete examples.

Our purpose in this paper is to introduce another mathematical quantity which is naturally related to the numerical map Φ_A. Given $A \in \mathbf{M}_n(\mathbb{C})$, the *numerical measure* of A is the probability measure μ_A on $\mathbb{C}$ defined by the formula

$$\int_{\mathbb{C}} \phi(z)\,\mathrm{d}\mu_A(z) \;=\; \int_{\partial\mathbb{B}^n} \phi(\langle Ax, x\rangle)\,\mathrm{d}\bar\sigma(x)\;, \tag{1}$$

for all continuous functions $\phi : \mathbb{C} \to \mathbb{C}$. Here $\bar\sigma$ denotes the Euclidian measure on the unit sphere $\partial\mathbb{B}^n$, normalized as a probability measure. In words, the numerical measure is thus the image under the numerical map of the normalized Euclidean measure on the unit sphere. Equivalently, if X is a random variable that is uniformly distributed on $\partial\mathbb{B}^n$, the numerical measure μ_A is just the distribution of the random variable $\langle AX, X\rangle \in \mathbb{C}$. This probabilistic intepretation will be useful later, especially in Section 8.

Our first goal is to establish a few general properties of the numerical measure μ_A. It is clear by construction that μ_A is invariant under unitary conjugations of A, namely $\mu_{U^*AU} = \mu_A$ for all $U \in \mathbf{U}_n(\mathbb{C})$. This is precisely the reason why we used the Euclidean measure on $\partial\mathbb{B}^n$ in the definition (1). It is also easy to verify that the support of μ_A is exactly the numerical range $W(A)$, see Section 2 below. Less obvious, perhaps, is the fact that μ_A is absolutely continuous with respect to the Lebesgue measure λ on $W(A)$, so that we can define the *numerical density* f_A as the Radon-Nikodym derivative of μ_A with respect to λ (in the particular situation where $W(A)$ reduces to a line segment Γ, we understand λ as the one-dimensional Lebesgue measure on Γ, see Section 2.) We also prove that the numerical density f_A is strictly positive in the interior of A, a property that can be interpreted as a strong version of Hausdorff's theorem [13]. Finally, we shall see that the singular support of μ_A is contained in the critical set Σ_A, which means that the numerical density f_A is smooth outside Σ_A. In fact, we conjecture that $\mathrm{sing\,supp}(\mu_A) = \Sigma_A$ for all $A \in \mathbf{M}_n(\mathbb{C})$, but this has not been proved yet.

After these general properties have been established, our next goal is to give a more precise description of the numerical density f_A. For this purpose, it is convenient to distinguish between various cases:

1. (*The scalar case*) If $W(A)$ is reduced to a single point $\{z\}$, then $A = zI_n$ (where $I_n \in \mathbf{M}_n(\mathbb{C})$ denotes the identity matrix) and $\mu_A = \delta_z$. In this trivial situation, there is of course no need to introduce a numerical density.

2. (*The Hermitian case*) Assume that $n \geq 2$ and that $W(A) \subset \mathbb{C}$ is a line segment. Then we can find $z \in \mathbb{C}$, $\theta \in [0, 2\pi]$, and a Hermitian matrix H such that $A = zI_n + e^{i\theta}H$. If $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ are the eigenvalues of H, we shall see in Section 3 that the numerical measure μ_H is absolutely continuous with respect to Lebesgue's measure on $\mathbb{R}$, and that the corresponding density f_H is exactly the normalized

B-spline of degree $n-2$ with knots $\lambda_1, \ldots, \lambda_n$ [7]. In particular, f_H is polynomial of degree $n-2$ on each interval $[\lambda_i, \lambda_{i+1}]$, vanishes identically outside $[\lambda_1, \lambda_n]$, and is continuous at each point λ_i together with its derivatives up to order $d_i = n-2-m_i$, where $m_i \ge 1$ is the multiplicity of λ_i as an eigenvalue of H (if $d_i < 0$, then f_H is discontinuous at λ_i.) This gives an explicit representation of the numerical measure μ_H, and the measure μ_A is the image of μ_H under the affine isometry $w \mapsto z + e^{i\theta} w$.

3. (*The normal case*) Suppose now that $n \ge 3$ and that $A \in \mathbf{M}_n(\mathbb{C})$ is a normal matrix whose spectrum $\sigma(A)$ is not contained in a line segment. Then $W(A) = \mathrm{conv}(\sigma(A))$ is a convex polygon with nonempty interior, and it turns out that the numerical density f_A is the *bivariate B-spline* of degree $n-3$ whose knots are the eigenvalues of A. Here we refer to the work of W. Dahmen [6] for the definition and the main properties of multivariate B-splines. In this particular case, the critical set Σ_A is thus the collection of all line segments joining pairs of eigenvalues of A, and the density f_A is polynomial of degree $n-3$ in each connected component of $\mathbb{C} \setminus \Sigma_A$. In the generic situation where no straight line contains more than two eigenvalues of A, one can show that f_A is continuous together with its derivatives up to order $n-4$ (and is discontinuous on $\partial W(A)$ if $n = 3$.)

4. (*The nonnormal case*) Finally, we consider the most interesting situation where the matrix $A \in \mathbf{M}_n(\mathbb{C})$ is not normal. In that case, there is no explicit formula for the numerical density, but the problem can be reduced in some sense to the Hermitian case by the following simple observation. For any $\theta \in S^1 = \mathbb{R}/(2\pi\mathbb{Z})$, let $H(\theta)$ be the Hermitian matrix defined by

$$H(\theta) \;=\; \frac{1}{2}(e^{-i\theta}A + e^{i\theta}A^*) \;=\; A_1 \cos(\theta) + A_2 \sin(\theta)\;, \tag{2}$$

where $A_1 = (A + A^*)/2$ and $A_2 = (A - A^*)/(2i)$. Then $\mathrm{Re}\,(e^{-i\theta}\langle Ax, x\rangle) = \langle H(\theta)x, x\rangle$ for all $x \in \partial\mathbb{B}^n$. Now, if the random variable X is uniformly distributed on $\partial\mathbb{B}^n$, the distribution of $\langle H(\theta)X, X\rangle$ is by definition the numerical measure $\mu_{H(\theta)}$, whereas the distribution of $\mathrm{Re}\,(e^{-i\theta}\langle AX, X\rangle)$ is easily identified as the two-dimensional *Radon transform* of the numerical measure μ_A, evaluated at $\theta \in [0, 2\pi]$. We thus have

$$\mathcal{R}\mu_A(\theta) \;=\; \mu_{H(\theta)}\;, \quad \theta \in S^1\;, \tag{3}$$

where $\mathcal{R}$ denotes the two-dimensional Radon transformation. Since the numerical density of $H(\theta)$ is known to be the B-spline based on the eigenvalues $\lambda_1(\theta), \ldots, \lambda_n(\theta)$ of $H(\theta)$, we can reconstruct the numerical measure μ_A by inverting the Radon transformation in (3), using the the well-known *backprojection* method which plays an important role in tomography [12]. This provides a useful representation formula for the numerical density, as well as an efficient algorithm for numerical calculations, see Section 4 for more details.

It is worth mentioning here that the critical set Σ_A can be conveniently characterized using the family of Hermitian matrices $H(\theta)$ associated with A. Indeed, if we define the eigenvalues $\lambda_j(\theta)$ in such a way that they depend analytically on θ, one can shown that the algebraic curve C_A which generates Σ_A is given by

$$C_A \;=\; \Big\{ e^{i\theta}(\lambda_j(\theta) + i\lambda_j'(\theta)) \,\Big|\, j \in \{1, \ldots, n\}\,,\; \theta \in [0, \pi] \Big\}\;,$$

see [26, 16, 17] and Section 5 below. In the generic case where $\lambda_1(\theta) < \lambda_2(\theta) < \cdots < \lambda_n(\theta)$ for all $\theta \in [0, \pi]$, the bitangent set C_A' is empty and $\Sigma_A = C_A$.

As was already mentioned, the numerical density f_A is smooth (in fact, real-analytic) on each connected component of $\mathbb{C} \setminus \Sigma_A$. The regularity across Σ_A is more difficult to study, but we shall show in Section 6.2 that f_A is everywhere of class C^{n-3} if $n \ge 3$ and $A \in \mathbf{M}_n(\mathbb{C})$ satisfies some generic

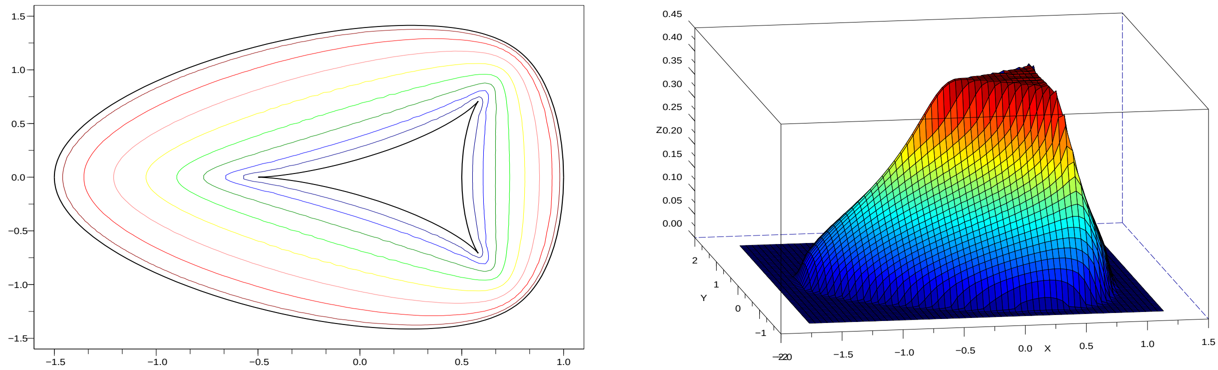

Figure 1: The numerical density f_A is represented for a typical matrix $A \in \mathbf{M}_3(\mathbb{R})$, given by (55) below. In the contour plot (left), the exterior ovate curve is the boundary of the numerical range $W(A)$, and the other component of the of the critical set Σ_A is the interior cuspidal triangle. The three-dimensional plot (right) confirms that the numerical density is continuous, positive inside the numerical range, and constant over the cuspidal triangle, in agreement with the results of Section 6.

hypotheses, which exclude in particular the case of normal matrices. In addition, for an arbitrary matrix of size n, we shall prove that all derivatives of f_A of order $n-2$ vanish identically in some distinguished regions, which have the following geometric characterization. For any $z \in \mathbb{C} \setminus \Sigma_A$, let $N(z)$ be the number of straight lines containing z which are tangent to the curve C_A, see (39) below for a precise definition where possible multiplicities are taken into account. It is easy to verify that $N(z)$ is constant in each connected component of $\mathbb{C} \setminus \Sigma_A$, and that $N(z) \le n$ [26]. The distinguished regions where f_A is polynomial of degree $n-3$ (if $n \ge 3$) or $f_A \equiv 0$ (if $n = 2$) are exactly those connected components of $\mathbb{C} \setminus \Sigma_A$ on which $N(z)$ *takes its maximal value* n. This remarkable property of the numerical density, which is one of our main results, will be established in Section 6.1. The geometric condition $N(z) = n$ is always satisfied in the complement of the numerical range, where f_A vanishes identically, but for many matrices of size $n \ge 3$ is it also met in some regions inside $W(A)$. For instance, in the three-dimensional example represented in Fig. 1, it is easy to verify that $N(z) = 3$ if z is outside $W(A)$ or inside the cuspidal triangle, and $N(z) = 1$ in the intermediate region where the numerical density is not constant.

At this point, it is necessary to make a connection with the theory of *lacunas* of symmetric hyperbolic systems of partial differential equations in $2+1$ variables. Given $A \in M_n(\mathbb{C})$, we consider the following system of linear PDE's in $\mathbb{R}_t \times \mathbb{R}^2_x$:

$$\partial_t u + A_1 \partial_{x_1} u + A_2 \partial_{x_2} u \;=\; 0\;, \tag{4}$$

where A_1, A_2 are as in (2) and $u = (u_1, \ldots, u_n)^\top : \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^n$. The *fundamental solution* of (4) is the unique (matrix-valued) distribution E supported in the half-space $\mathbb{R}_+ \times \mathbb{R}^2$ which satisfies

$$\partial_t E + A_1 \partial_{x_1} E + A_2 \partial_{x_2} E \;=\; I_n\, \delta_{t=0} \otimes \delta_{x=0}\;. \tag{5}$$

One can show that $E(t,x)$ is homogeneous of degree -2 in t and x, it is thus sufficient to consider the time-one trace $E_* = E(1,\cdot)$, which is a distribution on $\mathbb{R}^2$. Due to the finite speed of propagation, it is well-known that E_* is zero outside a compact set of $\mathbb{R}^2$, but it may also happen that E_* vanishes identically in some regions inside the domain of influence of the origin. Such regions are called *lacunas* of the hyperbolic system (4).

The properties of the fundamental solution of symmetric hyperbolic systems have been studied by many authors, see e.g. [22, 26, 4, 2, 3]. In the particular case of system (4), J. Bazer and D. Yen have shown that, if one identifies $\mathbb{C}$ with $\mathbb{R}^2$, the singular support of the distribution E_* is contained in the critical set Σ_A, and the (stable) lacunas of system (4) are exactly the regions described above where the numerical density f_A is polynomial of degree $n-3$. This remarkable coincidence is of course not fortuitous. In Section 6.3, we explain it by showing that the fundamental solution E can be expressed as a linear combination of derivatives of order $n-1$ of a homogeneous extension of the numerical density f_A. This connection allows us to recover some of the main results of [4], and therefore confirms that the numerical measure is a natural quantity attached to the matrix A. One might even argue that μ_A contains more information than E_*, since for instance $\mathrm{supp}(\mu_A) = W(A)$ while $\mathrm{supp}(E_*)$ is in general strictly smaller and not necessarily convex, see Section 7. Similarly, we believe that $\mathrm{sing\,supp}(\mu_A)$ always coincide with Σ_A, while $\mathrm{sing\,supp}(E_*)$ is usually smaller.

A final question that is worth investigating is the behavior of the numerical measure μ_A when the size of the matrix A goes to infnity. Here of course, specific assumptions have to be made in order to obtain convergence results. Suppose for instance that $\{A_n\}_{n\geq 1}$ is a sequence of complex matrices with $A_n \in \mathbf{M}_n(\mathbb{C})$, $\mathrm{Tr}(A_n) = 0$, and $\|A_n\| \leq C$ for all $n \geq 1$. If, for each $n \geq 1$, X_n is a random variable that is uniformly distributed on $\partial\mathbb{B}^n$, we show in Section 8 that the complex variable $\langle A_n X_n, X_n\rangle$ converges almost surely to zero as $n \to \infty$. This is reminiscent of the strong law of large numbers in probability theory. Under slightly stronger assumptions, we also establish the analog of the central limit theorem in this context. Our convergence results mean that μ_A is very close to δ_z when $\dim(A)$ is large, where z is the barycenter of $\sigma(A)$. This explains why plotting $\langle Ax, x\rangle$ for randomly chosen points $x \in \partial\mathbb{B}^n$ is a very unefficient algorithm for determining the numerical range $W(A)$ if A is a large matrix!

The rest of the paper is organized as follows. In Section 2, we establish some general properties of the numerical measure. Section 3 is devoted to the particular situations where the matrix A is Hermitian or normal. The nonnormal case is treated in Sections 4–6, which constitute the core of the paper. In Section 4, we derive a representation formula for the numerical density using the inversion of the Radon transformation. Section 5 collects a few results on the geometry of the critical set Σ_A, which are mainly borrowed from [19, 26, 17]. These informations are used in Section 6 to derive an explicit formula for the derivatives of order $n-2$ of the numerical density, which allows us to obtain generic regularity results and to express the fundamental solution of the hyperbolic system (4) in terms of derivatives of the numerical density. To illustrate our results, a few explicit examples are treated in Section 7. Finally, we investigate in Section 8 the concentration properties of the numerical density for large matrices, and we discuss in Section 9 a possible extension of our results to hyperbolic polynomials with an arbitrary number of variables.

Acknowledgements. This work has benefited of stimulating discussions with several of our colleagues, including Y. Colin de Verdière, F. Faure, and A. Joye.

2 General properties of the numerical measure

In this section, we establish a few general properties of the numerical measure of a complex matrix. In particular, we show that μ_A is absolutely continuous with respect to the Lebesgue measure on $W(A)$, and we prove a direct sum formula which will be useful later.

2.1 Support and regularity properties

We first show that the support of the numerical measure always coincides with the numerical range of the matrix.

Lemma 2.1 *For any $A \in \mathbf{M}_n(\mathbb{C})$, one has* $\operatorname{supp}(\mu_A) = W(A)$.

Proof. If $V = \mathbb{C} \setminus W(A)$, then $\Phi_A^{-1}(V) = \emptyset$, hence $\mu_A(V) = \bar\sigma(\Phi_A^{-1}(V)) = 0$. This shows that $\operatorname{supp}(\mu_A) \subset W(A)$. Conversely, if $V \subset \mathbb{C}$ is any open set such that $V \cap W(A) \neq \emptyset$, then $\Phi_A^{-1}(V)$ is a nonempty open subset of $\partial\mathbb{B}^n$, hence $\mu_A(V) = \bar\sigma(\Phi_A^{-1}(V)) > 0$. Thus $W(A) \subset \operatorname{supp}(\mu_A)$. ■

Our next goal is to locate the singular support of μ_A. We recall that $x \in \partial\mathbb{B}^n$ is a *regular point* of Φ_A if the differential map $\mathrm{d}_x\Phi_A : T_x^*\partial\mathbb{B}^n \to \mathbb{C}$ is onto. Otherwise, we say that x is a *critical point.* The following characterization will be useful:

Lemma 2.2 [16, 17] *Let $A \in \mathbf{M}_n(\mathbb{C})$.*
1) A point $x \in \partial\mathbb{B}^n$ is a critical point of the numerical map Φ_A if and only if x is an eigenvector of the Hermitian matrix $H(\theta)$ defined in (2) for some $\theta \in [0,\pi]$.
2) The differential of Φ_A vanishes at $x \in \partial\mathbb{B}^n$ if and only if x is an eigenvector of both A and A^*.

In other words, the range of the differential $\mathrm{d}_x\Phi_A$ has (real) dimension 1 if and only if x is an eigenvector of $H(\theta)$ for a *unique* $\theta \in [0,\pi)$, and is reduced to $\{0\}$ if and only if x is an eigenvector of $H(\theta)$ for *all* $\theta \in [0,\pi]$. The proof is neither new nor difficult, but we shall repeat it here in order to introduce some notation that will be needed later on.

Proof. Since $\Phi_A(e^{i\theta}x) = \Phi_A(x)$ for all $\theta \in [0,2\pi]$, we can consider the numerical map as acting on the quotient space $\partial\mathbb{B}^n/S^1 \simeq \mathbb{C}P^{n-1}$ [16]. Thus, to detect the critical points of Φ_A, we study the reduced map $\tilde\Phi_A : \partial\mathbb{B}^n/S^1 \to \mathbb{C}$ defined by

$$\tilde\Phi_A([x]) \,=\, \Phi_A(x) \,=\, \langle A_1x, x\rangle + i\langle A_2x, x\rangle \;, \qquad x \in \partial\mathbb{B}^n \;,$$

where $[x] = \{e^{i\theta}x \,|\, \theta \in S^1\}$ and A_1, A_2 are the Hermitian matrices introduced in (2).

If $x \in \partial\mathbb{B}^n$, the tangent space to $\partial\mathbb{B}^n/S^1$ at $[x]$ is just the $(2n-2)$-dimensional affine subspace $\{[x+y] \,|\, y \in \mathbb{C}^n\,,\ \langle x, y\rangle = 0\}$. Thus, using the definition above of $\tilde\Phi_A$, it is straightforward to verify that, for all $y \in \mathbb{C}^n$ with $\langle x, y\rangle = 0$, one has

$$\frac{1}{2}\,\mathrm{d}_{[x]}\tilde\Phi_A(y) \,=\, (A_1x|y) + i(A_2x|y) \,=\, (v_1(x)|y) + i(v_2(x)|y) \;, \tag{6}$$

where $(x|y) = \operatorname{Re}\langle x, y\rangle = (\operatorname{Re}x)^t(\operatorname{Re}y) + (\operatorname{Im}x)^t(\operatorname{Im}y)$ denotes the real scalar product in $\mathbb{C}^n \simeq \mathbb{R}^{2n}$, and

$$v_1(x) \,=\, A_1x - \langle A_1x, x\rangle x \;, \qquad v_2(x) \,=\, A_2x - \langle A_2x, x\rangle x \;.$$

Of course, replacing A_1x, A_2x with $v_1(x)$, $v_2(x)$ has no effect in (6) since $\langle x, y\rangle = 0$, but after this substitution we can let y run over the whole of $\mathbb{C}^n$ without increasing the range. So our task is reduced to computing the rank of the $\mathbb{R}$-linear map $y \mapsto (v_1(x)|y) + i(v_2(x)|y)$, which is just the rank of the 2×2 matrix

$$D(x) \,=\, \begin{pmatrix} (v_1(x)|v_1(x)) & (v_1(x)|v_2(x)) \\ (v_2(x)|v_1(x)) & (v_2(x)|v_2(x)) \end{pmatrix} \;. \tag{7}$$

By the Cauchy-Schwarz inequality, the positive matrix $D(x)$ is singular if and only if there exists $\theta \in [0, \pi]$ such that $v_1(x) \cos\theta + v_2(x) \sin\theta = 0$, which exactly means that x is an eigenvector of $H(\theta)$. Moreover, $D(x) = 0$ if and only if $v_1(x) = v_2(x) = 0$, which is equivalent to saying that x is an eigenvector of both A_1 and A_2, hence of both A and A^*. ■

Let $\Sigma_A \subset \mathbb{C}$ denote the set of all *critical values* of Φ_A, namely $\Sigma_A = \Phi_A(\Gamma(A))$ where $\Gamma(A) \subset \partial\mathbb{B}^n$ is the set of all critical points of Φ_A. Our next result is:

Lemma 2.3 *If* $A \in \mathbf{M}_n(\mathbb{C})$*, then* $\operatorname{sing\,supp}(\mu_A) \subset \Sigma_A$.

Proof. If the numerical range $W(A)$ is reduced to a line segment or to a single point, then $\Sigma_A = W(A)$, hence $\operatorname{sing\,supp}(\mu_A) \subset \operatorname{supp}(A) = \Sigma_A$ by Lemma 2.1. Thus, we assume from now on that $W(A)$ has nonempty interior. By Sard's lemma, the critical set Σ_A is then a compact subset of $W(A)$ with zero Lebesgue measure. We have to show that there exists a smooth density function $f_A \geq 0$ such that $\mathrm{d}\mu_A(z) = f_A(z)\,\mathrm{d}z$ on $\mathbb{C} \setminus \Sigma_A$. Clearly, we must have $f_A = 0$ on $\mathbb{C} \setminus W(A)$.

If $z \in W(A) \setminus \Sigma_A$, then $\mathcal{N}_z := \Phi_A^{-1}(z)$ is a compact submanifold of $\partial\mathbb{B}^n$ of codimension 2, which depends smoothly on z. Using classical arguments, involving a partition of unity and the Implicit Function Theorem, it is not difficult to verify that, for any continuous function ϕ with $\operatorname{supp}(\phi) \subset W(A) \setminus \Sigma_A$, one has

$$\int_{\partial\mathbb{B}^n} \phi(\langle Ax, x\rangle)\,\mathrm{d}\bar\sigma(x) \;=\; \frac{1}{\omega_n}\int_{\mathbb{C}} \phi(z)\left\{\int_{\mathcal{N}_z} \frac{\mathrm{d}\nu(x)}{2\Delta(x)^{1/2}}\right\}\mathrm{d}z \;,$$

where $\omega_n = 2\pi^n/((n-1)!)$ is the total measure of $\partial\mathbb{B}^n$, ν is the $(2n-3)$-dimensional Euclidean measure on the submanifold $\mathcal{N}_z$, and $\Delta(x) = \det D(x)$ where $D(x)$ is the 2×2 matrix defined in (7). Remark that $2\Delta(x)^{1/2} = \bar\lambda_1(x)\bar\lambda_2(x)$, where $\bar\lambda_1(x), \bar\lambda_2(x)$ are the singular values of the differential map $\mathrm{d}_x\Phi_A$. In view of (1), we conclude that $\mathrm{d}\mu_A(z) = f_A(z)\,\mathrm{d}z$ on $\mathbb{C} \setminus \Sigma_A$, where

$$f_A(z) \;=\; \frac{1}{\omega_n}\int_{\mathcal{N}_z} \frac{\mathrm{d}\nu(x)}{2\Delta(x)^{1/2}} \;, \qquad z \in \mathbb{C} \setminus \Sigma_A \;. \tag{8}$$

It is easily verified that the density f_A is smooth and strictly positive on $W(A) \setminus \Sigma_A$. ■

The results obtained so far are summarized in the following proposition, which also asserts that the numerical measure is absolutely continuous with respect to Lebesgue's measure on $W(A)$.

Proposition 2.4 *Let* $A \in \mathbf{M}_n(\mathbb{C})$.
1) If the numerical range $W(A)$ *has nonempty interior, the numerical measure* μ_A *is absolutely continuous with respect to the (two-dimensional) Lebesgue measure on* $\mathbb{C} \simeq \mathbb{R}^2$*. The numerical density* $f_A = \mathrm{d}\mu_A/\mathrm{d}z$ *is smooth outside the critical set* Σ_A.

2) If A *is a nonscalar Hermitian matrix, then* $W(A) \subset \mathbb{R}$ *and the numerical measure is absolutely continuous with respect to the (one-dimensional) Lebesgue measure on* $\mathbb{R}$*. The numerical density* $f_A = \mathrm{d}\mu_A/\mathrm{d}x$ *is smooth outside the spectrum* $\sigma(A)$.

Remark 2.5 *As is explained in the introduction, Proposition 2.4 covers all interesting cases. Indeed, if the numerical range has empty interior, then either* $W(A)$ *is reduced to a single point, in which case* A *is a scalar matrix and* μ_A *is just a Dirac mass, or* $W(A)$ *is a line segment of nonzero length, in which case* A *can be reduced to a nonscalar Hermitian matrix by a simple affine transformation.*

Proof. Using the same notations as in Lemmas 2.2 and 2.3, we observe that $\Gamma(A) = \{x \in \partial\mathbb{B}^n \,|\, \Delta(x) = 0\}$, where $\Delta(x) = \det D(x)$ is a polynomial in the $2n$ variables $\mathrm{Re}\, x_i, \mathrm{Im}\, x_i$ $(i = 1, \ldots, n)$. Thus one of the following two situations must occur:

1) $\Gamma(A)$ is an algebraic submanifold of $\partial\mathbb{B}^n$ of codimension at least 1. By Sard's lemma, this is the case if and only if $W(A)$ has nonempty interior. In that situation, since we already know that μ_A has a smooth density outside the critical set Σ_A, we only need to show that $\mu_A(\Sigma_A) = 0$. Given $\epsilon > 0$, let $\Gamma_\epsilon(A) = \{x \in \partial\mathbb{B}^n \,|\, \mathrm{dist}(x, \Gamma(A)) \le \epsilon\}$, where "dist" denotes here the geodesic distance on the unit sphere. We decompose

$$\Phi_A^{-1}(\Sigma_A) \,=\, \Big(\Phi_A^{-1}(\Sigma_A) \cap \Gamma_\epsilon(A)\Big) \cup \Big(\Phi_A^{-1}(\Sigma_A) \cap \Gamma_\epsilon(A)^c\Big) \,=\, E_1(\epsilon) \cup E_2(\epsilon) \;,$$

where $\Gamma_\epsilon(A)^c = \partial\mathbb{B}^n \setminus \Gamma_\epsilon(A)$. Since $\bar\sigma(\Gamma(A)) = 0$, we have $\bar\sigma(E_1(\epsilon)) \le \bar\sigma(\Gamma_\epsilon(A)) \to 0$ as $\epsilon \to 0$. Moreover, the proof of Lemma 2.3 shows $E_2(\epsilon)$ is a codimension two submanifold of $\partial\mathbb{B}^n$, so that $\bar\sigma(E_2(\epsilon)) = 0$ for any $\epsilon > 0$. Using the definition of the numerical measure, we conclude that $\mu_A(\Sigma_A) = \bar\sigma(\Phi_A^{-1}(\Sigma_A)) = 0$.

2) $\Gamma(A) = \partial\mathbb{B}^n$. This is the case if and only if $W(A)$ has empty interior, and without loss of generality we can then assume that the matrix A is Hermitian. Since $W(A) \subset \mathbb{R}$, it is more natural here to consider Φ_A as a map from $\partial\mathbb{B}^n$ into $\mathbb{R}$. If we do that, then repeating the proofs of Lemmas 2.2 and 2.3 we easily find that the critical points of Φ_A are exactly the eigenvectors of A. Moreover, the numerical measure has a smooth density on $\mathbb{R} \setminus \sigma(A)$, and is absolutely continuous with respect to the Lebesgue measure $\mathrm{d}x$ if A is not a scalar matrix. We skip the details here, because the Hermitian case will be treated in full details in Section 3 below. ∎

To conclude this section, we show that the numerical density is strictly positive in the interior of $W(A)$. A little care is needed in the formulation of that result, because as we shall see in Section 4.2 the numerical density need not be a continuous function.

Proposition 2.6 *If $z_0 \in \mathbb{C}$ is an interior point of $W(A)$, then*

$$\liminf_{\epsilon \to 0} \frac{1}{\epsilon^2}\, \mu_A(\{z \in \mathbb{C} \,|\, |z - z_0| \le \epsilon\}) \,>\, 0 \;. \tag{9}$$

Proof. If z_0 is an interior point of $W(A)$, it is shown in [17, Proposition 2.11] that the preimage $\Phi_A^{-1}(z_0)$ contains at least one regular point x_0. Let V be an open geodesic ball centerd at $x_0 \in \partial\mathbb{B}^n$ whose closure does not intersect $\Gamma(A)$. Proceeding as in the proof of Lemma 2.3, we find

$$\mu_A(\{z \in \mathbb{C} \,|\, |z - z_0| \le \epsilon\}) \,\ge\, \frac{1}{\omega_n} \int_{|z - z_0| \le \epsilon} \left\{ \int_{\mathcal{N}_z \cap V} \frac{\mathrm{d}\nu(x)}{2\Delta(x)^{1/2}} \right\} \mathrm{d}z \;.$$

If $\epsilon > 0$ is sufficiently small, the integral inside the curly brackets is a smooth and positive function of z, and (9) follows. ∎

2.2 The direct sum formula

Let $p, q \in \mathbb{N}^*$ and $n = p + q$. Given $A \in \mathbf{M}_p(\mathbb{C})$ and $B \in \mathbf{M}_q(\mathbb{C})$, the *direct orthogonal sum* of A and B is the matrix $A \oplus B \in \mathbf{M}_n(\mathbb{C})$ defined by

$$A \oplus B \,=\, \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \;.$$

In this situation, we have a formula for the numerical measure $\mu_{A \oplus B}$ in terms of μ_A and μ_B.

Proposition 2.7 *For any $\phi \in C^0(\mathbb{C})$, we have*

$$\int_{\mathbb{C}} \phi(z)\,\mathrm{d}\mu_{A\oplus B}(z) \;=\; \frac{1}{\mathcal{B}(p,q)} \int_{\mathbb{C}}\int_{\mathbb{C}}\int_0^1 \phi(tz' + (1-t)z'')\, t^{p-1}(1-t)^{q-1}\,\mathrm{d}t\,\mathrm{d}\mu_A(z')\,\mathrm{d}\mu_B(z'')\;, \tag{10}$$

where $\mathcal{B}(p,q)$ is Euler's beta function

$$\mathcal{B}(p,q) \;=\; \int_0^1 t^{p-1}(1-t)^{q-1}\,\mathrm{d}t \;=\; \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}\;.$$

Proof. Any unit vector $x \in \partial\mathbb{B}^n$ can be written as

$$x \;=\; \begin{pmatrix} \sqrt{t}\,u \\ \sqrt{1-t}\,v \end{pmatrix}\;,$$

where $u \in \partial\mathbb{B}^p$, $v \in \partial\mathbb{B}^q$, and $t \in [0,1]$. Up to negligible sets, the map $(u,v,t) \mapsto x$ defines a diffeomorphism from $\partial\mathbb{B}^p \times \partial\mathbb{B}^q \times [0,1]$ onto $\partial\mathbb{B}^n$. With this parametrization it is not difficult to verify that the Euclidean measure on $\partial\mathbb{B}^n$ has the following expression

$$\mathrm{d}\sigma_n(x) \;=\; \frac{1}{2}\, t^{p-1}(1-t)^{q-1}\,\mathrm{d}t\,\mathrm{d}\sigma_p(u)\,\mathrm{d}\sigma_q(v)\;.$$

Equivalently, since $\sigma_n(\partial\mathbb{B}^n) = \omega_n = 2\pi^n/\Gamma(n)$, the normalized Euclidean measure satisfies

$$\mathrm{d}\bar{\sigma}_n(x) \;=\; \frac{1}{\mathcal{B}(p,q)}\, t^{p-1}(1-t)^{q-1}\,\mathrm{d}t\,\mathrm{d}\bar{\sigma}_p(u)\,\mathrm{d}\bar{\sigma}_q(v)\;.$$

Thus, using definition (1) and the fact that $\langle (A\oplus B)x, x\rangle = t\langle Au,u\rangle + (1-t)\langle Bv,v\rangle$, we easily obtain

$$\begin{aligned}
\int_{\mathbb{C}} \phi(z)\,\mathrm{d}\mu_{A\oplus B}(z) \;&=\; \frac{1}{\mathcal{B}(p,q)} \int_{\partial\mathbb{B}^p}\int_{\partial\mathbb{B}^q}\int_0^1 \phi\Big(t\langle Au,u\rangle + (1-t)\langle Bv,v\rangle\Big)\, t^{p-1}(1-t)^{q-1}\,\mathrm{d}t\,\mathrm{d}\bar{\sigma}_q(v)\,\mathrm{d}\bar{\sigma}_p(u)\\
&=\; \frac{1}{\mathcal{B}(p,q)} \int_{\mathbb{C}}\int_{\mathbb{C}}\int_0^1 \phi(tz' + (1-t)z'')\, t^{p-1}(1-t)^{q-1}\,\mathrm{d}t\,\mathrm{d}\mu_A(z')\,\mathrm{d}\mu_B(z'')\;,
\end{aligned}$$

which is the desired result. ∎

As an application, if we choose $\phi(z) = e^{-i\xi\cdot z}$ in Proposition 2.7, we obtain the following relation between the Fourier transforms of the measures μ_A, μ_B and $\mu_{A\oplus B}$:

$$\hat{\mu}_{A\oplus B}(\xi) \;=\; \frac{1}{\mathcal{B}(p,q)} \int_0^1 \hat{\mu}_A(t\xi)\,\hat{\mu}_B((1-t)\xi)\, t^{p-1}(1-t)^{q-1}\,\mathrm{d}t\;, \quad \xi \in \mathbb{R}^2\;. \tag{11}$$

The formula given in Proposition 2.7 can be generalized in a straightforward way to a direct sum with an arbitrary number of terms. Assume that $A = A_1 \oplus \cdots \oplus A_k$ where $A_j \in \mathbf{M}_{p_j}(\mathbb{C})$, so that $A \in \mathbf{M}_n(\mathbb{C})$ with $n = p_1 + \cdots + p_k$. Setting $p = (p_1,\ldots,p_k)$, we denote

$$\mathcal{B}(p) \;=\; \int_{D_{k-1}} t_1^{p_1-1}\cdots t_k^{p_k-1}\,\mathrm{d}t_1\ldots\mathrm{d}t_{k-1} \;=\; \frac{\Gamma(p_1)\cdots\Gamma(p_k)}{\Gamma(p_1+\cdots+p_k)}\;,$$

where $t_k = 1 - (t_1 + \cdots + t_{k-1})$ and D_{k-1} denotes the $(k-1)$-dimensional simplex

$$D_{k-1} \;=\; \left\{ (t_1, \ldots, t_{k-1}) \in \mathbb{R}_+^{k-1} \,\middle|\, t_1 + \cdots + t_{k-1} \le 1 \right\} . \tag{12}$$

Using (10) and proceeding by induction over k, we easily obtain the general formula

$$\int_{\mathbb{C}} \phi(z)\,\mathrm{d}\mu_A(z) \;=\; \frac{1}{\mathcal{B}(p)} \int_{\mathbb{C}^k} \int_{D_{k-1}} \phi(t_1 z_1 + \cdots + t_k z_k)\, t_1^{p_1 - 1} \cdots t_k^{p_k - 1} \\ \mathrm{d}t_1 \ldots \mathrm{d}t_{k-1}\, \mathrm{d}\mu_{A_1}(z_1) \ldots \mathrm{d}\mu_{A_k}(z_k) \;, \tag{13}$$

where it is understood again that $t_k = 1 - (t_1 + \cdots + t_{k-1})$.

3 The numerical density of a normal matrix

If $A \in \mathbf{M}_n(\mathbb{C})$ is a normal matrix, the numerical measure μ_A is entirely determined by the spectrum $\sigma(A) = \{\lambda_1, \ldots, \lambda_n\}$. Indeed, we know that A is unitarily equivalent to the diagonal matrix $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, and that a unitary conjugation does not affect the numerical measure. Using this observation and the direct sum formula of Section 2.2, we shall prove that the numerical density of A is a piecewise polynomial function, which can be characterized as a multivariate B-spline whose knots are the eigenvalues of A. We begin with the important particular case where all eigenvalues of A are colinear.

3.1 The Hermitian case

If $A \in \mathbf{M}_n(\mathbb{C})$ is a Hermitian matrix, then $W(A) \subset \mathbb{R}$ and the numerical measure μ_A is therefore supported on the real axis. Assuming that A is not a multiple of the identity matrix, we show in this section that μ_A is absolutely continuous with respect to Lebesgue's measure on $\mathbb{R}$, and we give a simple characterization of the numerical density $f_A = \mathrm{d}\mu_A/\mathrm{d}x$. The result is:

Proposition 3.1 *If $A \in \mathbf{M}_n(\mathbb{C})$ is a nonscalar Hermitian matrix, the numerical density $f_A : \mathbb{R} \to \mathbb{R}_+$ is the normalized B-spline of degree $n-2$ whose knots are the eigenvalues of A.*

To make the statement clear, we briefly recall the definition and some elementary properties of the classical B-splines [7]. If $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ are pairwise distinct, the $(n-1)^{th}$ *divided difference* of a continuous function g at the points $\lambda_1, \ldots, \lambda_n$ is the quantity

$$\delta^{n-1}[\lambda_1, \ldots, \lambda_n] g \;=\; \sum_{j=1}^{n} \frac{g(\lambda_j)}{\prod_{k \ne j} (\lambda_j - \lambda_k)} \;. \tag{14}$$

This is the leading coefficient of the unique polynomial of degree at most $n-1$ which agrees with g at the points $\lambda_1, \ldots, \lambda_n$. It is easy to verify that the right-hand side of (14) is a completely symmetric function of the variables λ_j. If $g \in C^{n-1}(\mathbb{R})$, the divided difference can be extended by continuity to arbitrary (not necessarily distinct) values of $\lambda_1, \ldots, \lambda_n$, and we have the integral formula:

$$\delta^{n-1}[\lambda_1, \ldots, \lambda_n] g \;=\; \int_{D_{n-1}} g^{(n-1)}(t_1 \lambda_1 + \cdots + t_n \lambda_n)\, \mathrm{d}t_1 \ldots \mathrm{d}t_{n-1} \;, \tag{15}$$

where D_{n-1} is the $(n-1)$-dimensional simplex defined in (12) and $t_n = 1 - (t_1 + \cdots + t_{n-1})$. In what follows, we shall always assume that the set $S = \{\lambda_1, \ldots, \lambda_n\}$ is not reduced to a single point, so that the $(n-1)^{th}$ divided difference is well-defined as soon as g is of class C^{n-2} in a neighborhood of S.

With these notations, the *normalized B-spline of degree $n-2$ with knots* $\lambda_1, \ldots, \lambda_n$ is the function $B : \mathbb{R} \to \mathbb{R}$ defined by

$$B(x) \equiv B[\lambda_1, \ldots, \lambda_n](x) = (n-1)\, \delta^{n-1}[\lambda_1, \ldots, \lambda_n](\cdot - x)_+^{n-2} \,, \qquad x \in \mathbb{R} \,, \tag{16}$$

where $(\cdot - x)_+^{n-2}$ denotes the map $y \mapsto \max(0, y-x)^{n-2}$. If $\lambda_1 \le \cdots \le \lambda_n$, it is not difficult to show that $B(x)$ vanishes identically outside $[\lambda_1, \lambda_n]$, and coincides with a polynomial of degree at most $n-2$ on each nonempty interval $(\lambda_j, \lambda_{j+1})$. Moreover, if m_j denotes the multiplicity of λ_j in S, one can verify that $B(x)$ is continuous at $x = \lambda_j$ together with its derivatives up to order $d_j = n - 2 - m_j$, provided $d_j \ge 0$. If $d_j = -1$, then $B(x)$ is discontinuous at λ_j. Finally, we shall see below that $B(x)$ is positive on (λ_1, λ_n) and that $\int_{\mathbb{R}} B(x)\, \mathrm{d}x = 1$.

Proof of Proposition 3.1. Let $A \in \mathbf{M}_n(\mathbb{C})$ be a normal matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$. To compute the numerical density, we can assume without loss of generality that A is diagonal, namely $A = A_1 \oplus \cdots \oplus A_n$ with $A_j = \lambda_j \in \mathbf{M}_1(\mathbb{C})$. Thus we can use the direct sum formula (13) with $k = n$ and $p = (1, \ldots, 1)$. Since $\mathcal{B}(p)^{-1} = (n-1)!$ and $\mu_{A_j} = \delta_{\lambda_j}$ for $j = 1, \ldots, n$, we obtain the relation

$$\int_{\mathbb{C}} \phi(z)\, \mathrm{d}\mu_A(z) = (n-1)! \int_{D_{n-1}} \phi(t_1\lambda_1 + \cdots + t_n\lambda_n)\, \mathrm{d}t_1 \ldots \mathrm{d}t_{n-1} \,, \tag{17}$$

for any continuous function $\phi : \mathbb{C} \to \mathbb{C}$.

Assume now that A is Hermitian, so that $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$, and that $\sigma(A)$ is not reduced to a single point. If $B = B[\lambda_1, \ldots \lambda_n]$ is the normalized B-spline defined by (16), we claim that

$$\int_{\mathbb{R}} \phi(x) B[\lambda_1, \ldots, \lambda_n](x)\, \mathrm{d}x = (n-1)! \int_{D_{n-1}} \phi(t_1\lambda_1 + \cdots + t_n\lambda_n)\, \mathrm{d}t_1 \ldots \mathrm{d}t_{n-1} \,. \tag{18}$$

Indeed, it is clearly sufficient to prove (18) for compactly supported functions $\phi \in C^0(\mathbb{R})$. Moreover, since both members of (18) depend continuously on $\lambda_1, \ldots, \lambda_n$, we can also assume that the eigenvalues of A are all distinct. In that case, it follows immediately from (14), (16) that

$$\int_{\mathbb{R}} \phi(x) B[\lambda_1, \ldots, \lambda_n](x)\, \mathrm{d}x = (n-1)\, \delta^{n-1}[\lambda_1, \ldots, \lambda_n]\Phi \,,$$

where $\Phi(y) = \int_{\mathbb{R}} \phi(x)(y-x)_+^{n-2}\, \mathrm{d}x$. Since $\Phi^{(n-1)} = (n-2)!\, \phi$, we can use the integral formula (15) to evaluate the divided difference in the right-hand side, and we obtain (18).

Now, comparing (17) and (18), we conclude that μ_A is absolutely continuous with respect to Lebesgue's measure on $\mathbb{R}$, and that the numerical density $f_A = \mathrm{d}\mu_A/\mathrm{d}x$ is precisely the normalized B-spline $B[\lambda_1, \ldots, \lambda_n]$. Incidentally, the argument above shows that B is positive on its support (see Proposition 2.6) and that $\int_{\mathbb{R}} B(x)\, \mathrm{d}x = 1$. ∎

3.2 The quasi-Hermitian case

We say that a matrix $A \in \mathbf{M}_n(\mathbb{C})$ is *quasi-Hermitian* if the numerical range $W(A) \subset \mathbb{C}$ has empty interior, i.e. $W(A)$ is a single point or a line segment. In such a case, there exist $z \in \mathbb{C}$, $\theta \in S^1$, and a Hermitian matrix H such that $A = zI_n + e^{i\theta}H$ (in particular, A is normal). Indeed, if we choose z, θ such that $W(A) \subset z + e^{i\theta}\mathbb{R}$, the matrix $H = e^{-i\theta}(A - zI_n)$ satisfies $W(H) \subset \mathbb{R}$ and is therefore Hermitian. Since $\langle Ax, x\rangle = z + e^{i\theta}\langle Hx, x\rangle$ for all $x \in \partial\mathbb{B}^n$, it is clear from (1) that the numerical measure μ_A is just the image of μ_H under the affine isometry $w \mapsto z + e^{i\theta}w$. Combining this remark with Proposition 3.1, we thus obtain a precise characterization of the numerical measure of any quasi-Hermitian matrix.

3.3 The normal case

Finally, we consider the case of a normal matrix $A \in \mathbf{M}_n(\mathbb{C})$ whose numerical range $W(A)$ has nonempty interior. This, of course, is possible only if $n \geq 3$. By Proposition 2.4, the numerical measure μ_A is absolutely continuous with respect to Lebesgue's measure, and we have the following characterization of the numerical density $f_A = \mathrm{d}\mu_A/\mathrm{d}z$:

Proposition 3.2 *If $A \in \mathbf{M}_n(\mathbb{C})$ is a normal matrix whose numerical range $W(A) \subset \mathbb{C}$ has nonempty interior, the numerical density $f_A : \mathbb{C} \to \mathbb{R}_+$ is the bivariate B-spline of degree $n-3$ whose knots are the eigenvalues of A.*

The reader is referred here to the work of W. Dahmen [6], where multivariate B-splines are defined and studied in detail. To make the connection with the numerical density of a normal matrix, we use the relation (17), which corresponds to formula (2.2) in [6]. In the rest of this section, we assume that $A \in \mathbf{M}_n(\mathbb{C})$ is a normal matrix whose eigenvalues $\lambda_1, \dots, \lambda_n$ are not colinear, and we often identify the complex plane $\mathbb{C}$ with $\mathbb{R}^2$.

Proof of Proposition 3.2. Assume first that $n = 3$. Then $W(A) \subset \mathbb{C} \simeq \mathbb{R}^2$ is the 2-simplex with vertices $\lambda_1, \lambda_2, \lambda_3$, and using the change of variables $w = t_1\lambda_1 + t_2\lambda_2 + (1 - t_1 - t_2)\lambda_3$ in (17) we easily obtain

$$\int_{\mathbb{C}} \phi(z)\,\mathrm{d}\mu_A(z) \;=\; \frac{1}{|W(A)|}\int_{W(A)} \phi(w)\,\mathrm{d}w\ .$$

This shows that the numerical measure μ_A is uniformly distributed on $W(A)$. The numerical density is thus a multiple of the characteristic function of $W(A)$, which (by definition) is the bivariate B-spline of degree zero with knots $\lambda_1, \lambda_2, \lambda_3$.

We now assume that $n \geq 4$. Then we can choose n vectors $v_1, \dots, v_n \in \mathbb{R}^{n-1}$ such that
1) The $(n-1)$-simplex $S \subset \mathbb{R}^{n-1}$ with vertices $v_1, \dots, v_n$ has unit volume;
2) $Pv_i = \lambda_i$ for $i = 1, \dots, n$, where $P : \mathbb{R}^{n-1} \to \mathbb{R}^2$ is defined by $P(x_1, \dots, x_{n-1}) = (x_1, x_2)$.
This (elementary) claim is proved in [6, Section 2]. The simplex $S \subset \mathbb{R}^{n-1}$ is not uniquely defined, but any choice satisfies $PS = W(A) = \mathrm{conv}(\sigma(A))$. Returning to (17), we have

$$\int_{\mathbb{C}} \phi(z)\,\mathrm{d}\mu_A(z) \;=\; (n-1)!\int_{D_{n-1}} \phi(P(t_1v_1 + \cdots + t_nv_n))\,\mathrm{d}t_1 \dots \mathrm{d}t_{n-1} \;=\; \int_S \phi(Pw)\,\mathrm{d}w\ ,$$

where the second equality is obtained by applying the change of variables $(t_1, \dots, t_{n-1}) \mapsto w = t_1v_1 + \cdots + t_nv_n$, with $t_n = 1 - (t_1 + \cdots + t_{n-1})$. This shows that the numerical measure is the image under the projection P of the Lebesgue measure on the simplex $S \subset \mathbb{R}^{n-1}$. Given $z \in \mathbb{C} \simeq \mathbb{R}^2$, the numerical density $f_A(z)$ is thus the $(n-3)$-dimensional measure of the simplex $P^{-1}z \cap S$. This is precisely the definition given in [6] of a bivariate B-spline of degree $n-3$ with knots $\lambda_1, \dots, \lambda_n$. ∎

We conclude by listing a few properties of the numerical density f_A, which follow from [6, Theorem 4.1]. Let $\Sigma_A \subset \mathbb{C}$ be the union of all line segments joining pairs of eigenvalues of A. Using Lemma 2.2, it is straightforward to verify that Σ_A is exactly the set of critical values of the numerical map Φ_A. Then f_A is a polynomial of total degree $n-3$ in each connected component of $\mathbb{C} \setminus \Sigma_A$. Moreover, if $\ell \subset \mathbb{C}$ is a straight line passing through a pair of eigenvalues of A, the numerical density is continuous across $\ell \cap W(A)$ together with its derivatives up to order $n-2-m \geq 0$, where $m \geq 2$ is the number of eigenvalues of A (counted with multiplicities) which belong to ℓ. If $m = n-1$, then f_A is discontinuous on $\ell \cap W(A)$. In particular, in the generic case where no straight line contains more than two eigenvalues of A, the numerical density f_A is of class C^{n-4} if $n \geq 4$.

Remark. In the Hermitian and normal cases, the fact that the numerical density f_A is a projection of the characteristic function of a convex set, which is a log-concave function, implies that the density is itself a log-concave function. We thus have the following inequality

$$f_A(\lambda z + (1-\lambda)z') \;\geq\; f_A(z)^\lambda \, f_A(z')^{1-\lambda}\ , \tag{19}$$

for all $z, z' \in \mathbb{C}$ and all $\lambda \in (0,1)$. This follows from the Prékopa–Leindler inequality, see e.g. [10, Section 9]. We warn the reader that this property does not extend to nonnormal matrices, as we can see already from the two-dimensional case considered in Section 4.1.

4 The Radon transform of the numerical measure

Our purpose in this section is to derive a representation formula for the numerical density of a nonnormal matrix $A \in \mathbf{M}_n(\mathbb{C})$. Our approach is based on a natural expression of the *Radon transform* of the numerical measure μ_A in terms of the Hermitian matrices $H(\theta)$ defined in (2). By definition, the Radon transform of μ_A is the family $\mathcal{R}\mu_A = \{\mathcal{R}\mu_A(\theta)\,|\,\theta \in S^1\}$, where $\mathcal{R}\mu_A(\theta)$ denotes the Borel measure on $\mathbb{R}$ defined by

$$(\mathcal{R}\mu_A(\theta))(I) \;=\; \mu_A(\{z \in \mathbb{C}\,|\,\mathrm{Re}\,(e^{-i\theta}z) \in I\})\ ,$$

for any open set $I \subset \mathbb{R}$. In other words, $\mathcal{R}\mu_A(\theta)$ is the image of the measure μ_A under the orthogonal projection in $\mathbb{C} \simeq \mathbb{R}^2$ onto the line $e^{i\theta}\mathbb{R}$. The fundamental observation is:

Proposition 4.1 *For any $\theta \in S^1 = \mathbb{R}/(2\pi\mathbb{Z})$, one has $\mathcal{R}\mu_A(\theta) = \mu_{H(\theta)}$.*

Proof. The definition (2) implies that $\mathrm{Re}\,(e^{-i\theta}\langle Ax,x\rangle) = \langle H(\theta)x,x\rangle$ for any $x \in \partial\mathbb{B}^n$. Thus, for any open set $I \subset \mathbb{R}$, we have

$$\begin{aligned}(\mathcal{R}\mu_A(\theta))(I) \;&=\; \bar{\sigma}(\{x \in \partial\mathbb{B}^n\,|\,\mathrm{Re}\,(e^{-i\theta}\langle Ax,x\rangle) \in I\}) \\ &=\; \bar{\sigma}(\{x \in \partial\mathbb{B}^n\,|\,\langle H(\theta)x,x\rangle \in I\}) \;=\; \mu_{H(\theta)}(I)\ ,\end{aligned}$$

which proves the claim. ∎

Proposition 4.1 shows that the numerical measure of an arbitrary matrix $A \in \mathbf{M}_n(\mathbb{C})$ is entirely determined by the one-dimensional measures associated with the Hermitian matrices $\{H(\theta)\,|\,\theta \in S^1\}$. If we assume that $W(A)$ has nonempty interior, which is always the case if A is nonnormal, the matrix $H(\theta)$ is nonscalar for every $\theta \in S^1$ and Proposition 3.1 show that its numerical density is the B-spline

$B[\lambda_1(\theta),\dots,\lambda_n(\theta)]$, where $\lambda_1(\theta),\dots,\lambda_n(\theta)$ are the eigenvalues of $H(\theta)$. In that case, the result of Proposition 4.1 can be stated in the following equivalent form

$$\int_{\mathbb{R}} f_A(e^{i\theta}(x+iy))\,\mathrm{d}y \;=\; B[\lambda_1(\theta),\dots,\lambda_n(\theta)](x)\;, \tag{20}$$

where equality holds for all $\theta \in S^1$ and almost all $x \in \mathbb{R}$ since the numerical density f_A belongs to $L^1(\mathbb{C})$.

Our goal is to invert the Radon transform (20) to obtain a representation formula for the numerical density f_A. The general results established in [12] show that

$$(Jf_A)(x+iy) \;=\; \frac{1}{4\pi}\int_{S^1} B[\lambda_1(\theta),\dots,\lambda_n(\theta)](x\cos\theta + y\sin\theta)\,\mathrm{d}\theta\;, \tag{21}$$

where $J = (-\Delta)^{-1/2}$ is the Riesz potential defined by $(Jf)(z) = \frac{1}{2\pi}\int_{\mathbb{C}} |z-z'|^{-1} f(z')\,\mathrm{d}z'$. The idea is thus to apply the nonlocal operator $(-\Delta)^{1/2}$ to both sides of (21), but since we are not dealing with smooth functions we have to differentiate in the sense of distributions. As a preliminary remark, if $f(x,y) = g(x\cos\theta + y\sin\theta)$ for some test function $g : \mathbb{R} \to \mathbb{R}$ and some fixed $\theta \in S^1$, a direct calculation shows that $(\Delta f)(x,y) = g''(x\cos\theta + y\sin\theta)$, and a standard interpolation argument allows us to conclude that $(-\Delta)^{1/2} f(x,y) = \mathcal{H}g'(x\cos\theta + y\sin\theta)$, where $\mathcal{H}g'$ denotes the *Hilbert transform* of the derivative g'. Using this observation, we easily obtain the representation formula

$$f_A(x+iy) \;=\; \frac{1}{4\pi}\int_{S^1} \mathcal{H}B'[\lambda_1(\theta),\dots,\lambda_n(\theta)](x\cos\theta + y\sin\theta)\,\mathrm{d}\theta\;, \tag{22}$$

where both sides define integrable functions of $z = x+iy \in \mathbb{C}$, and equality holds almost everywhere. So we have shown:

Proposition 4.2 *If the numerical range of a matrix $A \in \mathbf{M}_n(\mathbb{C})$ has nonempty interior, the numerical density of A can be represented as in* (22) *for almost all $(x,y) \in \mathbb{R}^2$.*

In the rest of this section, we shall apply Proposition 4.2 to compute the numerical density of a two-dimensional nonnormal matrix. We shall also consider the interesting particular situation where the numerical density is radially symmetric, in which case the representation formula takes a simpler form. Proposition 4.2 will be used again in Section 6 to derive some important properties of the numerical measure in the general case.

4.1 The two-dimensional case

Let $A \in \mathbf{M}_2(\mathbb{C})$, and assume that the numerical range $W(A)$ has nonempty interior. As is well-known [15], $W(A)$ is then a filled ellipse, and without loss of generality we can assume that this ellipse is centered at the origin and that its major axis is aligned with the real axis of the complex plane. In that case, up to a unitary conjugation, the matrix A has the following form

$$A \;=\; \begin{pmatrix} -c & 2b \\ 0 & c \end{pmatrix}\;, \qquad W(A) \;=\; \Big\{ x+iy \in \mathbb{C} \,\Big|\, \frac{x^2}{a^2} + \frac{y^2}{b^2} \le 1 \Big\}\;, \tag{23}$$

where $b > 0$, $c \ge 0$, and $a = (b^2+c^2)^{1/2}$. For $\theta \in S^1 = \mathbb{R}/(2\pi\mathbb{Z})$, let

$$H(\theta) \;=\; \frac{1}{2}(e^{-i\theta}A + e^{i\theta}A^*) \;=\; \begin{pmatrix} -c\cos\theta & b\,e^{-i\theta} \\ b\,e^{i\theta} & c\cos\theta \end{pmatrix}\;.$$

The eigenvalues of $H(\theta)$ are $\pm\lambda(\theta)$, where $\lambda(\theta) = (b^2 + c^2\cos^2\theta)^{1/2}$. It follows that

$$B[-\lambda(\theta), \lambda(\theta)](s) \;=\; \frac{1}{2\lambda(\theta)}\mathbf{1}_{[-\lambda(\theta),\lambda(\theta)]}(s) \;.$$

According to (22), we have to differentiate this expression with respect to s, which yields a linear combination of Dirac masses, and to apply the Hilbert transformation, which is the convolution with the distribution p.v.$\frac{1}{\pi s}$. We obtain

$$\mathcal{H}B'[-\lambda(\theta), \lambda(\theta)](s) \;=\; \frac{1}{2\pi\lambda(\theta)}\Big(\text{p.v.}\frac{1}{s+\lambda(\theta)} - \text{p.v.}\frac{1}{s-\lambda(\theta)}\Big) \;=\; \frac{1}{\pi}\,\text{p.v.}\frac{1}{\lambda(\theta)^2 - s^2}\;,$$

hence

$$f_A(x+iy) \;=\; \frac{1}{4\pi^2}\,\text{p.v.}\int_{S^1}\frac{1}{b^2 + c^2\cos^2\theta - (x\cos\theta + y\sin\theta)^2}\,\mathrm{d}\theta\;. \tag{24}$$

It remains to compute the right-hand side of (24), which is a simple exercise in complex analysis. Setting $z = x + iy$ and $w = e^{2i\theta}$, we first observe that

$$b^2 + c^2\cos^2\theta - (x\cos\theta + y\sin\theta)^2 \;=\; \frac{1}{4w}\Big((c^2 - z^2) + 2w(a^2 + b^2 - |z|^2) + w^2(c^2 - \bar{z}^2)\Big)\;,$$

hence

$$f_A(z) \;=\; \frac{1}{i\pi^2}\,\text{p.v.}\oint_{|w|=1}\frac{1}{(c^2 - z^2) + 2w(a^2 + b^2 - |z|^2) + w^2(c^2 - \bar{z}^2)}\,\mathrm{d}w\;. \tag{25}$$

In (25), the roots of the denominator are

$$\begin{aligned} w_\pm \;&=\; \frac{1}{c^2 - \bar{z}^2}\Big(-a^2 - b^2 + |z|^2 \mp \sqrt{(a^2 + b^2 - |z|^2)^2 - (c^2 - z^2)(c^2 - \bar{z}^2)}\Big) \\ &=\; \frac{1}{c^2 - \bar{z}^2}\Big(-a^2 - b^2 + |z|^2 \mp \sqrt{4(a^2b^2 - b^2x^2 - a^2y^2)}\Big)\;. \end{aligned} \tag{26}$$

We can therefore distinguish between two cases:

1. The point $z = x + iy$ belongs to the interior of $W(A)$. Then the expression under the square root is positive, and it is easy to verify that $|w_+| > 1$, $|w_-| < 1$ (in the limiting case where $z = \pm c$ is a focus of the ellipse, one can set $w_- = 0$ and $w_+ = \infty$.) Thus the principal value in (25) is not needed, and the residue theorem shows that

$$f_A(z) \;=\; \frac{1}{\pi}\frac{1}{(a^2 + b^2 - |z|^2) + (c^2 - \bar{z}^2)w_-} \;=\; \frac{1}{2\pi}\frac{1}{\sqrt{a^2b^2 - b^2x^2 - a^2y^2}}\;. \tag{27}$$

2. The point $z = x + iy$ lies outside $W(A)$. Then the expression under the square root in (26) is negative, and one verifies that $|w_\pm| = 1$. Thus the integrand in (25) is holomorphic outside the unit circle $\{|w| = 1\}$ and decreases like $1/|w|^2$ at infinity. It follows that

$$I_z(r) \;:=\; \frac{1}{i\pi^2}\oint_{|w|=r}\frac{1}{(c^2 - z^2) + 2w(a^2 + b^2 - |z|^2) + w^2(c^2 - \bar{z}^2)}\,\mathrm{d}w \;=\; 0\;,$$

for all $r \neq 1$, hence $f_A(z) = \frac{1}{2}(I_z(1+) + I_z(1-)) = 0$.

Summarizing, we have shown that the numerical density of the matrix (23) is the function $f_A \in L^1(\mathbb{C})$ defined by (27) inside $W(A)$, and vanishing identically outside $W(A)$. Note that $\operatorname{sing\,supp}(\mu_A) = \partial W(A) = \Sigma_A$, in agreement with Lemma 2.3, and that $f_A(z)$ blows up when z converges to the boundary of $W(A)$ from inside. In particular f_A is not log-concave, in contrast to what happens when A is normal.

4.2 The radially symmetric case

It sometimes happens that the numerical range of a matrix $A \in \mathbf{M}_n(\mathbb{C})$ is a disk in the complex plane and that the numerical density is radially symmetric about the center. In such a case, it is possible to obtain a representation formula which is simpler than (22). Indeed, assume that the disk $W(A)$ is centered at the origin, and let $R > 0$ denote the numerical radius of A (or any larger positive number). We set

$$f_A(z) \;=\; F_A(R^2 - |z|^2) \;, \qquad z \in \mathbb{C} \;, \quad |z| \le R \;, \tag{28}$$

where $F_A : [0, R^2] \to \mathbb{R}_+$ has to be determined. Since the numerical measure is invariant under rotations about the origin, it follows from Proposition 4.1 that the projected measure $\mu_{H(\theta)}$ does not depend on θ. In analogy with (28), if f_H denotes the numerical density of $H(\theta)$ for any $\theta \in S^1$, we set

$$f_H(x) \;=\; F_H(R^2 - x^2) \;, \qquad x \in \mathbb{R} \;, \quad |x| \le R \;. \tag{29}$$

We then have the following result:

Proposition 4.3 *If the numerical density of a nonscalar matrix $A \in \mathbf{M}_n(\mathbb{C})$ is radially symmetric, the functions F_A, F_H defined in* (28), (29) *satisfy the relations*

$$F_H(t) \;=\; \int_0^t F_A(t-s)\frac{1}{\sqrt{s}}\,\mathrm{d}s \;, \qquad F_A(s) \;=\; \frac{1}{\pi}\frac{\mathrm{d}}{\mathrm{d}s}\int_0^s F_H(s-t)\frac{1}{\sqrt{t}}\,\mathrm{d}t \;, \tag{30}$$

for $t, s \in [0, R^2]$.

Proof. By the definition of the Radon transformation, we have $f_H(x) = \int_{\mathbb{R}} f_A(x+iy)\,\mathrm{d}y$ for all $x \in \mathbb{R}$. Using (28), (29) and the support property, we thus find

$$F_H(R^2 - x^2) \;=\; \int_{y^2 \le R^2 - x^2} F_A(R^2 - x^2 - y^2)\,\mathrm{d}y \;=\; \int_0^{R^2 - x^2} F_A(R^2 - x^2 - s)\frac{1}{\sqrt{s}}\,\mathrm{d}s \;,$$

for all $x \in [-R, R]$. Setting $t = R^2 - x^2$, we obtain the first relation in (30). So far, we have shown that $F_H = \pi^{1/2} I F_A$, where I is the Riesz potential

$$(If)(t) \;=\; \frac{1}{\sqrt{\pi}}\int_0^t f(t-s)\frac{1}{\sqrt{s}}\,\mathrm{d}s \;, \qquad t > 0 \;.$$

Now, it is well known that $(I^2 f)(t) = \int_0^t f(s)\,\mathrm{d}s$, see e.g. [12, Chapter V.5], thus the second relation in (30) follows from the first one. ■

Examples. Let $A = (a_{ij})$ be a complex matrix, and assume that there exists a nonzero integer k such that $a_{ij} = 0$ whenever $j - i \neq k$. Then $W(A)$ is a disk centered at the origin, and the numerical density f_A is radially symmetric. Indeed, given any $\theta \in S^1$, the map

$$x \;=\; \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \;\mapsto\; x_\theta \;=\; \begin{pmatrix} e^{i\theta}x_1 \\ \vdots \\ e^{in\theta}x_n \end{pmatrix}$$

is a measure-preserving isomorphism of the unit sphere $\partial\mathbb{B}^n$, and our assumption on A implies that $\langle Ax_\theta, x_\theta\rangle = e^{ik\theta}\langle Ax, x\rangle$ for all $x \in \partial\mathbb{B}^n$. This proves that a rotation of angle $k\theta$ about the origin does not affect the numerical measure μ_A. Since θ is arbitrary, the numerical density f_A is necessarily radially symmetric.

The simplest example in this category is the 2×2 Jordan block

$$A_2 \;=\; \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} . \tag{31}$$

Here $W(A_2) = \{z \in \mathbb{C} \,|\, |z| \le 1\}$, and applying (27) with $a = b = 1$ we find

$$f_{A_2}(z) \;=\; \frac{1}{2\pi}\,\frac{1}{\sqrt{1-|z|^2}}\,\mathbf{1}_{\{|z|<1\}} . \tag{32}$$

Alternatively, since the eigenvalues of $H_2 = \frac{1}{2}(A_2 + A_2^*)$ are ± 1, we have $f_{H_2} = \frac{1}{2}\,\mathbf{1}_{[-1,1]}$ and applying Proposition 4.3 we easily obtain (32).

As a more interesting application, consider the 3×3 matrix

$$A_3 \;=\; \begin{pmatrix} 0 & a & 0 \\ 0 & 0 & b \\ 0 & 0 & 0 \end{pmatrix} , \tag{33}$$

where $a, b \in \mathbb{C}$ and $|a| + |b| > 0$. Multiplying A_3 with a positive constant, we can assume that $|a|^2 + |b|^2 = 4$. Then $W(A_3) = \{z \in \mathbb{C} \,|\, |z| \le 1\}$, and

$$f_{A_3}(z) \;=\; \frac{1}{\pi}\,\log\frac{1+\sqrt{1-|z|^2}}{|z|}\,\mathbf{1}_{\{0<|z|<1\}} . \tag{34}$$

Indeed, the eigenvalues of $H_3 = \frac{1}{2}(A_3 + A_3^*)$ are 0 and ± 1, hence $f_{H_3}(x) = B[-1,0,1](x) = (1-|x|)_+$. Applying Proposition 4.3, we obtain (34) by a straightforward calculation. Alternatively, if we take $(a, b) = (0, 2)$, we can use the direct sum formula (10) to deduce (34) from (32). Here again, we have $\operatorname{sing\,supp}(\mu_{A_3}) = \partial W(A_3) \cup \{0\} = \Sigma_{A_3}$, in agreement with Lemma 2.3. Remark that $f_{A_3}(z) = \mathcal{O}((1-|z|)^{1/2})$ as $|z| \to 1-$, so that the singularity of the numerical measure at the boundary is weaker than it was for A_2. This reflects the fact that $f_{H_3}(x) = (1 - |x|)_+$ is Lipschitz continuous, whereas $f_{H_2}(x) = \mathbf{1}_{[-1,+1]}(x)$ had jump discontinuities. However, we observe that the (logarithmic) singularity of f_{A_3} at the origin is much stronger than the (square root) singularity at the boundary. As we shall see in Section 5, this is a nongeneric concentration phenomenon due to the fact that the component of the critical set Σ_{A_3} associated with the eigenvalue 0 of $H(\theta)$ is reduced to a single point. We have here the rare instance of an unbounded numerical density for a matrix of size $n \ge 3$.

Another interesting conclusion that can be drawn from this example is that the numerical density μ_A *does not* determine the matrix $A \in \mathbf{M}_n(\mathbb{C})$ up to unitary conjugations if $n \ge 3$. Indeed, if we set $(a, b) = (0, 2)$ and $(a, b) = (\sqrt{2}, \sqrt{2})$ in the definition (33) of A_3, the resulting matrices are not even similar, yet they have the same numerical density, given by (34).

5 The geometry of the singular set

We know from Lemma 2.3 that the numerical density of a matrix $A \in \mathbf{M}_n(\mathbb{C})$ is smooth outside the set Σ_A of all critical values of the numerical map Φ_A. In this section, we describe a few geometrical properties of the singular set Σ_A which will be needed to formulate our main results in Section 6. We do not claim much originality here: the material of this section is essentially borrowed from [19, 4, 16, 17], and is reproduced below for the reader's convenience.

As was shown by Kippenhahn [19], the singular set Σ_A has a natural description in terms of the eigenvalues $\lambda_1(\theta), \dots, \lambda_n(\theta)$ of the Hermitian matrices $H(\theta)$ defined in (2). To see that, we first recall that these eigenvalues can be numbered in such a way that they are real-analytic functions of $\theta \in \mathbb{R}$, see [26]. By analyticity, any two eigenvalues either coincide for all $\theta \in \mathbb{R}$ or cross at most a finite number of times on each compact interval. As a consequence, there exists an integer $m \le n$ such that $H(\theta)$ has exactly m distinct eigenvalues for all $\theta \in [0, \pi) \setminus \Theta$, where $\Theta \subset [0, \pi)$ is a finite set. Since $H(\theta + \pi) = -H(\theta)$, it follows that we can number the eigenvalues in such a way that

$$\sigma(H(\theta)) = \{\lambda_1(\theta), \dots, \lambda_m(\theta)\} , \qquad \text{for all } \theta \in \mathbb{R} ,$$

where $\lambda_1(\theta), \dots, \lambda_m(\theta)$ are pairwise distinct and have constant multiplicities outside the crossing set $\Theta + \pi\mathbb{Z}$. Let $\tau : \{1, \dots, m\} \to \{1, \dots, m\}$ be the permutation defined by

$$\{\lambda_{\tau(1)}(\theta), \dots, \lambda_{\tau(m)}(\theta)\} = \{-\lambda_1(\theta + \pi), \dots, -\lambda_m(\theta + \pi)\} , \tag{35}$$

for any $\theta \in [0, \pi) \setminus \Theta$ (hence for all $\theta \in \mathbb{R}$). If we decompose τ into disjoint cycles $\mathscr{C}_1, \dots, \mathscr{C}_k$, we can associate to each cycle $\mathscr{C}_J$ its *length* ℓ_J and its *multiplicity* m_J, the latter being defined as the multiplicity of $\lambda_j(\theta)$ as an eigenvalue of $H(\theta)$ for any $j \in \mathscr{C}_J$. By construction, we have

$$\ell_1 + \dots + \ell_k = m , \qquad \text{and} \quad \ell_1 m_1 + \dots + \ell_k m_k = n .$$

Moreover, if $j \in \mathscr{C}_J$, then $\lambda_j(\theta + \ell_J \pi) = (-1)^{\ell_J} \lambda_j(\theta)$, hence $\lambda_j(\theta)$ is periodic with period $\ell_J \pi$ if ℓ_J is even and $2\ell_J \pi$ if ℓ_J is odd. Note however that these periods are not necessarily minimal.

Now, we associate to each cycle $\mathscr{C}_J$ of the permutation τ a closed curve $C_J \subset \mathbb{C}$ defined by

$$C_J = \left\{ e^{i\theta}(\lambda_j(\theta) + i\lambda_j'(\theta)) \,\middle|\, \theta \in \mathbb{R} \right\} , \tag{36}$$

where $j \in \{1, \dots, m\}$ is any element of the cycle $\mathscr{C}_J$. Since $\lambda_j(\theta + 2\ell_J\pi) = \lambda_j(\theta)$, it is clear that C_J is indeed a closed curve, and the definition of the permutation τ shows that the right-hand side of (36) does not depend on the choice of $j \in \mathscr{C}_J$. Equivalently, we can define C_J as the union over all $j \in \mathscr{C}_J$ of the curve segments $\{e^{i\theta}(\lambda_j(\theta) + i\lambda_j'(\theta)) \,|\, \theta \in [0, \pi]\}$. Let

$$C_A = C_1 \cup \ldots \cup C_k \subset \mathbb{C} ,$$

and let $C_A' \subset \mathbb{C}$ be the *bitangent set* of C_A, namely the set of all line segments joining pairs of points of C_A at which C_A has the same tangent line. With these definitions, we have the following useful characterization of the singular set Σ_A:

Proposition 5.1 [17, Theorem 3.5] $\Sigma_A = C_A \cup C_A'$.

Proof. According to Lemma 2.2, Σ_A is the set of all complex numbers of the form $\langle Ax, x\rangle$ where $x \in \partial\mathbb{B}^n$ is a normalized eigenvector of the Hermitian matrix $H(\theta)$ for some $\theta \in [0, \pi]$. To describe that set, fix $j \in \{1, \dots, m\}$, $\theta_0 \in \mathbb{R}$, and for θ in a neighborhood of θ_0 let

$$z_j(\theta) = \langle Ax_j(\theta), x_j(\theta)\rangle ,$$

where $x_j(\theta)$ is a normalized eigenvector of $H(\theta)$ associated with the eigenvalue $\lambda_j(\theta)$ and depending smoothly on θ. General results in perturbation theory imply that such an eigenvector indeed exists [26]. Using the definition (2) of $H(\theta)$, we find

$$\begin{aligned}\mathrm{Re}\,(e^{-i\theta}z_j(\theta)) \;&=\; \mathrm{Re}\,\langle e^{-i\theta}Ax_j(\theta), x_j(\theta)\rangle \;=\; \langle H(\theta)x_j(\theta), x_j(\theta)\rangle \;=\; \lambda_j(\theta)\ ,\\ \mathrm{Im}\,(e^{-i\theta}z_j(\theta)) \;&=\; \mathrm{Im}\,\langle e^{-i\theta}Ax_j(\theta), x_j(\theta)\rangle \;=\; \langle H'(\theta)x_j(\theta), x_j(\theta)\rangle \;=\; \lambda_j'(\theta)\ ,\end{aligned}$$

since $\langle H(\theta)x_j'(\theta), x_j(\theta)\rangle + \langle H(\theta)x_j(\theta), x_j'(\theta)\rangle = 2\lambda_j(\theta)\,\mathrm{Re}\,\langle x_j(\theta), x_j'(\theta)\rangle = 0$ due to the normalization condition. Thus

$$z_j(\theta) \;=\; e^{i\theta}(\lambda_j(\theta) + i\lambda_j'(\theta))\ , \qquad \text{hence} \qquad z_j'(\theta) \;=\; ie^{i\theta}(\lambda_j(\theta) + \lambda_j''(\theta))\ . \tag{37}$$

These relations show that the curve $C_J = \{z_j(\theta)\,|\,\theta \in \mathbb{R}\}$ is tangent, at each point $z_j(\theta)$, to the straight line

$$L_j(\theta) \;=\; \{e^{i\theta}(\lambda_j(\theta) + i\alpha)\,|\,\alpha \in \mathbb{R}\} \;=\; \{z \in \mathbb{C}\,|\,\mathrm{Re}\,(ze^{-i\theta}) = \lambda_j(\theta)\}\ . \tag{38}$$

In other words C_J is the *envelope* of the family of straight lines $L_j(\theta)$, for any $j \in \mathscr{C}_J$. Since $C_J \subset \Sigma_A$ by construction, we have shown that Σ_A contains the curve $C_A = C_1 \cup \ldots \cup C_k$.

However, it is important to realize that Σ_A can be larger than C_A if the crossing set $\Theta \subset [0, \pi)$ defined above is nonempty. Indeed, assume that $\lambda_j(\theta_0) = \lambda_p(\theta_0)$ for some $\theta_0 \in \Theta$ and some $j, p \in \{1, \ldots, m\}$ with $j \neq p$. For θ in a neighborhood of θ_0, let $x_j(\theta), x_p(\theta)$ be smooth, normalized eigenvectors of $H(\theta)$ corresponding to $\lambda_j(\theta), \lambda_p(\theta)$ respectively. Using the same notations as above, we have $z_j(\theta_0) \neq z_p(\theta_0)$ in general, because $\lambda_j'(\theta_0) \neq \lambda_p'(\theta_0)$. Now, since $\langle x_j(\theta), x_p(\theta)\rangle = 0$ whenever $\lambda_j(\theta) \neq \lambda_p(\theta)$, we also have $\langle x_j(\theta_0), x_p(\theta_0)\rangle = 0$ by continuity. In particular, if $\alpha, \beta \in \mathbb{C}$ satisfy $|\alpha|^2 + |\beta|^2 = 1$, then $x = \alpha x_j(\theta_0) + \beta x_p(\theta_0)$ is a normalized eigenvector of $H(\theta_0)$, and a direct calculation yields

$$\mathrm{Im}\,\langle e^{-i\theta_0}Ax, x\rangle \;=\; \langle H'(\theta_0)x, x\rangle \;=\; |\alpha|^2\lambda_j'(\theta_0) + |\beta|^2\lambda_p'(\theta_0)\ ,$$

whereas $\mathrm{Re}\,\langle e^{-i\theta_0}Ax, x\rangle = \langle H(\theta_0)x, x\rangle = \lambda_j(\theta_0) = \lambda_p(\theta_0)$. This shows that Σ_A contains the line segment $[z_j(\theta_0), z_p(\theta_0)]$, which by construction is tangent to the curve C_J at $z_j(\theta_0)$ and to the curve C_P at $z_p(\theta_0)$. Repeating the same argument for all eigenvalue crossings, we conclude that Σ_A contains the whole bitangent set C_A'. Finally, it is clear from Lemma 2.2 that all points of Σ_A either belong to C_A or to C_A'. ■

Examples:

1. (*The generic case*) As is well known, within the space of all Hermitian matrices of size $n \geq 2$, the set of matrices having a multiple eigenvalue is a finite union of submanifolds of codimension at least three. This implies that, for a generic matrix $A \in \mathbf{M}_n(\mathbb{C})$, the Hermitian matrices $H(\theta)$ defined by (2) will have simple eigenvalues for all $\theta \in S^1$ [16]. In that situation, we denote by $\lambda_1(\theta) < \lambda_2(\theta) < \ldots < \lambda_n(\theta)$ the eigenvalues of $H(\theta)$. Using the notations introduced above, we have $m = n$ and the permutation τ defined in (35) is simply

$$\tau \;=\; \begin{pmatrix} 1 & 2 & \ldots & n \\ n & n-1 & \ldots & 1 \end{pmatrix}\ .$$

If $n = 2k$ is even, τ has k cycles with $\ell_1 = \ldots \ell_k = 2$; if $n = 2k-1$ is odd, τ has one fixed point and $k-1$ cycles of length 2. In all cases, the multiplicities $m_1, \ldots, m_k$ are all equal to 1, and the bitangent set C_A' is empty by assumption. Thus $\Sigma_A = C_A$ is the union of $k = [\frac{n+1}{2}]$ closed curves. It is not difficult to prove that the curve C_1 associated with the cycle $\mathscr{C}_1 = (1\ n)$ is smooth, strictly convex, and contains all the other curves $C_2, \ldots, C_k$ in its interior [19]. In particular, $C_1 = \partial W(A)$. Consider now the curve $C_J = \{z_j(\theta)\,|\,\theta \in \mathbb{R}\}$ for some $j \neq 1, n$. If $\delta_j(\theta) = \lambda_j(\theta) + \lambda_j''(\theta)$ is not identically zero, the formulas (37) show that the curvature of C_J at any regular point $z_j(\theta)$ is strictly positive:

$$\kappa_j(\theta) \;=\; \frac{1}{|z_j'(\theta)|} \;=\; \frac{1}{|\lambda_j(\theta) + \lambda_j''(\theta)|} \;>\; 0 \;.$$

This means that the tangent vector $z_j'(\theta)$ always rotates counterclockwise when θ is increased. Nevertheless, the whole curve C_J is not convex in general, because it may have a finite number of singular points corresponding to zeros of $\delta_j(\theta)$. As is easily verified, simple zeros of $\delta_j(\theta)$ correspond to *cusp points* of the curve C_J, see e.g. Fig. 1 where a generic example with $n = 3$ is represented. On the other hand, if $\delta_j(\theta)$ vanishes identically, then $z_j'(\theta) \equiv 0$ and the curve C_J reduces to a single point. Under our generic assumptions, this can happen only if n is odd and $j = (n+1)/2$. As an example, the singular set of the matrix A_3 defined in (33) includes the isolated point $\{0\}$.

2. (*The normal case*) In contrast with the previous example, we now consider the particular case where the matrix A is normal. If $\lambda_1, \ldots, \lambda_m$ denote the distinct eigenvalues of A, it is straightforward to verify that the eigenvalues of the Hermitian matrices $H(\theta)$ defined in (2) are simply $\lambda_j(\theta) = \mathrm{Re}\,(\lambda_j\, e^{-i\theta})$, $j = 1, \ldots, m$. In view of (37), this means that $z_j(\theta) = \lambda_j$ for all $\theta \in \mathbb{R}$, hence the curve C_j is reduced to the single point $\{\lambda_j\}$ for all $j = 1, \ldots, m$. Needless to say, the permutation τ defined by (35) is the identity here. It follows that $C_A = \sigma(A) = \{\lambda_1, \ldots, \lambda_m\}$, and proceeding as in the proof of Proposition 5.1 we easily see that C_A' is the set of all line segments joigning pairs of eigenvalues of A. This is in full agreement with the conclusions of Section 3.3.

Remark: We have seen that the curve C_A which generates the singular set Σ_A is the envelope of the family of straight lines $L_j(\theta)$ defined in (38). This geometric construction can be formulated in an equivalent way [1], which is more conceptual and worth mentioning here. Assume for simplicity that the matrix $A \in \mathbf{M}_n(\mathbb{C})$ is "generic" in the sense of Example 1 above, and let $a(\xi)$ be the homogeneous polynomial of degree n defined by

$$a(\xi) = \det\Bigl(\xi_0 I_n + \xi_1 A_1 + \xi_2 A_2\Bigr)\;, \qquad \xi = (\xi_0, \xi_1, \xi_2) \in \mathbb{R}^3\;,$$

where A_1, A_2 are as in (2). Since $a(\xi)$ has real coefficients, the equation $a(\xi) = 0$ defines, in homogeneous coordinates, an algebraic curve Γ in the projective plane $\mathbb{R}P^2$. Moreover, our genericity assumption on A implies that this curve is nonsingular: for each $\bar\xi \in \Gamma$, the tangent line to Γ at $\bar\xi$ is uniquely defined and satisfies, in homogeneous coordinates, an equation of the form $x_0\xi_0 + x_1\xi_1 + x_2\xi_2 = 0$. The set of all $x = (x_0, x_1, x_2)$ obtained in this way is again a curve Γ' in $\mathbb{R}P^2$ (the *dual* curve of Γ) given by the equation $b(x) = 0$ for some homogeneous polynomial b. As is shown in [9], the degree of b does not exceed $n(n-1)$ if $n \geq 2$. Now, it is rather straightforward to verify that the curve C_A defined as the envelope of the family of straight lines (38) is nothing but the restriction of the projective curve Γ' to the subspace $x_0 = 1$, namely $z = x + iy \in C_A$ if and only if $b(1, x, y) = 0$. Thus C_A is a real algebraic curve in $\mathbb{C} \simeq \mathbb{R}^2$ of degree at most $n(n-1)$ if $n \geq 2$. In the language of partial differential equations, the algebraic variety Γ is the *characteristic variety* of the symmetric hyperbolic system (4), and we shall see in Section 6 that the dual variety Γ' is related to the singular support of the fundamental solution of (4). Note that our genericity assumption on A precisely means that system (4) is *strictly hyperbolic*.

We conclude this section with a brief discussion of the number of tangent lines to the algebraic curve C_A which can be drawn from a given point. We recall that $C_A = C_1 \cup \ldots \cup C_k$, where each C_J is a closed curve associated with the cycle $\mathscr{C}_J$ of the permutation (35). For all $J \in \{1, \ldots, k\}$ and all $z \in \mathbb{C} \setminus C_J$, we denote by $N_J(z)$ the number of straight lines that are tangent to the curve C_J and contain the point z. Note that, since C_J was itself defined as the envelope of a family of straight lines, the tangent line to C_J is well defined even at singular points. In the degenerate case where C_J reduces to a single point $\{z_J\}$, the set of tangent lines should be understood as the pencil of all straight lines through z_J. Now, if $z \in \mathbb{C} \setminus C_A$, we denote by $N(z)$ the total number of tangents to the curve $C_A = C_1 \cup \ldots \cup C_k$ that can be drawn from the point z, with multiplicities taken into account:

$$N(z) \;=\; m_1 N_1(z) + \cdots + m_k N_k(z)\;, \qquad z \in \mathbb{C} \setminus C_A\;. \tag{39}$$

The following elementary properties of $N_J(z)$ and $N(z)$ will be useful.

Proposition 5.2 *For each $J \in \{1, \ldots, k\}$, the number $N_J(z)$ is constant in each connected component of $\mathbb{C} \setminus C_J$. Moreover $N(z) \le n$ for all $z \in \mathbb{C} \setminus C_A$.*

Proof. Fix $J \in \{1, \ldots, k\}$ and pick $j \in \mathscr{C}_J$. For any $z \in \mathbb{C} \setminus C_J$, $N_J(z)$ is the number of zeros of the function

$$f_j(\theta, z) \;=\; \lambda_j(\theta) - \mathrm{Re}\,(z e^{-i\theta})$$

for θ in the interval $[0, \ell_J \pi)$. When z is varied, this number can only change if $f_j(\theta, z)$ has a double zero for some θ, but this would mean that $z = z_j(\theta) = e^{i\theta}(\lambda_j(\theta) + i\lambda_j'(\theta)) \in C_J$, thus contradicting our assumption. Therefore $N_J(z)$ is necessarily constant in each connected component of $\mathbb{C} \setminus C_J$. On the other hand, we have the identity

$$\det\Big(H(\theta) - \mathrm{Re}\,(z e^{-i\theta}) I_n\Big) \;=\; \prod_{J=1}^{k} \prod_{j \in \mathscr{C}_J} f_j(\theta, z)^{m_J}\;. \tag{40}$$

Fix $z \in \mathbb{C} \setminus C_A$, and consider both sides of (40) as functions of θ. If multiplicities are taken into account, the number of zeros of the right-hand side on the interval $[0, \pi)$ is precisely $N(z)$. But the left-hand side, being a trigonometric polynomial of degree at most n, cannot have more than n zeros in $[0, \pi)$. This proves the claim. ■

Remark 5.3 *As was mentioned in the proof of Proposition 5.1, if $z \in \mathbb{C} \setminus C_A$ it is possible that $f_j(\theta, z) = f_p(\theta, z) = 0$ for some $\theta \in [0, \pi)$ and some $j \ne p$. This is the case, in particular, whenever $z \in C_A'$. However, if $z \in \mathbb{C} \setminus \Sigma_A$, we have*

$$\partial_\theta f_j(\theta, z) \cdot \partial_\theta f_p(\theta, z) \;>\; 0 \qquad \textit{whenever} \quad f_j(\theta, z) = f_p(\theta, z) = 0\;.$$

Indeed, replacing A with $A - zI_n$, we can assume without loss of generality that $z = 0$. If $\lambda_j(\theta) = \lambda_p(\theta) = 0$ for some $\theta \in [0, \pi)$ and some $j \ne p$, then $\lambda_j'(\theta)$ and $\lambda_p'(\theta)$ have necessarily the same sign, otherwise (37) *would imply that the origin belongs to the line segment $[z_j(\theta), z_p(\theta)]$, thus contradicting our assumption that $z \notin C_A'$.*

Proposition 5.2 asserts that the integer $N_J(z)$ can only change if z crosses the curve C_J. In fact, if the crossing occurs at a regular point $\bar{z} \in C_J$, it is not difficult to verify that $N_J(z)$ is decreased by two units if z crosses C_J from the convex to the concave side (i.e., in the direction of the local center

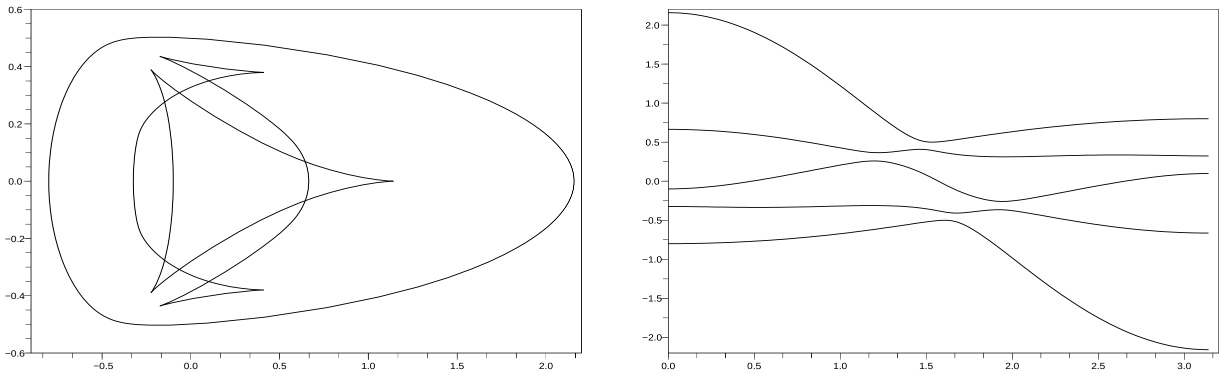

Figure 2: The singular set Σ_A (left) and the eigenvalues $\lambda_1(\theta), \dots, \lambda_5(\theta)$ (right) are represented for a typical matrix $A \in M_5(\mathbb{R})$. The singular set consists of three closed curves, one of which (the boundary of $W(A)$) is smooth, and the other two have cusps. In this example, the set Π_A defined in (46) has six connected components.

of curvature), and increased by two units if z crosses C_J in the converse direction, see [4, Section 4.1] or Section 6.2 below. These simple rules give an efficient algorithmic way to compute $N_J(z)$, and hence $N(z)$, in concrete examples. Consider for instance Fig. 2, where the singular set Σ_A of a generic matrix $A \in M_5(\mathbb{R})$ is represented. Here $\Sigma_A = C_A = C_1 \cup C_2 \cup C_3$, where C_1 is the boundary of $W(A)$, C_2 is the closed curve with two swallowtails, and C_3 is the triangle with three cusps. The set $\mathbb{C} \setminus \Sigma_A$ has 11 connected components, on which $N(z)$ is equal to 5, 3, or 1. Using the crossing rules above, it is easy to verify that $N(z) = 5$ in six different regions: inside both swallowtails, inside the three tips of the triangle, and outside $W(A)$. Such a result is definitely more cumbersome to obtain by counting directly the number of tangents to C_A from a given point.

6 Qualitative properties of the numerical density

Equipped with the results of Sections 4 and 5, we now derive some of the main properties of the numerical density f_A of an arbitrary matrix $A \in \mathbf{M}_n(\mathbb{C})$. We first establish an explicit formula for the derivatives of order $n-2$, which allows us to prove that f_A is polynomial in some distinguished regions of the complex plane which can be characterized geometrically. We next show that, for a generic matrix $A \in \mathbf{M}_n(\mathbb{C})$, the density f_A is of class C^{n-3} if $n \geq 3$. Finally, as announced in the introduction, we prove that the fundamental solution of the linear hyperbolic system (4) can be represented in terms of derivatives of the numerical density f_A. In particular, the lacunas of system (4) are precisely the polynomial regions described in Corollary 6.2.

6.1 Polynomial regions

Most of what we know about the numerical density of nonnormal matrices is based on the following result:

Proposition 6.1 *Let $A \in \mathbf{M}_n(\mathbb{C})$, where $n \geq 2$, and let $\mathcal{P} = P(\partial_x, \partial_y)$ be a homogeneous differential operator of degree $n-2$. Then, for all $z = x + iy \in \mathbb{C} \setminus \Sigma_A$, we have*

$$(\mathcal{P} f_A)(z) \;=\; -\frac{(n-1)!}{4\pi^2}\,\text{f.p.}\int_{S^1} \frac{P(\cos\theta, \sin\theta)}{\Delta(\theta, z)}\,\mathrm{d}\theta\;, \tag{41}$$

where

$$\Delta(\theta, z) = \det\Big(H(\theta) - \mathrm{Re}\,(ze^{-i\theta})I_n\Big)\ , \quad \theta \in S^1\ . \tag{42}$$

In (41) the symbol f.p. denotes the finite part of the integral in the sense of Hadamard, but in many situations it is sufficient to take simply the Cauchy principal value, as in (24). This is the case in particular for generic matrices in the sense of Section 5 (Example 1), because all multiplicities $m_1, \dots, m_k$ are all equal to one and formula (40) then shows that the map $\theta \mapsto \Delta(\theta, z)$ has only simple zeros on S^1 if $z \in \mathbb{C} \setminus \Sigma_A$.

In fact, using the analyticity of the integrand, it is possible to rewrite (41) in a slightly different form which is appropriate for further analysis. As in Section 4.1, we set $w = e^{2i\theta}$ and we observe that

$$P(\cos\theta, \sin\theta) = w^{1-\frac{n}{2}}\tilde{P}(w)\ , \qquad \Delta(\theta, z) = w^{-\frac{n}{2}}\tilde{\Delta}(w, z)\ ,$$

where (if $z = x + iy$)

$$\tilde{P}(w) = P\Big(\frac{w+1}{2}, \frac{w-1}{2i}\Big)\ , \qquad \tilde{\Delta}(w, z) = \det\left(\frac{A}{2} + \frac{wA^*}{2} - \Big(x\frac{w+1}{2} + y\frac{w-1}{2i}\Big)I_n\right)\ .$$

Thus (41) is equivalent to

$$\begin{aligned}
(\mathcal{P}f_A)(z) &= \frac{i(n-1)!}{4\pi^2}\,\mathrm{f.p.}\oint_{|w|=1}\frac{\tilde{P}(w)}{\tilde{\Delta}(w,z)}\,\mathrm{d}w \\
&\equiv \frac{i(n-1)!}{8\pi^2}\oint_{|w|=1-\epsilon}\frac{\tilde{P}(w)}{\tilde{\Delta}(w,z)}\,\mathrm{d}w + \frac{i(n-1)!}{8\pi^2}\oint_{|w|=1+\epsilon}\frac{\tilde{P}(w)}{\tilde{\Delta}(w,z)}\,\mathrm{d}w\ ,
\end{aligned} \tag{43}$$

where ϵ is any sufficiently small positive number, depending on z. In the particular case where $P = 1$ and $A \in \mathbf{M}_2(\mathbb{C})$ is given by (23), we recover (25).

Proof of Proposition 6.1. We first consider the generic situation where the Hermitian matrices $H(\theta)$ have simple eigenvalues $\lambda_1(\theta) < \lambda_2(\theta) < \dots < \lambda_n(\theta)$ for all $\theta \in S^1$. In that case, the B-spline representing the numerical density of $\mu_H(\theta)$ can be given an explicit expression using the divided difference formula (14), (16):

$$B[\lambda_1(\theta), \dots, \lambda_n(\theta)](s) = (n-1)\sum_{j=1}^{n}\frac{(\lambda_j(\theta) - s)_+^{n-2}}{\prod_{k\neq j}(\lambda_j(\theta) - \lambda_k(\theta))}\ .$$

In particular, differentiating $(n-1)$ times with respect to s, we find

$$B^{(n-1)}[\lambda_1(\theta), \dots, \lambda_n(\theta)](s) = (n-1)!\sum_{j=1}^{n}\frac{\delta(s - \lambda_j(\theta))}{\prod_{k\neq j}(\lambda_k(\theta) - \lambda_j(\theta))}\ . \tag{44}$$

Let $\mathcal{P} = P(\partial_x, \partial_y)$ be a homogeneous differential operator of degree $n-2$. From the representation formula (22), we deduce at least formally

$$(\mathcal{P}f_A)(x+iy) = \frac{1}{4\pi}\int_{S^1}P(\cos\theta, \sin\theta)\,\mathcal{H}B^{(n-1)}[\lambda_1(\theta), \dots, \lambda_n(\theta)](x\cos\theta + y\sin\theta)\,\mathrm{d}\theta\ . \tag{45}$$

To evaluate the integrand in the right-hand side, we start from (44) and recall that the Hilbert transform (with respect to the variable s) of the Dirac measure $\delta(s-\lambda)$ is the distribution $\frac{1}{\pi}\,\text{p.v.}\frac{1}{s-\lambda}$. We also use the identity

$$\sum_{j=1}^{n} \frac{1}{\mu_j \prod_{k\neq j}(\mu_k-\mu_j)} \;=\; \prod_{j=1}^{n}\frac{1}{\mu_j}\;,$$

which holds for any collection of pairwise distinct nonzero complex numbers $\mu_1,\dots,\mu_n$. We thus find

$$\mathcal{H}B^{(n-1)}[\lambda_1(\theta),\dots,\lambda_n(\theta)](s) \;=\; -\frac{(n-1)!}{\pi}\,\text{p.v.}\prod_{j=1}^{n}\frac{1}{\lambda_j(\theta)-s} \;=\; -\frac{(n-1)!}{\pi}\,\text{p.v.}\frac{1}{\det(H(\theta)-sI_n)}\;.$$

Setting $s = x\cos\theta + y\sin\theta$ and inserting this expression into (45), we obtain (41).

The calculations so far are formal, but they can be justified if we assume that $z = x+iy \in \mathbb{C}\setminus\Sigma_A$. In that case, we know from the proof of Proposition 5.2 that the map $\theta \mapsto \Delta(\theta,z)$ has only simple zeros on S^1, because this is the case for each of the factors $f_j(\theta,z)$ in (40) and, by assumption, the eigenvalues $\lambda_j(\theta)$ are all distinct for $\theta\in S^1$. It follows that the integral in (41) is well-defined in the sense of Cauchy's principal value, and depends smoothly on $z\in\mathbb{C}\setminus\Sigma_A$. This in turn implies that the density f_A given by (22) is smooth on $\mathbb{C}\setminus\Sigma_A$, as it should be, and that the calculations above are correct.

To conclude the proof of Proposition 6.1, it remains to verify that (41) or (43) holds for an arbitrary matrix $A\in\mathbf{M}_n(\mathbb{C})$. To do that, we first observe that the singular set $\Sigma_A\subset\mathbb{C}$ is an upper-semicontinuous function of A in the sense that

$$\delta(\Sigma_B,\Sigma_A) \;\equiv\; \sup_{z\in\Sigma_B}\,\text{dist}(z,\Sigma_A) \;\xrightarrow[B\to A]{}\; 0\;.$$

Moreover the numerical density f_A, together with its derivatives, depends continuously on A in $\mathbb{C}\setminus\Sigma_A$. These rather classical facts can be established using, for instance, the representation formula (8) (we omit the details). On the other hand, it is not difficult to verify that the right-hand side of (43) depends continuously on A for each $z\in\mathbb{C}\setminus\Sigma_A$. The crucial point here is that the polynomial map $w\mapsto\tilde{\Delta}(w,z)$ keeps the same number of zeros on the unit circle (counted with multiplicities) if the matrix A is slightly varied; in particular, we can choose the same $\epsilon>0$ in (43) for all matrices in a neighborhood of A. This property can be established using the factorization (40), Remark 5.3, and general results for perturbations of eigenvalues of Hermitian matrices, see [18]. Now, given $A\in\mathbf{M}_n(\mathbb{C})$, there exists a sequence $\{A_\ell\}_{\ell\in\mathbb{N}}$ of generic matrices converging to A as $\ell\to\infty$. If $z\in\mathbb{C}\setminus\Sigma_A$, we know that equation (43) holds for A_ℓ if ℓ is sufficiently large, hence taking the limit $\ell\to\infty$ and using the continuity properties mentioned above we obtain the desired equality. ■

As a consequence of Proposition 6.1, we now establish an important property of the numerical density in the regions of the complex plane where the number $N(z)$ defined in (39) takes its maximal value n.

Corollary 6.2 *Given $A\in\mathbf{M}_n(\mathbb{C})$, let*

$$\Pi_A \;=\; \Big\{z\in\mathbb{C}\setminus\Sigma_A \,\Big|\, N(z)=n\Big\}\;. \tag{46}$$

If $n\geq 3$, the numerical density f_A is polynomial of degree at most $n-3$ on each connected component of Π_A. If $n=2$, then $f_A=0$ on Π_A.

Proof. We shall prove that $\mathcal{P}f_A$ vanishes identically on Π_A for any homogeneous differential operator of order $n-2$. Indeed, in view of (40), (42), the assumption $z \in \Pi_A$ implies that the map $\theta \mapsto \Delta(\theta, z)$ has exactly n zeros (counting multiplicities) on $[0, \pi)$. Equivalently, the polynomial $\tilde{\Delta}(w, z)$ has exactly n zeros on the unit circle $\{|w| = 1\}$. But since $\tilde{\Delta}(w, z)$ has degree n, this polynomial has no zeros outside the unit circle, and using Cauchy's theorem we conclude that the first integral in the last member of (43) vanishes. The second integral is also zero, because the numerator is a polynomial of degree at most $n-2$, while the denominator has degree exactly n, hence the integrand $\tilde{P}(w)/\tilde{\Delta}(w, z)$ decays at least like $|w|^{-2}$ as $|w| \to \infty$. Thus $\mathcal{P}f_A \equiv 0$ on Π_A, and the conclusion follows. ■

Remark 6.3 *If $z = x + iy \in \sigma(A)$, then the polynomial $\tilde{\Delta}(w, z)$ has degree strictly less than n, and it follows from the above proof that $z \notin \Pi_A$. Thus $\sigma(A) \cap \Pi_A = \emptyset$.*

The set Π_A is never empty, because it always contains the complement of the numerical range $W(A)$, where the density f_A vanishes identically [26, 4]. Moreover, in many situations, one or several components of Π_A are contained in $W(A)$, in which case Corollary 6.2 gives nontrivial informations on the numerical density. For instance, if $A \in \mathbf{M}_n(\mathbb{C})$ is a normal matrix whose numerical range has nonempty interior, then $N(z) = n$ for all $z \in \mathbb{C} \setminus \sigma(A)$, and it follows from Corollary 6.2 that f_A is polynomial of degree at most $n-3$ in each connected component of $\mathbb{C} \setminus \Sigma_A$, in agreement with Proposition 3.2. In the same spirit, if $A = A_1 \oplus A_2$ with $A_1 \in \mathbf{M}_{n_1}(\mathbb{C})$ and $A_2 \in \mathbf{M}_{n_2}(\mathbb{C})$, it is easily verified that $N(z) = n = n_1 + n_2$ if $z \notin W(A_1) \cup W(A_2)$, hence f_A is piecewise polynomial outside $W(A_1) \cup W(A_2)$. Finally, Fig. 1 shows a typical example of a matrix $A \in \mathbf{M}_3(\mathbb{C})$ for which Π_A has a component inside $W(A)$, on which the density f_A is identically constant by Corollary 6.2.

6.2 Generic regularity results

Our purpose here is to establish regularity results for the numerical density f_A in the whole complex plane, and not only outside the singular set Σ_A. For simplicity, we assume henceforth that our matrix $A \in \mathbf{M}_n(\mathbb{C})$ enjoys the following (generic) properties:

H1: The eigenvalues of (2) satisfy $\lambda_1(\theta) < \ldots < \lambda_n(\theta)$ for all $\theta \in S^1$.

H2: For all $j \in \{1, \ldots, n\}$, the function $\lambda_j(\theta) + \lambda_j''(\theta)$ is not identically zero.

The second assumption guarantees that the curve (36) associated with λ_j is not reduced to a single point. This of course is possible only if $n \geq 2$, and the two-dimensional case $n = 2$ is completely treated in Section 4.1. So we can assume that $n \geq 3$, and we have the following result:

Proposition 6.4 *Assume that $A \in \mathbf{M}_n(\mathbb{C})$ satisfies H1, H2 above. If $n \geq 3$, the numerical density $f_A : \mathbb{C} \to \mathbb{R}_+$ is of class C^{n-3}.*

We have already seen that both hypotheses H1, H2 are necessary, in general, for the conclusion of Proposition 6.4 to hold. For instance, if $A \in M_3(\mathbb{C})$ is a normal matrix whose numerical range $W(A)$ has nonempty interior, the numerical density f_A is proportional to the characteristic function of $W(A)$ and is therefore discontinuous on $\partial W(A)$. A more subtle example is provided by the matrix A_3 defined in (33): if $|a|^2 + |b|^2 = 4$, we have here $\lambda_1(\theta) = -1$, $\lambda_2(\theta) = 0$, $\lambda_3(\theta) = 1$ for all $\theta \in S^1$. Thus H1 is satisfied, but obviously not H2, and the explicit formula (34) shows that the numerical density of A_3 is discontinuous at the origin.

Proof of Proposition 6.4. Assume that $n \geq 3$ and that $A \in \mathbf{M}_n(\mathbb{C})$ satisfies H1, H2. If $\mathcal{Q} = Q(\partial_x, \partial_y)$ is a homogeneous differential operator of degree $n-3$, we have as in (45):

$$(\mathcal{Q}f_A)(x+iy) \;=\; \frac{1}{4\pi}\int_{S^1} Q(\cos\theta,\sin\theta)\,\mathcal{H}B^{(n-2)}[\lambda_1(\theta),\dots,\lambda_n(\theta)](x\cos\theta + y\sin\theta)\,\mathrm{d}\theta\;, \tag{47}$$

where

$$B^{(n-2)}[\lambda_1(\theta),\dots,\lambda_n(\theta)](s) \;=\; -(n-1)!\sum_{j=1}^{n}\frac{\mathbf{H}(\lambda_j(\theta)-s)}{\prod_{k\neq j}(\lambda_k(\theta)-\lambda_j(\theta))}\;. \tag{48}$$

Here $\mathbf{H} : \mathbb{R} \to [0,1]$ denotes the Heaviside function. Taking the Hilbert transform of (48) with respect to s and using (47), we thus find

$$\begin{aligned}(\mathcal{Q}f_A)(z) \;&=\; \frac{(n{-}1)!}{4\pi^2}\sum_{j=1}^{n}\int_{S^1}\frac{Q(\cos\theta,\sin\theta)}{\prod_{k\neq j}(\lambda_k(\theta)-\lambda_j(\theta))}\,\log|f_j(\theta,z)|\,\mathrm{d}\theta\\ &\equiv\; \sum_{j=1}^{n}\int_{S^1} h_j(\theta)\log|f_j(\theta,z)|\,\mathrm{d}\theta\;,\end{aligned} \tag{49}$$

where $f_j(\theta,z) = \lambda_j(\theta) - \mathrm{Re}(ze^{-i\theta})$ and $h_j : S^1 \to \mathbb{R}$ is a smooth function. Note that the integrand in (49) is 2π-periodic, because by assumption H1 this is the case for all eigenvalues $\lambda_1(\theta),\dots,\lambda_n(\theta)$.

It remains to show that each integral in the right-hand side of (49) defines a continuous function of $z \in \mathbb{C}$. Replacing A by $A - zI_n$ (an operation which does not affect the properties H1, H2), we see that it is sufficient to prove continuity at $z = 0$. This in turn is obvious if $\lambda_j(\theta) = f_j(\theta,0)$ does not vanish, so from now on we focus on the case where $\lambda_j(\theta)$ has (isolated) zeros on S^1. Using a partition of unity, we can treat each zero separately, so it is sufficient to consider the case where $\lambda_j(\theta)$ has a single zero of order $q \geq 1$ at $\theta = 0$, and $h_j(\theta)$ is localized near the origin. By analyticity, for z close to zero we have the factorization

$$f_j(\theta,z) \;=\; g(\theta,z)\prod_{p=1}^{q}(\theta-\mu_p(z))\;,$$

where the (possibly complex) roots $\mu_p(z)$ depend continuously on z, with $\mu_p(0) = 0$, and $g(\cdot,z)$ does not vanish in a neighborhood of zero. The quantity we have to study is therefore

$$\sum_{p=1}^{q}\int_{\mathbb{R}} h_j(\theta)\log|\theta-\mu_p(z)|\,\mathrm{d}\theta + \int_{\mathbb{R}} h_j(\theta)\log|g(\theta,z)|\,\mathrm{d}\theta\;.$$

The last term is clearly a continuous function of z. In the integral involving μ_p, we make the change of variables $\theta = t + \mathrm{Re}\,\mu_p(z)$ and observe that $\log|t| \leq \log|t - i\,\mathrm{Im}\,\mu_p(z)| \leq 0$ in a neighborhood of $(t,z) = (0,0)$, hence continuity with respect to z follows from Lebesgue's dominated convergence theorem. ■

Once continuity of the derivatives of order $n-3$ has been established, we can obtain further regularity results for the numerical density by using the representation formula (41) or (43). In view of Lemma 2.3, it is sufficient to study the density in a neighborhood of a point $\bar{z} \in \Sigma_A$. Since $\Sigma_A = C_A = C_1 \cup \ldots \cup C_k$ by assumption H1, there exists $j \in \{1,\dots,n\}$ such that $\bar{z} \in C_j$, and for simplicity we also assume that $\bar{z} \notin C_p$ for all $p \neq j$. This means that the function $\theta \mapsto f_j(\theta,\bar{z})$ has a zero of order $m \geq 2$ at some point $\bar{\theta} \in S^1$, and only simple zeros for $\theta \neq \bar{\theta} \mod \pi$; moreover, if $p \neq j$, $\theta \mapsto f_p(\theta, \bar{z})$ has only simple zeros and does not vanish for $\theta = \bar{\theta}$.

Without loss of generality, we assume from now on that $\bar{z} = 0$ and $\bar{\theta} = 0$, and we first consider the simplest case where $m = 2$. This means that $\lambda_j(\theta) = \frac{\alpha}{2}\theta^2 + \mathcal{O}(\theta^3)$ near $\theta = 0$, for some $\alpha \neq 0$. If $z = x+iy$ is sufficiently small, it follows that the analytic function $\theta \mapsto f_j(\theta) = \lambda_j(\theta) - (x\cos\theta + y\sin\theta)$ has exactly two zeros $\theta_\pm(z)$ in a neighborhood of the origin, which satisfy

$$\theta_\pm(z) \;=\; \frac{1}{\alpha}\Big(y \pm \mathrm{sign}(\alpha)\sqrt{y^2 + 2\alpha x}\Big) + \mathcal{O}(|x| + |y|^2)\ .$$

The critical curve C_j, which is the set of all points z for which $\theta_+(z) = \theta_-(z)$, is therefore given by the equation $x = -\frac{1}{2\alpha}y^2 + \mathcal{O}(y^3)$ in a small ball B around the origin. Moreover $B \setminus C_j = B_r \cup B_c$, where B_r is the set of all $z \in B$ for which the roots $\theta_\pm(z)$ are real and distinct, whereas $z \in B_c$ when $\theta_\pm(z)$ are complex conjugate with nonzero imaginary part. As is easily verified, the local center of curvature of C_j is located on the side of B_c.

Now, let $\mathcal{P} = P(\partial_x, \partial_y)$ be a homogeneous differential operator of degree $n-2$, and consider the expression of $(\mathcal{P}f_A)(z)$ given by (43). Assume that, in the right-hand side, the parameter $\epsilon > 0$ is chosen in such a way that the slit annulus $\mathcal{A}_\epsilon = \{w \in \mathbb{C}\,|\, 0 < |1 - |w|| < \epsilon\}$ contains the points $w_\pm(z) = e^{2i\theta_\pm(z)}$ for all $z \in B_c$, but that $\mathcal{A}_{2\epsilon}$ does not contain any other root of the determinant $\tilde{\Delta}(w,z)$ for $z \in B$ (these conditions are easily achieved by choosing first ϵ and then B sufficiently small). Under these assumptions, the right-hand side of (43) defines a smooth function of $z \in B$, which coincides with $(\mathcal{P}f_A)(z)$ if $z \in B_r$ but *not* if $z \in B_c$. Indeed, in the latter case, we have to consider in addition the roots $w_\pm(z)$ of $\tilde{\Delta}(w,z)$ which are not taken into account by the fixed integration contours in (43) since $|1 - |w_\pm(z)|| < \epsilon$. Using Cauchy's theorem, we easily obtain

$$(\mathcal{P}f_A)(z) = (\mathcal{P}f_A)_{\mathrm{reg}}(z) + \frac{(n{-}1)!}{4\pi}\,\frac{\tilde{Q}(w_-(z),z) + \tilde{Q}(w_+(z),z)}{w_-(z) - w_+(z)}\ , \quad z \in B_c\ ,$$

where $(\mathcal{P}f_A)_{\mathrm{reg}}(z)$ denotes the regular part of $(\mathcal{P}f_A)(z)$, given by the right-hand side of (43) with fixed ϵ, and

$$\tilde{Q}(w,z) \;=\; \frac{\tilde{P}(w)(w - w_+(z))(w - w_-(z))}{\tilde{\Delta}(w,z)}\ .$$

In particular, we have for all $z \in B_c$

$$\Big|(\mathcal{P}f_A)(z) - (\mathcal{P}f_A)_{\mathrm{reg}}(z)\Big| \;\leq\; \frac{C}{|w_+(z) - w_-(z)|} \;\leq\; \frac{C}{|\theta_+(z) - \theta_-(z)|} \;\leq\; \frac{C}{\mathrm{dist}(z, C_j)^{1/2}}\ ,$$

and this estimate is sharp if $\tilde{P}(1) = P(1,0) \neq 0$, because in that case $\tilde{Q}(w,z)$ does not vanish near $w = 1$ if $z \in B$. Summarizing, we have reached the following important conclusion: Near a regular point $\bar{z} \in C_j$ of the critical set Σ_A of a generic matrix $A \in \mathbf{M}_n(\mathbb{C})$, the derivatives of order $n-2$ of the numerical density f_A are smooth on the convex side of the curve C_j, and blow up like $\mathrm{dist}(z, C_j)^{-1/2}$ on the concave side. If $n = 2$, this is in full agreement with the explicit formula (27) obtained in Section 4.1. If $n \geq 3$, we deduce after integrating that the numerical density f_A is of class $C^{n-5/2}$ in a neighborhood of such a point $\bar{z}$.

Using the same techniques, it is also possible to study the regularity of the numerical density near more singular points $\bar{z} \in \Sigma_A$. On the typical example represented in Fig. 2, we see that the following two cases have to be analyzed:

i) *Crossings*, which occur when two critical curves C_j and C_p intersect transversally at $\bar{z}$. This situation can be treated exactly as before, except that one has to consider four distinct regions near $\bar{z}$, instead of two.

ii) *Cusps*, which arise whenever one of the functions $\lambda_j(\theta) + \lambda_j''(\theta)$ has a simple zero. Here we can repeat the analysis above, assuming that $\lambda_j(\theta) = \frac{\alpha}{3}\theta^3 + \mathcal{O}(\theta^4)$ near $\theta = 0$. In a neighborhood of $\bar{z} = 0$, we find that $f_j(\theta, z)$ has either three real roots (for z inside a cuspidal domain with tip at $\bar{z}$), or one real and two complex conjugate roots (outside the cusp). Using the same argument as before, we conclude that $\mathcal{P}f_A$ is smooth inside the cusp, but blows up on the other side of the critical curve C_j. Altogether, the numerical density f_A is of class $C^{n-8/3}$ near the cusp, see [26, Section 4.3] for a similar analysis of the singularities of the fundamental solution of (4).

Under generic assumptions on the matrix A, all intersections are transversal and all cusps are non degenerate, so that the singular set C_A is a generic curve in the sense of real algebraic geometry, and the singularities of the numerical density f_A can be completely analyzed using the techniques described bove. In particular, the derivatives of order $n-2$ of the numerical density f_A are locally integrable, and since we know from Proposition 6.4 that $f_A \in C^{n-3}$ we conclude that the relation (41) holds everywhere (in the sense of distributions), and not only on the complement of Σ_A.

6.3 Connexion with the fundamental solution

As was explained in the introduction, the numerical measure of a matrix $A \in \mathbf{M}_n(\mathbb{C})$ is related to the fundamental solution of the hyperbolic system (4). This connexion can be established rigorously by comparing the expression (41) for the derivatives of the numerical density f_A with the representation formulas for the fundamental solution $E(t,x)$ of (4), which can be found e.g. in [4, 26].

In what follows, we identify $\mathbb{C}$ with $\mathbb{R}^2$, and we denote by $x = (x_1, x_2)$ the points of the Euclidean plane. With a slight abuse of notation, we write $f_A(x)$ instead of $f_A(z)$, and we consider as subsets of $\mathbb{R}^2$ the various regions associated with A, such as $W(A)$ or Σ_A. To derive a representation formula for the fundamental solution $E(t,x)$, we take the Radon transform of (5) with respect to $x \in \mathbb{R}^2$, and obtain the one-dimensional hyperbolic system

$$\partial_t \tilde{E}(t,s,\theta) + H(\theta)\partial_s \tilde{E}(t,s,\theta) \;=\; I_n\, \delta_{t=0} \otimes \delta_{s=0}\ ,$$

where $H(\theta)$ is given by (2) and $\tilde{E}(t,s,\theta)$ denotes the Radon transform of $E(t,x)$. Using the method of characteristics, we easily find

$$\tilde{E}(t,s,\theta) \;=\; \sum_{j=1}^{n} \delta(s - t\lambda_j(\theta)) P_j(\theta)\ , \quad t \ge 0,$$

where $\lambda_1(\theta), \dots, \lambda_n(\theta)$ are the eigenvalues of $H(\theta)$ and $P_1(\theta), \dots, P_n(\theta)$ the corresponding spectral projections. Now, if we invert the Radon transform as in (22) and use the identity

$$\sum_{j=1}^{n} \frac{1}{s - t\lambda_j(\theta)}\, P_j(\theta) \;=\; (sI_n - tH(\theta))^{-1}\ ,$$

we arrive at the representation formula

$$E(t,x) \;=\; -\frac{1}{4\pi^2}\,\mathrm{f.p.}\int_{S^1} \Big((x_1\cos\theta + x_2\sin\theta)I_n - tH(\theta)\Big)^{-2} \mathrm{d}\theta\ , \quad t \ge 0\ , \tag{50}$$

which coincides with Eq. (4.4a) in [4]. Arguing as in Section 6.1, one can show that equation (50) is rigorously satisfied for all $(t,x) \in (0,\infty) \times \mathbb{R}^2$ with $\frac{x}{t} \in \mathbb{R}^2 \setminus \Sigma_A$, and that $E(t,x)$ is smooth in that region of space-time.

To compare the numerical density with the fundamental solution, it is natural to extend f_A to a homogeneous function of space and time by setting

$$\mathcal{F}_A(t,x) \,=\, t^{n-3} f_A\Bigl(\frac{x}{t}\Bigr)\,, \qquad t > 0\,, \quad x \in \mathbb{R}^2\,. \tag{51}$$

We then have the following result:

Proposition 6.5 *For any $A \in \mathbf{M}_n(\mathbb{C})$, there exists a matrix-valued homogeneous polynomial Q of degree $n-1$ such that*

$$E(t,x) \,=\, Q(\partial_t, \partial_{x_1}, \partial_{x_2}) \mathcal{F}_A(t,x)\,, \tag{52}$$

for all $(t,x) \in (0,\infty) \times \mathbb{R}^2$ with $\frac{x}{t} \in \mathbb{R}^2 \setminus \Sigma_A$.

In particular, if Ω is a connected component of the region $\Pi_A \subset \mathbb{R}^2$ defined by (46), we know from Corollary 6.2 that $f_A(x)$ is polynomial of degree at most $n-3$ in Ω, and (51) then shows that $\mathcal{F}_A(t,x)$ is also polynomial of degree at most $n-3$ in the half-cone $C_+(\Omega) = \{(t,x) \in (0,\infty) \times \mathbb{R}^2 \,|\, \frac{x}{t} \in \Omega\}$. By Proposition 6.5, we conclude that $E(t,x) = 0$ in $C_+(\Omega)$. We thus have

Corollary 6.6 *The fundamental solution $E(t,x)$ of a matrix $A \in \mathbf{M}_n(\mathbb{C})$ vanishes for all $(t,x) \in (0,\infty) \times \mathbb{R}^2$ such that $\frac{x}{t}$ belongs to the region $\Pi_A \subset \mathbb{R}^2$ defined by (46).*

In the language of partial differential equations, the *domain of influence* of the origin for system (4) is the half-cone $C_+(D) = \{(t,x) \in (0,\infty) \times \mathbb{R}^2 \,|\, \frac{x}{t} \in D\}$, where $D \subset \mathbb{R}^2$ is the complement of the largest connected open region on which $E_* = E(1,\cdot)$ vanishes. If in addition $E_* = 0$ in some open region $L \subset D$, we say that L is a *lacuna* of the hyperbolic system (4). With this terminology, Corollary 6.6 asserts that each connected component of Π_A either lies outside the domain of influence of the origin, or is a lacuna. This important geometric characterization of lacunas is originally due to Petrovsky [22], and was thoroughly discussed in [4, 26] and from a more algebraic point of view in [2, 3]. Strictly speaking, it gives only a sufficient condition for the occurence of lacunas, but further work allows to show that all *stable* lacunas satisfy this criterion. In other words, in generic situations, the converse of Corollary 6.6 also holds: each open region on which E_* vanishes (in particular, any lacuna) belongs to Π_A.

As a typical example, consider the matrix $A \in M_3(\mathbb{C})$ defined by (55), whose numerical density is represented in Fig. 1. The polynomial region Π_A has just two components here: the exterior of $D = W(A)$, where both f_A and E_* vanish, and the interior of the cuspidal triangle, where $E_* = 0$ and f_A is identically constant, which is therefore a lacuna. A more complicated situation is depicted in Fig. 2, where the polynomial region Π_A has six connected components, among which five correspond to lacunas. In a different spirit, it is also interesting to consider the nongeneric matrix (57) whose numerical density is studied in Section 7. The domain of influence of the origin is the union of the closed unit disk and a single point $\{a\}$, so the numerical range $W(A) = \mathrm{conv}(D)$ is substantially larger than D if $|a| > 1$, see Fig. 5. In that case, the region Π_A has again two components, none of which is a lacuna: the exterior of $W(A)$, and the interior of the triangular region $W(A) \setminus D$ where f_A is identically constant. We see on these examples that the numerical density allows to distinguish between various regions where the fundamental solution vanishes identically, and which are nevertheless of rather different nature.

Before proving Proposition 6.5, we briefly verify its conclusion on a simple example. If $A \in M_2(\mathbb{C})$ is defined by (31), system (4) reduces to

$$\partial_t u_1 + \partial_{x_1} u_2 - i\partial_{x_2} u_2 \,=\, 0\;, \qquad \partial_t u_2 + \partial_{x_1} u_1 + i\partial_{x_2} u_1 \,=\, 0\;.$$

Combining both equations, one verifies that $\partial_t^2 u_j = \Delta u_j$ for $j = 1, 2$, hence the fundamental solution $E(t, x)$ can easily be computed using Poisson's formula for the solution of the wave equation in two dimensions. In agreement with Proposition 6.5, the result is:

$$E(t,x) \,=\, \begin{pmatrix} \partial_t & -\partial_{x_1} + i\partial_{x_2} \\ -\partial_{x_1} - i\partial_{x_2} & \partial_t \end{pmatrix} \mathcal{F}_A(t,x)\;, \qquad |x| < t\;,$$

where according to (32), (51)

$$\mathcal{F}_A(t,x) \,=\, \frac{1}{t} f_A\Bigl(\frac{x}{t}\Bigr) \,=\, \frac{1}{2\pi}\,\frac{1}{\sqrt{t^2 - |x|^2}}\,\mathbf{1}_{\{|x|<t\}}\;.$$

Proof of Proposition 6.5. If $n = 1$ the conclusion is trivial, because both E_* and f_A vanish identically outside Σ_A (which is reduced to a single point), so we assume henceforth that $n \ge 2$. If $\mathcal{P} = P(\partial_t, \partial_{x_1}, \partial_{x_2})$ is a homogeneous differential operator of degree $n - 2$, then using (51) it is straightforward to verify that

$$(\mathcal{P}\mathcal{F}_A)(t,x) \,=\, \frac{1}{t}\Bigl[P\Bigl(-\frac{x}{t}\cdot\nabla_\xi, \partial_{\xi_1}, \partial_{\xi_2}\Bigr) f_A\Bigr]\Big|_{\xi = \frac{x}{t}}\;, \tag{53}$$

whenever $\frac{x}{t} \in \mathbb{R}^2 \setminus \Sigma_A$. Here $\xi \in \mathbb{R}^2$ denotes the argument of the function f_A, which has to be replaced by $\frac{x}{t}$ after differentiation. We warn the reader that equality (53) holds only if P is of degree $n - 2$. Applying Proposition 6.1, we deduce that

$$P(\partial_t, \partial_{x_1}, \partial_{x_2})\mathcal{F}_A(t,x) \,=\, -\frac{(n{-}1)!}{4\pi^2 t}\,\text{f.p.}\int_{S^1} \frac{P(-\frac{x}{t}\cdot e_\theta, \cos\theta, \sin\theta)}{\det(H(\theta) - \frac{x}{t}\cdot e_\theta\, I_n)}\,\mathrm{d}\theta\;, \tag{54}$$

where $e_\theta = (\cos\theta, \sin\theta)$. On the other hand, starting from (50), we observe that

$$-\Bigl(\frac{x}{t}\cdot e_\theta\, I_n - H(\theta)\Bigr)^{-2} \,=\, \frac{\partial}{\partial s}\Bigl((s + \frac{x}{t}\cdot e_\theta)I_n - H(\theta)\Bigr)^{-1}\Big|_{s=0}\;.$$

By Cramer's rule, the inverse of the matrix $SI_n - H(\theta) = SI_n - A_1\cos\theta - A_2\sin\theta$, with $S = s + \frac{x}{t}\cdot e_\theta$, has the following form

$$\Bigl(SI_n - H(\theta)\Bigr)^{-1} \,=\, \frac{1}{\bar\Delta(\theta,S)}\Bigl(S\,P_0(S,\cos\theta,\sin\theta) + \cos\theta\,P_1(S,\cos\theta,\sin\theta) + \sin\theta\,P_2(S,\cos\theta,\sin\theta)\Bigr)\;,$$

where $\bar\Delta(\theta, S) = \det(H(\theta) - SI_n)$ and P_0, P_1, P_2 are matrix-valued homogeneous polynomials of degree $n - 2$. The idea is now to insert this expansion into the right-hand side of (50) and to use (54) to express the result as a derivative of order $n - 1$ of the numerical density $\mathcal{F}_A$.

We begin with the term involving P_1, and remark that

$$\cos\theta\,\frac{\partial}{\partial s}\left(\frac{P_1(s + \frac{x}{t}\cdot e_\theta, \cos\theta, \sin\theta)}{\bar\Delta(\theta, s + \frac{x}{t}\cdot e_\theta)}\right)\Big|_{s=0} \,=\, t\,\frac{\partial}{\partial x_1}\left(\frac{P_1(\frac{x}{t}\cdot e_\theta, \cos\theta, \sin\theta)}{\bar\Delta(\theta, \frac{x}{t}\cdot e_\theta)}\right)\;.$$

The corresponding contribution to (50) is thus

$$E_1(t,x) \;=\; \frac{1}{4\pi^2 t}\,\frac{\partial}{\partial x_1}\,\text{f.p.}\!\int_{S^1} \frac{P_1(\frac{x}{t}\cdot e_\theta, \cos\theta, \sin\theta)}{\bar{\Delta}(\theta, \frac{x}{t}\cdot e_\theta)}\,\mathrm{d}\theta \;=\; -\frac{1}{(n{-}1)!}\,\frac{\partial}{\partial x_1} P_1(-\partial_t, \partial_{x_1}, \partial_{x_2})\mathcal{F}_A(t,x)\;.$$

Similarly, the term involving P_2 gives the contribution

$$E_2(t,x) \;=\; -\frac{1}{(n{-}1)!}\,\frac{\partial}{\partial x_2} P_2(-\partial_t, \partial_{x_1}, \partial_{x_2})\mathcal{F}_A(t,x)\;.$$

Finally, to treat the expression containing P_0, we observe that

$$\frac{\partial}{\partial s}\left(\frac{(s+\frac{x}{t}\cdot e_\theta)P_0(s+\frac{x}{t}\cdot e_\theta, \cos\theta, \sin\theta)}{\bar{\Delta}(\theta, s+\frac{x}{t}\cdot e_\theta)}\right)\Big|_{s=0} \;=\; (1+x\cdot\nabla_x)\left(\frac{P_0(\frac{x}{t}\cdot e_\theta, \cos\theta, \sin\theta)}{\bar{\Delta}(\theta, \frac{x}{t}\cdot e_\theta)}\right)\;,$$

and using (53), (54) we obtain the contribution

$$E_0(t,x) \;=\; -\frac{1}{(n{-}1)!\,t}\,(1+x\cdot\nabla_x)P_0(-\partial_t, \partial_{x_1}, \partial_{x_2})\mathcal{F}_A(t,x) \;=\; \frac{1}{(n{-}1)!}\,\frac{\partial}{\partial t} P_0(-\partial_t, \partial_{x_1}, \partial_{x_2})\mathcal{F}_A(t,x)\;.$$

Recalling that $E(t,x) = E_0(t,x) + E_1(t,x) + E_2(t,x)$, we arrive at (52). ∎

7 Three-dimensional examples

To illustrate the results of the previous sections, we consider here four concrete examples which, according to [19, Section I.7], give a complete picture of what can happen for three-dimensional matrices. The two-dimensional case, which is much simpler, was already treated in Section 4.1, and the references [3, 4, 16] include a detailed study of the singular set Σ_A for a few higher-dimensional examples.

Example 1. We first consider the 3×3 matrix

$$A \;=\; \begin{pmatrix} -1.5 & 1 & 0 \\ -1 & 1 & 1 \\ 0 & -1 & 0.5 \end{pmatrix}\;, \tag{55}$$

which is generic in the sense that hypotheses H1, H2 in Section 6.2 are fulfilled. The eigenvalues of the associated Hermitian matrix $H(\theta)$ satisfy $\lambda_1(\theta) < \lambda_2(\theta) < \lambda_3(\theta)$ for all $\theta \in [0,\pi]$, as can be seen from Fig. 3 (right). The critical set $\Sigma_A = C_A = C_1 \cup C_2$ is an algebraic curve of degree 6 consisting of a smooth ovate curve C_1 enclosing a cuspidal triangle C_2 (Fig. 3, left). The component C_1 corresponds to the eigenvalues $\lambda_1(\theta), \lambda_3(\theta)$ while C_2 is associated with $\lambda_2(\theta)$. All multiplicities are equal to one, and the number $N(z)$ defined by (39) is equal to 3 outside C_1 and inside C_2, and to 1 in the intermediate region. The numerical density f_A is continuous, identically constant inside C_2, and vanishes outside C_1. Moreover f_A is Hölder continuous with exponent 1/2 across C_1 and C_2, except at the cusps. The level lines of f_A are represented in Fig. 1.

Example 2. We next consider a nongeneric matrix

$$A \;=\; \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}\;, \tag{56}$$

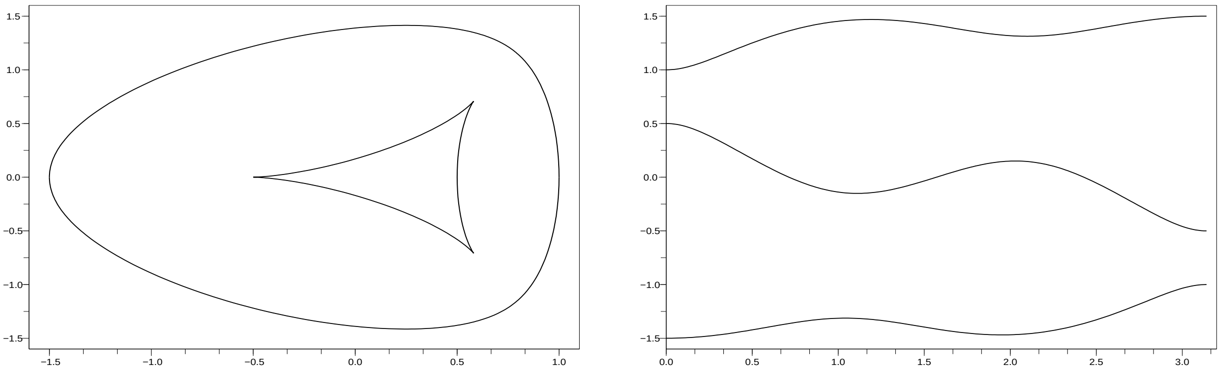

Figure 3: The critical set Σ_A (left) and the eigenvalues $\lambda_j(\theta)$ (right) are represented for the matrix (55).

for which the critical set Σ_A can be computed exactly. Indeed, if $H(\theta)$ is the Hermitian matrix (2), it is easy to verify that $\det(\lambda I_3 - H(\theta)) = \lambda^3 - \frac{3}{4}\lambda - \frac{1}{4}\cos\theta$, hence

$$\lambda_1(\theta) \,=\, \cos\Bigl(\frac{\theta + 2\pi}{3}\Bigr)\,, \quad \lambda_2(\theta) \,=\, \cos\Bigl(\frac{\theta - 2\pi}{3}\Bigr)\,, \quad \lambda_3(\theta) \,=\, \cos\Bigl(\frac{\theta}{3}\Bigr)\,.$$

Thus the permutation (35) is just a cycle $\tau = (1\,2\,3)$, and applying (37) we easily find that the critical curve C_A is the cardioid defined by $C_A = \{\frac{1}{3}(2e^{i\phi} + e^{2i\phi}) \,|\, \phi \in S^1\}$. Since $\lambda_1(0) = \lambda_2(0)$, the bitangent set C_A' is not empty and consists of the line segment joining the points $-1/2 \pm i/(2\sqrt{3})$, see Fig. 4 (left). Altogether we have $\Sigma_A = C_A \cup C_A'$, and we observe that Σ_A encloses a convex region of the complex plane which is of course the numerical range $W(A)$. The index $N(z)$ defined by (39) is equal to 3 outside C_A and to 1 inside. The numerical density f_A vanishes outside $W(A)$ and is equal to a nonzero constant inside the cuspidal region, in agreement with Propositions 2.6 and 6.1. In particular, f_A is discontinuous along the line segment C_A'.

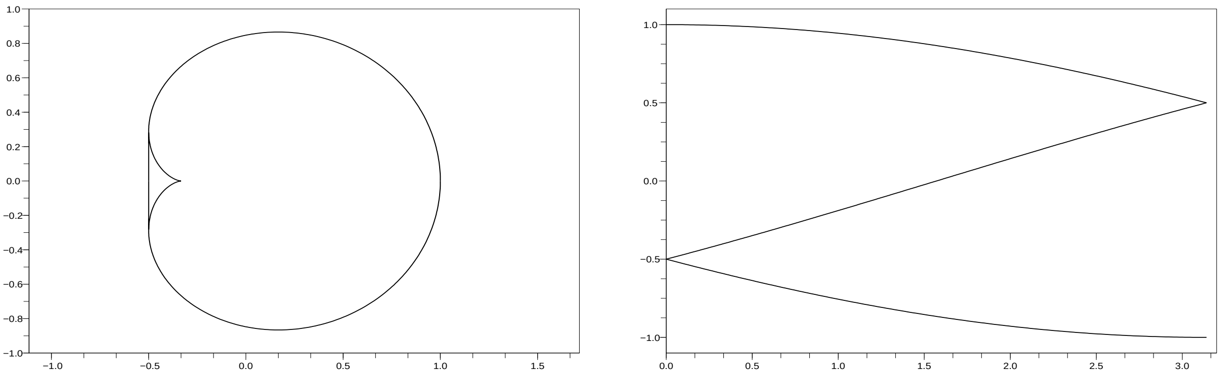

Figure 4: The critical set Σ_A (left) and the eigenvalues $\lambda_j(\theta)$ (right) are represented for the matrix (56).

Example 3. The matrices considered so far were unitarily irreducible. In contrast, the matrix

$$A \,=\, \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & a \end{pmatrix}\,, \qquad \text{where } a \in \mathbb{C}\,, \tag{57}$$

is the direct orthogonal sum of the two-dimensional Jordan block (31) and the one-dimensional matrix (a). The numerical density f_A can therefore be computed using Proposition 2.7, and without loss of generality we can assume that $a \geq 0$. However, we have to distinguish between three cases:

i) If $0 \leq a < 1$, the numerical range $W(A)$ is the closed unit disk, and the numerical density has the following expression:

$$f(z) \;=\; \frac{1}{\pi\sqrt{1-a^2}}\,\mathrm{argch}\left(\frac{1-az_1}{\sqrt{(1-az_1)^2-(1-|z|^2)(1-a^2)}}\right)\;, \qquad |z|<1\;,$$

which reduces to (34) when $a = 0$. In particular f_A vanishes on the unit circle, has a logarithmic singularity at the point $\{a\}$, and is otherwise smooth. The singular set Σ_A is the union of the unit circle and the point $\{a\}$.

ii) In the limiting case $a = 1$, the numerical range is still the closed unit disk, but the formula

$$f(z) \;=\; \frac{1}{\pi}\frac{\sqrt{1-|z|^2}}{1-z_1}\;, \qquad |z|<1\;,$$

shows that the numerical density has now an algebraic singularity at the boundary point $z = 1$.

iii) When $a > 1$, the numerical range $W(A)$ is the convex hull of the union of the unit disk and the exterior point $\{a\}$. Within this region, the numerical density satisfies

$$f(z) \;=\; \frac{1}{\pi\sqrt{a^2-1}}\;\arccos\left(\frac{1-az_1}{\sqrt{(1-az_1)^2+(1-|z|^2)(a^2-1)}}\right)\;, \qquad \text{when} \quad |z|<1\;,$$

and $f(z) = 1/\sqrt{a^2-1}$ when $|z| > 1$. As is easily verified, the eigenvalues of the Hermitian matrix $H(\theta)$ are $\lambda_1(\theta) = -1$, $\lambda_2(\theta) = 1$, and $\lambda_3(\theta) = a\cos\theta$, see Fig. 5 (right). The algebraic curve C_A consists of the unit circle (associated with λ_1, λ_2) and the point $\{a\}$ (corresponding to λ_3), but the bitangent set C_A' is not empty and consists of two line segments, see Fig. 5 (left).

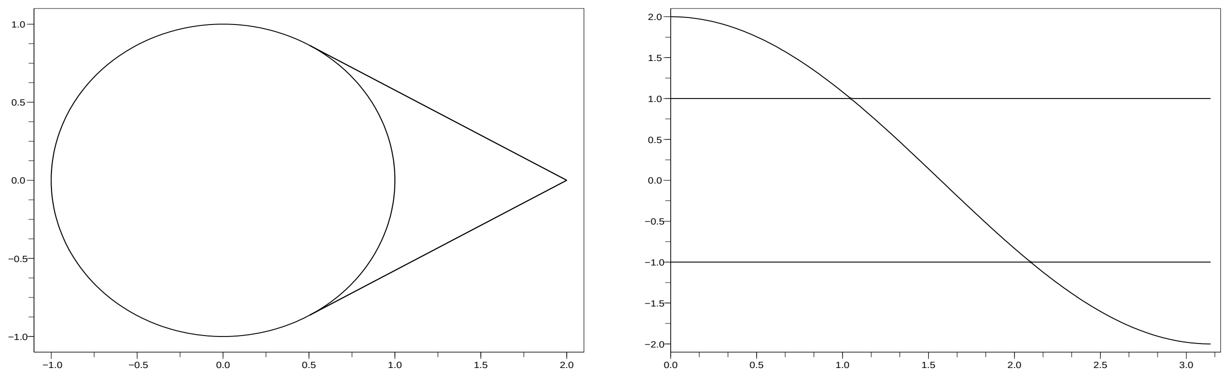

Figure 5: The critical set Σ_A (left) and the eigenvalues $\lambda_j(\theta)$ (right) are represented for the matrix (57) with $a = 2$.

Example 4. As a final example, we consider the case of a normal matrix $A \in M_3(\mathbb{C})$ whose eigenvalues $\lambda_1, \lambda_2, \lambda_3$ are not colinear. Then the numerical range is the triangle with vertices $\{\lambda_1, \lambda_2, \lambda_3\}$ and the numerical density is a multiple of the characteristic function of $W(A)$. In that situation $\Sigma_A = C_A \cup C_A'$, where C_A is the set of all vertices and C_A' the set of all edges of the triangle.

8 Statistical properties of the numerical measure

In this section, we study the numerical measure from a statistical point of view, and we establish various convergence results which show that the measure μ_A of a large matrix $A \in \mathbf{M}_n(\mathbb{C})$ is concentrated in a neighborhood of the barycenter of the spectrum $\sigma(A)$.

8.1 Concentration phenomena

Proposition 8.1 *For any $A \in \mathbf{M}_n(\mathbb{C})$, the first moment of the probability measure μ_A is the normalized trace of A:*

$$\overline{\mu}_A \;=\; \int_{\mathbb{C}} z \,\mathrm{d}\mu_A(z) \;=\; \frac{1}{n}\,\mathrm{Tr}(A)\;,$$

and the variance of μ_A is given by

$$\mathrm{Var}(\mu_A) \;=\; \int_{\mathbb{C}} |z-\overline{\mu}_A|^2 \,\mathrm{d}\mu_A(z) \;=\; \frac{1}{n+1}\Bigl(\frac{1}{n}\,\mathrm{Tr}(A^*A) - \Bigl|\frac{1}{n}\,\mathrm{Tr}(A)\Bigr|^2\Bigr)\;. \tag{58}$$

Proof. Let a_{jk} denote the entries of A. Applying definition (1) with $\phi(z) = z$, we have to compute the average of $\sum_{jk} a_{jk}\overline{x_j}x_k$ over the unit sphere. By symmetry, the average of $\overline{x_j}x_k$ is equal to zero if $j \neq k$ and to $1/n$ if $j = k$. Thus

$$\overline{\mu}_A \;=\; \frac{1}{n}(a_{11} + \cdots + a_{nn}) \;=\; \frac{1}{n}\,\mathrm{Tr}(A)\;.$$

To compute the second moment, we take $\phi(z) = |z|^2$ and proceed in exactly the same way. By symmetry, the average of $\overline{x_j}x_k\overline{x_\ell}x_m$ is zero unless $j = k$ and $\ell = m$, or $j = m$ and $\ell = k$. Moreover, it is easy to verify that the average of $|x_j|^2|x_\ell|^2$ is equal to r_n if $j = \ell$ and to s_n if $j \neq \ell$, where

$$r_n \;=\; \int_{\partial\mathbb{B}^n} |x_1|^4 \,\mathrm{d}\bar{\sigma}(x) \;=\; \frac{2}{n(n+1)}\;, \qquad s_n \;=\; \int_{\partial\mathbb{B}^n} |x_1|^2|x_2|^2 \,\mathrm{d}\bar{\sigma}(x) \;=\; \frac{1}{n(n+1)}\;.$$

Thus

$$\begin{aligned}
\int_{\mathbb{C}} |z|^2 \,\mathrm{d}\mu_A(z) \;&=\; \int_{\partial\mathbb{B}^n} |\langle Ax, x\rangle|^2 \,\mathrm{d}\bar{\sigma}(x) \;=\; \sum_{j,k,\ell,m} \overline{a_{kj}}a_{\ell m}\int_{\partial\mathbb{B}^n} \overline{x_j}x_k\overline{x_\ell}x_m \,\mathrm{d}\bar{\sigma}(x)\\
&=\; s_n\sum_{j\neq\ell}(\overline{a_{jj}}a_{\ell\ell} + |a_{\ell j}|^2) + r_n\sum_{j=1}^{n}|a_{jj}|^2\\
&=\; s_n\sum_{j,\ell=1}^{n}(\overline{a_{jj}}a_{\ell\ell} + |a_{\ell j}|^2) \;=\; s_n(|\,\mathrm{Tr}\,A|^2 + \mathrm{Tr}(A^*A))\;.
\end{aligned}$$

This gives the desired result, since $\mathrm{Var}(\mu_A) = \int_{\mathbb{C}} |z|^2 \,\mathrm{d}\mu_A(z) - |\overline{\mu}_A|^2$. ∎

Now we consider a sequence of matrices $\{A_n\}_{n\geq 1}$ such that $A_n \in \mathbf{M}_n(\mathbb{C})$ for each $n \geq 1$. As is well known, we have

$$\frac{1}{n}\,\mathrm{Tr}(A_n^*A_n) \;\leq\; \|A_n\|^2 \;\leq\; \mathrm{Tr}(A_n^*A_n)\;,$$

where $\|A_n\| = \sup\{\|A_nx\| \,|\, x \in \partial\mathbb{B}^n\}$. As a consequence, if we suppose that $\|A_n\|^2 = o(n)$ as $n \to \infty$, it follows from (58) that the variance of μ_{A_n} converges to zero as $n \to \infty$. This gives:

Corollary 8.2 *Assume that $A_n \in \mathbf{M}_n(\mathbb{C})$ and that $\|A_n\|^2/n \to 0$ as $n \to \infty$. Then the measure $\mu_{A_n} - \delta_{\overline{\mu}_{A_n}}$ converges weakly to zero as $n \to \infty$.*

We recall that the *numerical radius* of a matrix $A \in \mathbf{M}_n(\mathbb{C})$ is defined by

$$R(A) = \sup\{|z| \,|\, z \in W(A)\} = \sup\{|\langle Ax, x\rangle| \,|\, x \in \partial\mathbb{B}^n\} ,$$

and satisfies $R(A) \le \|A\| \le 2R(A)$ [15]. Thus, a sequence of matrices $A_n \in \mathbf{M}_n(\mathbb{C})$ is uniformly bounded (in the operator norm) if and only if the numerical ranges $W(A_n)$ are all contained in a bounded region of the complex plane. Under this assumption, Proposition 8.1 shows that the variance of μ_{A_n} is $\mathcal{O}(1/n)$ as $n \to \infty$, so that the numerical measure is asymptotically concentrated in a disk of radius $\mathcal{O}(1/\sqrt{n})$ around the mean $\overline{\mu}_{A_n}$.

In the introduction, we have observed that the numerical measure μ_{A_n} is the distribution of the random variable $\langle A_n X_n, X_n\rangle \in \mathbb{C}$ when the vector X_n is uniformly distributed on the unit sphere $\partial\mathbb{B}^n$. With this interpretation, Corollary 8.2 is reminiscent of the weak law of large numbers in probability theory. Under slightly stronger assumptions, it is also possible to obtain a pointwise convergence result in the spirit of the strong law of large numbers. Without loss of generality, we assume from now on that $\mathrm{Tr}(A_n) = 0$ for all $n \ge 1$, so that the measure μ_{A_n} is centered at the origin.

Proposition 8.3 *Assume that $A_n \in \mathbf{M}_n(\mathbb{C})$ satisfies $\mathrm{Tr}(A_n) = 0$ for all $n \ge 1$ and*

$$\sup_{n\ge 1} \frac{(\log n)\,\|A_n\|}{n^{1/2}} < \infty . \tag{59}$$

If for each $n \ge 1$ the random variable X_n is uniformy distributed on the unit sphere $\partial\mathbb{B}^n$, then $\langle A_n X_n, X_n\rangle$ converges almost surely to zero as $n \to \infty$.

Proof. It is clearly sufficient to prove the result for Hermitian matrices A_n, because the general case then follows by considering the real and imaginary parts of $\langle A_n X_n, X_n\rangle$. We thus assume that $A_n = A_n^*$ for all $n \ge 1$, and we denote by $\lambda_{n,1}, \dots, \lambda_{n,n}$ the eigenvalues of A_n. For each $n \ge 1$, let $Y_{n,1}, \dots, Y_{n,n}$ be independent and identically distributed complex random variables with density function $f_Y(z) = \pi^{-1}e^{-|z|^2}$, $z \in \mathbb{C}$. In particular, we have $E(|Y_{n,m}|^{2k}) = k!$ for each $k \in \mathbb{N}$. Since the Euclidean measure on $\partial\mathbb{B}^n$ is the projection on the unit sphere of the standard Gaussian measure in $\mathbb{C}^n$, we obtain a uniformly distributed random variable on $\partial\mathbb{B}^n$ by setting $X_n = U_n Y_n/\|Y_n\|$, where $Y_n = (Y_{n,1}, \dots, Y_{n,n})^\top$ and $U_n \in \mathbf{U}_n(\mathbb{C})$ is a unitary matrix such that $U_n^* A_n U_n = \mathrm{diag}(\lambda_{n,1}, \dots, \lambda_{n,n})$. Thus

$$\langle A_n X_n, X_n\rangle = \frac{\lambda_{n,1}|Y_{n,1}|^2 + \cdots + \lambda_{n,n}|Y_{n,n}|^2}{|Y_{n,1}|^2 + \cdots + |Y_{n,n}|^2} = \frac{P_n}{Q_n} ,$$

where

$$P_n = \frac{1}{n}\sum_{m=1}^n \lambda_{n,m}|Y_{n,m}|^2 = \frac{1}{n}\sum_{m=1}^n \lambda_{n,m}(|Y_{n,m}|^2 - 1) , \qquad Q_n = \frac{1}{n}\sum_{m=1}^n |Y_{n,m}|^2 .$$

By the strong law of large numbers, the denominator Q_n converges almost surely to 1 as $n \to \infty$, hence it remains to show that the numerator P_n converges almost surely to zero. But this follows from classical theorems on the limiting behavior of weighted sums of independent random variables, see [5, 23]. Since for each $n \ge 1$ the random variables $X_{n,m} = |Y_{n,m}|^2 - 1$ $(1 \le m \le n)$ are independent, have zero mean and finite second order moment, and since by (59) the coefficients $a_{n,m} = n^{-1}\lambda_{n,m}$ satisfy

$$\max_{1\le m\le n} |a_{n,m}| \le \frac{C}{n^{1/2}\log n} ,$$

the results of [23, Section 3] imply that $P_n = \sum_{m=1}^n a_{n,m}X_{n,m}$ converges almost surely to zero as $n \to \infty$. ■

8.2 Central limit theorems

The results established so far show that for a sequence of traceless matrices $A_n \in \mathbf{M}_n(\mathbb{C})$ the numerical measure μ_{A_n} tends to concentrate on the origin as $n \to \infty$. Under stronger assumptions, we now prove that the rescaled measure $\mu_{\sqrt{n}A_n}$ converges to a Gaussian distribution, as in the classical central limit theorem. We first consider the Hermitian case, which is somewhat simpler.

Proposition 8.4 *Let $\{A_n\}_{n\ge1}$ be a sequence of Hermitian matrices such that $A_n \in \mathbf{M}_n(\mathbb{C})$ and* $\mathrm{Tr}(A_n) = 0$. *We assume that*

$$\frac{(\log n)\,\|A_n\|}{n^{1/2}} \xrightarrow[n\to\infty]{} 0 , \qquad \textit{and} \quad \frac{1}{n}\,\mathrm{Tr}(A_n^2) \xrightarrow[n\to\infty]{} \sigma^2 > 0 . \tag{60}$$

Then the rescaled numerical measure $\mu_{\sqrt{n}A_n}$ converges weakly to the normal distribution $\mathcal{N}(0,\sigma^2)$ as $n \to \infty$.

Proof. We use the same notations as in the proof of Proposition 8.3. For each $n \ge 1$, the numerical measure μ_{A_n} is the distribution of the random variable

$$Z_n = \frac{\lambda_{n,1}|Y_{n,1}|^2 + \cdots + \lambda_{n,n}|Y_{n,n}|^2}{|Y_{n,1}|^2 + \cdots + |Y_{n,n}|^2} = \frac{P_n}{Q_n} ,$$

where $\lambda_{n,1}, \dots, \lambda_{n,n}$ denote the eigenvalues of A_n and $Y_{n,1}, \dots, Y_{n,n}$ are independent complex random variables with density function $f_Y(z) = \pi^{-1}e^{-|z|^2}$. Since Q_n converges almost surely to 1 as $n \to \infty$, we have to show that $n^{1/2}P_n$ converges in law to $\mathcal{N}(0,\sigma^2)$.

To do that, we use the Lindeberg-Feller theorem for triangular arrays of random variables [8, Section 2.4.b]. Let

$$X_{n,m} = \frac{1}{\sqrt{n}}\,\lambda_{n,m}\,(|Y_{n,m}|^2 - 1) , \qquad 1 \le m \le n ,$$

so that $n^{1/2}P_n = X_{n,1} + \cdots + X_{n,n}$. For each fixed $n \ge 1$, the random variables $X_{n,m}$ are independent and satisfy $E(X_{n,m}) = 0$ for $m = 1, \dots, n$. Moreover,

$$\sum_{m=1}^n E(|X_{n,m}|^2) = \frac{1}{n}\sum_{m=1}^n \lambda_{n,m}^2 = \frac{1}{n}\,\mathrm{Tr}(A_n^2) \xrightarrow[n\to\infty]{} \sigma^2 > 0 .$$

Finally, for any $\epsilon > 0$, we have

$$E(|X_{n,m}|^2\,;\,|X_{n,m}| \ge \epsilon) = \frac{\lambda_{n,m}^2}{\pi n}\int_{D_{n,m,\epsilon}} (|z|^2 - 1)^2\,e^{-|z|^2}\,\mathrm{d}z ,$$

where $D_{n,m,\epsilon} = \{z \in \mathbb{C} \,|\, \lambda_{n,m}^2(|z|^2-1)^2 \geq n\epsilon^2\}$. Thus, using the first assumption in (60), we obtain by a direct calculation

$$\sum_{m=1}^{n} E(|X_{n,m}|^2\,;\,|X_{n,m}|\geq\epsilon) \;\leq\; n \sup_{1\leq m\leq n} E(|X_{n,m}|^2\,;\,|X_{n,m}|\geq\epsilon) \xrightarrow[n\to\infty]{} 0\ .$$

Invoking the Lindeberg-Feller theorem, we conclude that $n^{1/2}P_n = X_{n,1}+\cdots+X_{n,n}$ converges in law to $\mathcal{N}(0,\sigma^2)$, which is the desired result. ■

Remark. Since the numerical density of a Hermitian matrix is a B-spline, Proposition 8.4 shows under very general assumptions that B-splines of degree n satisfy a central limit theorem in the limit $n\to\infty$. In the particular case of uniform B-splines, this result was obtained by Unser *et. al.* in [25].

Before considering more general matrices, we would like to mention an alternative proof of Proposition 8.4 which has its own interest. The starting point is a very nice formula for the moments of the numerical measure of a Hermitian matrix $A \in \mathbf{M}_n(\mathbb{C})$. Fix $k\in\mathbb{N}$ and let $\lambda_1,\ldots,\lambda_n\in\mathbb{R}$ denote the eigenvalues of A. Using (17) or (18) with $\phi(x)=x^k$, we find

$$\int_{\mathbb{R}} x^k\,\mathrm{d}\mu_A(x) \;=\; (n-1)!\int_{D_{n-1}} (t_1\lambda_1+\cdots+t_n\lambda_n)^k\,\mathrm{d}t_1\ldots\,\mathrm{d}t_{n-1}\ ,$$

where D_{n-1} is the $(n-1)$-dimensional simplex defined in (12) and $t_n = 1-(t_1+\cdots+t_{n-1})$. To evaluate the right-hand side, we apply the multinomial formula

$$(X_1+\cdots+X_n)^k \;=\; \sum_{|\alpha|=k}\frac{k!}{\alpha!}\,X^\alpha\ ,$$

where the sum runs over all multi-indices $\alpha\in\mathbb{N}^n$ of order $|\alpha|=\alpha_1+\cdots+\alpha_n=k$. Here we use the standard notations $X^\alpha = X_1^{\alpha_1}\cdots X_n^{\alpha_n}$ and $\alpha! = (\alpha_1!)\cdots(\alpha_n!)$. Now, it is not difficult to verify that

$$\int_{D_{n-1}} t^\alpha\,\mathrm{d}t_1\ldots\,\mathrm{d}t_{n-1} \;=\; \frac{\alpha!}{(n+k-1)!}\ ,$$

for any $\alpha\in\mathbb{N}^n$ with $|\alpha|=k$. We thus obtain the following identity

$$\int_{\mathbb{R}} x^k\,\mathrm{d}\mu_A(x) \;=\; \frac{k!\,(n-1)!}{(n+k-1)!}\sum_{|\alpha|=k}\lambda^\alpha\ , \tag{61}$$

which shows that the k-th moment of the numerical measure μ_A is the complete symmetric homogeneous polynomial of degree k in the variables $\lambda_1,\ldots,\lambda_n$, divided by the combinatorial factor $\binom{n+k-1}{k}$ which is just the number of terms in the sum.

Using the Newton identities, the right-hand side of (61) can be decomposed into as a sum of products of elementary symmetric polynomials of the form $p_\ell = \lambda_1^\ell+\cdots+\lambda_n^\ell$, see [21, Eq. (2.14')]. If we assume that $\mathrm{Tr}(A)=0$, then $p_1=0$ and the number of nonzero terms in the sum is considerably reduced. Using these remarks, it is not difficult to show that, under the assumptions of Proposition 8.3, the k-th moment of the rescaled numerical measure $\mu_{\sqrt{n}A_n}$ satisfies

$$\int_{\mathbb{R}} x^k\,\mathrm{d}\mu_{\sqrt{n}A_n}(x) \xrightarrow[n\to\infty]{} \begin{cases} 0 & \text{if } k \text{ is odd}\ ,\\ 2^{-k/2}\sigma^k\frac{k!}{(k/2)!} & \text{if } k \text{ is even}\ .\end{cases}$$

Since the moments in the right-hand side are those of the normal law, we conclude that $\mu_{\sqrt{n}A_n}$ converges weakly to $\mathcal{N}(0,\sigma^2)$ as $n \to \infty$.

We now consider general matrices $A_n \in \mathbf{M}_n(\mathbb{C})$, and obtain a central limit theorem by applying Proposition 8.4 to the Radon transform of A_n.

Theorem 8.5 *Let $\{A_n\}_{n\ge1}$ be a sequence of matrices satisfying $A_n \in \mathbf{M}_n(\mathbb{C})$, $\mathrm{Tr}(A_n) = 0$, and $(\log n)\|A_n\|/n^{1/2} \to 0$ as $n \to \infty$. We assume that*

$$\frac{1}{n}\,\mathrm{Tr}(A_n^*A_n) \xrightarrow[n\to\infty]{} a > 0\ , \qquad \frac{1}{n}\,\mathrm{Tr}(A_n^2) \xrightarrow[n\to\infty]{} b \in [0,a)\ . \tag{62}$$

Then the rescaled numerical measure $\mu_{\sqrt{n}A_n}$ converges weakly to the Gaussian measure $f_\infty(z)\,\mathrm{d}z$ as $n \to \infty$, where

$$f_\infty(x+iy) \;=\; \frac{1}{\pi\sqrt{a^2-b^2}}\,e^{-\frac{x^2}{a+b}-\frac{y^2}{a-b}}\ . \tag{63}$$

Proof. For each $\theta \in S^1$, the Hermitian matrices $H_n(\theta) = \frac{1}{2}(e^{-i\theta}A_n + e^{i\theta}A_n^*)$ satisfy the assumptions of Proposition 8.4, with

$$\sigma^2 \;=\; \sigma(\theta)^2 \;=\; \lim_{n\to\infty}\frac{1}{2n}\Big(\mathrm{Tr}(A_n^*A_n) + \mathrm{Re}(\mathrm{Tr}(A_n^2)e^{-2i\theta})\Big) \;=\; \frac{1}{2}(a + b\cos(2\theta))\ .$$

Let us denote by μ_n and $\mu_{n,\theta}$ the numerical measures of $\sqrt{n}\,A_n$ and $\sqrt{n}\,H_n(\theta)$, respectively. Since $\mu_{n,\theta}$ is the Radon transform of μ_n, the two-dimensional Fourier transform $\hat{\mu}_n(\xi)$ for $\xi = re^{i\theta}$ is precisely the one-dimensional Fourier transform of $\mu_{n,\theta}$ evaluated at r [12]. Thus, applying Proposition 8.4, we find

$$\hat{\mu}_n(re^{i\theta}) \;=\; \int_{\mathbb{R}} e^{-ixr}\,\mathrm{d}\mu_{n,\theta}(x) \xrightarrow[n\to\infty]{} \frac{1}{\sigma(\theta)\sqrt{2\pi}}\int_{\mathbb{R}} e^{-ixr}e^{-x^2/(2\sigma(\theta)^2)}\,\mathrm{d}x \;=\; e^{-\sigma(\theta)^2r^2/2}\ ,$$

for any $r \ge 0$, $\theta \in S^1$. This shows that μ_n converges weakly as $n \to \infty$ to the measure μ_∞ on $\mathbb{C}$ defined by

$$\hat{\mu}_\infty(\xi) \;=\; e^{-\frac{1}{2}|\xi|^2\sigma(\theta)^2} \;=\; e^{-\frac{a+b}{4}\,\mathrm{Re}(\xi)^2}e^{-\frac{a-b}{4}\,\mathrm{Im}(\xi)^2}\ , \quad \xi \in \mathbb{C}\ .$$

Inverting the Fourier transform, this gives $\mathrm{d}\mu_\infty = f_\infty(z)\,\mathrm{d}z$ with f_∞ as in (63). ■

Remarks.

1. If we assume for simplicity that $\|A_n\|$ is uniformly bounded, we can suppose (up to extracting a subsequence) that $\frac{1}{n}\,\mathrm{Tr}(A_n^*A_n)$ converges as $n \to \infty$ to some $a \ge 0$. However, we have to assume in (62) that $a > 0$ in order to get a universal Gaussian limit.

2. Similarly, the first assumption in (62) implies that $\frac{1}{n}\,\mathrm{Tr}(A_n^2)$ converges, after extracting a subsequence, to some $b \in \mathbb{C}$ with $|b| \le a$. Multiplying A_n by a unit complex number, we can assume that $0 \le b \le a$, but in the borderline case where $b = a$ the limiting measure $f_\infty(z)\,\mathrm{d}z$ should be replaced $(2\pi a)^{-1/2}e^{-x^2/(2a)}\,\mathrm{d}x \otimes \delta_{y=0}$.

Example. Let us consider the Jordan block of size n:

$$A_n \;=\; \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix}\ .$$

Since $\mathrm{Tr}(A_n^* A_n) = n-1$ and $\mathrm{Tr}(A_n^2) = 0$, the assumptions of Theorem 8.5 are satisfied with $a = 1$ and $b = 0$. Thus the numerical measure of $\sqrt{n}A_n$ converges weakly to the normal distribution $\pi^{-1}e^{-|z|^2}\,\mathrm{d}z$ as $n \to \infty$. In this example, the measure μ_{A_n} has in fact a radially symmetric density for all $n \geq 1$, see Section 4.2.

9 Perspectives

As a conclusion, we briefly mention a natural extension of our work, which is left for future investigation. Recall that a homogeneous polynomial $P \in \mathbb{R}[X_0, X_1, \ldots, X_d]$ of total degree n is *hyperbolic* in the direction $e_0 = (1, 0, \ldots, 0)$ if, on the one hand, it has partial degree n with respect to the first variable X_0, and on the other hand the n roots of the univariate polynomial $t \mapsto P(t, y)$ are real for every vector $y \in \mathbb{R}^d$. Hyperbolic polynomials arise as principal symbols of hyperbolic differential operators of order n in d space variables, see [11]. As an example, if $A_1, \ldots, A_d \in \mathbf{M}_n(\mathbb{C})$ are Hermitian matrices, the polynomial

$$P(X_0, X_1, \ldots, X_d) \,=\, \det(X_0 I_n - X_1 A_1 - \cdots - X_d A_d) \tag{64}$$

is hyperbolic. In the particular case where $d = 2$, it has been conjectured in [20], and proved in [14], that all monic hyperbolic polynomials are of the form (64). This is no longer true if $d \geq 3$.

It might be argued that a large part of our work is not really about matrices, but rather concerns hyperbolic polynomials in $2 + 1$ variables. Indeed, if $A \in \mathbf{M}_n(\mathbb{C})$ and A_1, A_2 are as in (2), the eigenvalues $\lambda_1(\theta), \ldots, \lambda_n(\theta)$ of the Hermitian matrix $H(\theta)$ are the solutions of the equation

$$P_A(\lambda, \cos\theta, \sin\theta) \,=\, 0\ , \qquad \theta \in S^1\ ,$$

where $P_A(X_0, X_1, X_2) = \det(X_0 I_n - X_1 A_1 - X_2 A_2)$ is the hyperbolic polynomial associated with A_1, A_2. As was shown in Section 4, the numerical measure μ_A is entirely determined by the eigenvalues $\lambda_j(\theta)$, hence by the polynomial P_A.

This in turn suggests a natural way to associate to any hyperbolic polynomial P of degree n in $d+1$ variables a probability measure μ_P on $\mathbb{R}^d$, which coincides with the numerical measure μ_A when $d = 2$ and $P = P_A$. Given a unit vector $\omega \in S^{d-1} \subset \mathbb{R}^d$, let $\lambda_1(\omega), \ldots, \lambda_n(\omega) \in \mathbb{R}$ be the roots of the polynomial equation $P(\lambda, \omega) = 0$, and let $B_\omega(s) = B[\lambda_1(\omega), \ldots, \lambda_n(\omega)](s)$ be the normalized B-spline with knots $\lambda_1(\omega), \ldots, \lambda_n(\omega)$. The "numerical measure" of P is then the unique probability measure μ_P on $\mathbb{R}^d$ whose Radon transform satisfies

$$(\mathcal{R}\mu_P)(\omega,\, \mathrm{d}s) \,=\, B_\omega(s)\,\mathrm{d}s\ ,$$

where, by definition, $(\mathcal{R}\mu_P)(\omega, I) = \mu_P(\{x \in \mathbb{R}^d \,|\, x \cdot \omega \in I\})$ for any interval $I \subset \mathbb{R}$.

In this generalized setting, the counterpart of normal matrices is the case where the polynomial P split into linear factors

$$P(X) \,=\, \prod_{k=1}^{n} (X_0 - v^k \cdot (X_1, \ldots, X_d))\ , \qquad \text{with } v^k \in \mathbb{R}^d\ .$$

When $n \geq d+1$ and the vectors v^k span the affine space $\mathbb{R}^d$, the density of μ_P with respect to the Lebesgue measure is the multivariate B-spline in d variables, whose nodes are the v^k's. It is piecewise polynomial of degree $n - d - 1$. In the generic situation where any $(d+1)$-uplet of vectors v^k is an affine basis, it is of class C^{n-d-2}, see [6]. Again the density is log-concave in this case.

As in the two-dimensional case, the measure μ_P can be expressed in terms of B_ω using the back-projection method. The example above suggests that, as the space dimension increases, the measure μ_P becomes more singular. Then, the inversion formula has to be understood in the sense of distributions. In the three-dimensional case $d = 3$, we arrive at the simple expression

$$\mu_P = -\frac{1}{8\pi^2}\,\Delta_x \int_{S^2} B_\omega(x \cdot \omega)\,\mathrm{d}\sigma(\omega)\ , \tag{65}$$

where $\mathrm{d}\sigma$ denotes the Euclidean measure on the unit sphere S^2. As an example, in the particular situation where $P = X_0^2 - X_1^2 - X_2^2 - X_3^2$, which corresponds to the differential operator $\partial_t^2 - \Delta$ of the wave equation, we obtain $\mu_P = \frac{1}{4\pi}\,\mathrm{d}\sigma(x)$. This shows that, when $d \geq 3$, the support of μ_P does not need to be convex. Because the wave equation satisfies the Huyghens Principle, this example suggests that the link between polynomial regions of the density and lacunas of differential operators persists in higher dimensions.

References

[1] V. M. Atiyah. Hyperbolic differential equations and algebraic geometry (after Petrowsky), *Séminaire Bourbaki* **319** (1967), 1–12.

[2] M. Atiyah, R. Bott, and L. Gårding. Lacunas for hyperbolic differential operators with constant coefficients. I. *Acta Math.* **124** (1970), 109–189.

[3] M. Atiyah, R. Bott, and L. Gårding. Lacunas for hyperbolic differential operators with constant coefficients. II. *Acta Math.* **131** (1973), 145–206.

[4] J. Bazer and D. Yen. Lacunas of the Riemann matrix of symmetric-hyperbolic systems in two space variables. *Comm. Pure Appl. Math.* **22** (1969), 279–333.

[5] Y. S. Chow and T. L. Lai, Limiting behavior of weighted sums of independent random variables. *Ann. Probability* **1** (1973), 810–824.

[6] W. Dahmen. On multivariate B-splines. *SIAM J. Numer. Anal.*, **17** (1980), pp 179–191.

[7] C. de Boor. *A Practical Guide to Splines.* Applied Mathematical Sciences **27**, Springer, New York, 1978.

[8] R. Durrett. *Probability: theory and examples* (second edition). Duxbury Press, Belmont, 1996.

[9] M. Fiedler. Geometry of the numerical range of matrices. *Linear Algebra Appl.* **37** (1981), 81–96.

[10] R. J. Gardner. The Brunn–Minkoski inequality. *Bull. (new series) A. M. S.*, **39** (2002), pp 355–405.

[11] L. Gårding. Linear hyperbolic partial differential equations with constant coefficients. *Acta Math.*, **85** (1951), pp 1–62. An inequality for hyperbolic polynomials. *J. Math. Mech.*, **8** (1959), pp 957–965.

[12] S. Helgason. *The Radon transform*, 2nd edition. Progress in Mathematics. Birkhäuser, Basel, 1999.

[13] F. Hausdorff. Der Wertevorrat einer Bilinearform. *Math. Zeitschrift*, **3** (1919), pp 314–316.

[14] J. W. Helton, V. Vinnikov. Linear matrix inequality representation of sets. *Comm. Pure Appl. Math.*, **60** (2007), pp 654–674.

[15] R.A. Horn and C.R. Johnson. *Topics in matrix analysis*, Cambridge university press, Cambridge, 1991.

[16] E. A. Jonckheere, F. Ahmad, E. Gutkin. Differential topology of numerical range. *Linear Algebra Appl.*, **279** (1998), pp. 227–254

[17] M. Joswig and B. Straub. On the numerical range map. *J. Austral. Math. Soc. Ser. A* **65** (1998), 267–283.

[18] T. Kato. *Perturbation theory for linear operators*, Reprint of the 1980 edition. Springer, Berlin, 1995.

[19] R. Kippenhahn. Über den Wertevorrat einer Matrix. *Math. Nachrichten*, **6** (1951), pp 193–228.

[20] P. D. Lax. Differential equations, difference equations and matrix theory. *Comm. Pure Appl. Math.*, **11** (1958), pp 175–194.

[21] I. G. Macdonald. *Symmetric functions and Hall polynomials.* Second edition. Oxford Mathematical Monographs, Oxford University Press, 1995.

[22] I. Petrovsky, On the diffusion of waves and lacunas for hyperbolic equations, *Math. Sbornik* **17** (1945), 289–368.

[23] H. Teicher. Almost certain convergence in double arrays. *Z. Wahrsch. Verw. Gebiete* **69** (1985), 331–345.

[24] O. Toeplitz. Das algebraische Analogon zu einem Satze von Fejér. *Math. Zeitschrift*, **2** (1918), pp 187–197.

[25] M. Unser, A. Aldroubi, M. Eden, On the Asymptotic Convergence of B-Spline Wavelets to Gabor Functions. *IEEE Trans. Inform. Theo.*, **38** (1992), pp. 864–872.

[26] D. Yen and J. Bazer. The Riemann matrix of $(2+1)$-dimensional symmetric-hyperbolic systems. *Comm. Pure Appl. Math.* **20** (1967), 329–363.