

A direct proof of the complete Crouzeix inequality

OpenAI

September 2026

Abstract

Every square matrix A comes with a compact convex subset $W(A)$ of the complex plane, called its numerical range. The complete Crouzeix conjecture says: if you plug A into a matrix-valued polynomial P , the result is never larger than twice the biggest value of P on $W(A)$, and this is uniform in all matrix dimensions. We give a self-contained direct proof. It uses an exterior conformal map of an analytic convex domain containing $W(A)$, a positive kernel on the unit circle, and a two-by-two positive block comparison. Degenerate cases, where $W(A)$ is a single point or a line segment, are included.

1 Introduction

How to read this paper

Sections 2–6 prove the theorem. The one technical input is the exterior conformal map of Appendix A; a reader who wants to accept it may take that appendix as a black box.

The numerical range

Fix a positive integer n and a matrix $A \in M_n(\mathbb{C})$. A vector $u \in \mathbb{C}^n$ with $u^*u = 1$ is a *unit vector*. Its *Rayleigh quotient* is the complex number u^*Au . The *numerical range* of A is the set

$$W(A) = \{ u^*Au : u \in \mathbb{C}^n, u^*u = 1 \}. \quad (1.1)$$

The numerical range of A is always a compact convex subset of $\mathbb{C}$, and it always contains the spectrum of A : a unit eigenvector u of an eigenvalue λ satisfies $u^*Au = \lambda u^*u = \lambda$.

Two extremes: (a) if $A = \text{diag}(\lambda_1, \dots, \lambda_n)$ with all $\lambda_i \in \mathbb{R}$, then $W(A) = [\min \lambda_i, \max \lambda_i]$; (b) if $A = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}$, then $|u^*Au| = 2|u_1u_2| \leq |u_1|^2 + |u_2|^2 = 1$, so $W(A)$ is the closed unit disk $\overline{\mathbb{D}}$, even though the spectrum is just $\{0\}$.

For a normal matrix (one with $AA^* = A^*A$), the spectral theorem gives a unitary U with $A = U\Lambda U^*$, and then u^*Au is a convex combination of the eigenvalues. So $W(A)$ is exactly the convex hull of the eigenvalues. For a general matrix, this statement fails badly: the numerical range is genuinely larger than the spectrum, and that extra size is what makes the spectral-set constant 2 possible.

The main theorem

We allow *matrix-valued* polynomials. If $P(z) = \sum_{k=0}^d B_k z^k$ with coefficient matrices $B_k \in M_m(\mathbb{C})$, define

$$P[A] = \sum_{k=0}^d A^k \otimes B_k \in M_n(\mathbb{C}) \otimes M_m(\mathbb{C}) \cong M_{nm}(\mathbb{C}). \quad (1.2)$$

Here $\otimes$ is the Kronecker product: the factor A^k acts on $\mathbb{C}^n$ and the factor B_k acts on $\mathbb{C}^m$, and the two actions commute. Equivalently, $P[A]$ is obtained by replacing the scalar variable z with the matrix A , keeping the coefficients B_k in the exact order written. The two tensor factors are the *base* $\mathbb{C}^n$ and the *coefficient* $\mathbb{C}^m$.

All norms are the operator norms from the Euclidean inner products and their tensor products (equivalently, the largest singular value). The Hilbert–Schmidt norm of a matrix C is denoted $\|C\|_2 = (\text{tr } C^* C)^{1/2}$.

Theorem 1.1 (Complete Crouzeix inequality). *Let $n, m \geq 1$, let $A \in M_n(\mathbb{C})$, let $d \geq 0$, and let $P(z) = \sum_{k=0}^d B_k z^k$ with $B_k \in M_m(\mathbb{C})$. Then*

$$\|P[A]\| \leq 2 \max_{z \in W(A)} \|P(z)\|. \quad (1.3)$$

The constant 2 is optimal: it cannot be lowered uniformly in n, m, d .

The word *complete* refers to the uniformity in the coefficient dimension m . A scalar proof (the case $m = 1$) does not automatically extend, because scalar coefficients commute but matrix coefficients need not: $F(z)G(z)$ and $G(z)F(z)$ are different $M_m(\mathbb{C})$ -valued functions in general, and $F[A]G[A]$ is not $G[A]F[A]$. The proof below keeps these two products separate all the way through, which is why it works uniformly in m .

Why the constant cannot be less than 2

Take $n = 2$, $m = 1$, $A = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}$, and $P(z) = z$. The eigenvalue 0 of A has a two-dimensional generalised eigenvector space, and $\|A\| = 2$. Meanwhile the previous subsection showed that $W(A)$ is the closed unit disk, so $\max_{W(A)} |P(z)| = \max_{W(A)} |z| = 1$. The ratio $\|P[A]\| / \max_{W(A)} |P| = 2$, so no universal constant below 2 works. The rest of the paper proves that 2 is enough.

Roadmap

There are five steps. First, enclose $W(A)$ in a bounded convex domain Ω with real-analytic Jordan boundary. Second, build an exterior conformal map $h : \{|\lambda| > 1\} \rightarrow \mathbb{C} \setminus \overline{\Omega}$ parametrising $\partial\Omega$ by the unit circle. Third, expand the Cauchy kernel $q(\lambda)/(h(\lambda) - z)$, $q(\lambda) = \lambda h'(\lambda)$, at infinity to obtain scalar *Faber polynomials* $b_k(z)$ and a positive kernel K on $\mathbb{T}$; convexity enters here. Fourth, deduce from K the two-sided coefficient estimate

$$\sum_k \|C_k\|_2^2 \leq \int_{\mathbb{T}} \|G(h(\mu))\|_2^2 d\tau(\mu) \leq \|C_0\|_2^2 + 2 \sum_{k \geq 1} \|C_k\|_2^2, \quad (1.4)$$

whose factor of 2 is the sole source of the sharp constant. Fifth, pair with the exterior resolvent $R(\lambda) = q(\lambda)(h(\lambda)I - A)^{-1}$ and use positivity of $R + R^*$ to pass the coefficient estimate to $F[A]$ and force $\|F[A]\| \leq 2$. Finally, shrink Ω to $W(A)$.

Historical context

Crouzeix posed the scalar problem in 2004 [7] and gave a universal complete bound of 11.08 in 2007 [8]. Delyon and Delyon introduced positive boundary representations for the numerical range [10]. Crouzeix and Palencia obtained the complete bound $1 + \sqrt{2}$ [9], later simplified by Ransford and Schwenninger [16]. The complete constant 2 was obtained for base dimension 2 by Badea, Crouzeix, Delyon [3] and for base dimension 3 by Åhag, Czyż, Virtanen [1]. The scalar case $m = 1$ was settled by Jin [12], Lorist and Schwenninger [13], and Luo [14]. A structural proof of the complete case is in [15]. The present paper gives a direct, self-contained argument with all matrix dimensions independent.

2 Exterior coordinates and Faber coefficients

Throughout this section, $\Omega \subset \mathbb{C}$ is a bounded open convex set whose boundary is a *regular real-analytic Jordan curve*: a simple closed curve parametrised real-analytically with nowhere-zero derivative. The unit circle is $\mathbb{T} = \{\lambda : |\lambda| = 1\}$. The exterior unit disk is $\mathbb{C} \setminus \overline{\mathbb{D}} = \{\lambda : |\lambda| > 1\}$. Every circle integral uses $d\tau(e^{it}) = dt/(2\pi)$, so constant functions of modulus one have integral one.

The exterior coordinate

By Lemma A.1 in the Appendix, there is a number $0 < r_0 < 1$ and a map

$$h(\lambda) = c\lambda + c_0 + \sum_{j \geq 1} c_j \lambda^{-j}, \quad c \neq 0, \quad (2.1)$$

meromorphic on $\{\lambda : |\lambda| > r_0\} \cup \{\infty\}$, one-to-one there as a map of the Riemann sphere, with a simple pole at infinity, mapping $|\lambda| > 1$ conformally onto $\mathbb{C} \setminus \overline{\Omega}$ and $|\lambda| = 1$ onto $\partial\Omega$. This is the *exterior conformal map*, or *exterior coordinate* of Ω .

The Appendix gives a complete derivation from Riemann's mapping theorem. The reader who wants to accept it can take (2.1) and the statements below as basic inputs.

The supporting line and the outward normal

Define

$$q(\lambda) = \lambda h'(\lambda). \quad (2.2)$$

The two partial derivatives of h at a point $e^{i\theta} \in \mathbb{T}$ are $\partial_\theta h = iq$ (tangent to the boundary) and $\partial_r h = q$ where $r = |\lambda|$ (radial direction, pointing outward from $\overline{\Omega}$). So $q(\lambda)$ is a nonzero outward normal vector at $h(\lambda) \in \partial\Omega$, and $iq(\lambda)$ is a positively oriented tangent vector.

Because Ω is convex, the entire set $\overline{\Omega}$ lies on the inner side of the supporting line through $h(\lambda)$ with normal direction $q(\lambda)$:

$$\Re(\overline{q(\lambda)}(h(\lambda) - z)) \geq 0 \quad (\lambda \in \mathbb{T}, z \in \overline{\Omega}). \quad (2.3)$$

This is the only geometric input from convexity, and it is what makes the kernel below nonnegative. Figure 1 shows the picture.

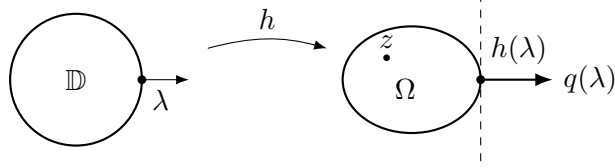


Figure 1: The map h takes the exterior unit circle to the exterior of $\overline{\Omega}$. The vector $q(\lambda) = \lambda h'(\lambda)$ is the outward normal at $h(\lambda)$. Convexity of Ω places every $z \in \overline{\Omega}$ on the inner side of the supporting line $\Re(\overline{q(\lambda)}(w - h(\lambda))) = 0$.

Faber polynomials

Define scalar polynomials $b_k(z)$, $k \geq 0$, by the expansion at infinity

$$\frac{q(\lambda)}{h(\lambda) - z} = \sum_{k \geq 0} b_k(z) \lambda^{-k}, \quad |\lambda| \gg 1. \quad (2.4)$$

The left side is holomorphic at $\lambda = \infty$ (since $h(\lambda) \sim c\lambda$ with $c \neq 0$), so the expansion makes sense. These b_k are the *Faber polynomials* of the exterior coordinate h ; they are the natural generalisation of the monomials z^k to the domain Ω . In the simplest case $\Omega = \mathbb{D}$ with $h(\lambda) = \lambda$, one has $b_k(z) = z^k$; for a general Ω , the b_k have corrections coming from the shape of $\partial\Omega$.

We can be more precise: after factoring $c\lambda$ from the denominator, we find $b_0 = 1$ and, for $k \geq 1$, b_k is a polynomial of degree k with leading coefficient c^{-k} . So $(b_k)_{k \geq 0}$ really is a basis of the space of polynomials, and polynomial degrees match indices. This is useful because it lets us rewrite any matrix polynomial $G(z) = \sum_{k=0}^N b_k(z) C_k$ uniquely in the Faber basis, with $C_k \in M_m(\mathbb{C})$ in the natural way.

Boundary values and the mixed coefficient

The boundary restrictions $b_k \circ h$ of the Faber polynomials have a remarkably clean Fourier form. There are scalars s_{kj} , indexed by $k, j \geq 1$, such that on the exterior collar one has

$$b_k(h(\mu)) = \mu^k + \sum_{j \geq 1} s_{kj} \mu^{-j}, \quad k \geq 1. \quad (2.5)$$

Equivalently, the *mixed correction kernel*

$$s(\lambda, \mu) = \frac{q(\lambda)}{h(\lambda) - h(\mu)} - \frac{\lambda}{\lambda - \mu} = \sum_{k, j \geq 1} s_{kj} \lambda^{-k} \mu^{-j} \quad (2.6)$$

collects exactly the difference between the true boundary kernel and the boundary kernel of the unit disk.

The function s has a nice structure. It extends holomorphically across the diagonal $\lambda = \mu$ and across both infinities, so it is regular on a full product collar of the exterior. Its mixed Taylor coefficients decay geometrically:

$$|s_{kj}| \leq C_\rho \rho^{k+j} \quad (r_0 < \rho < 1, \quad k, j \geq 1). \quad (2.7)$$

Consequently the double series converges normally on the exterior product collar, hence absolutely and uniformly on $\mathbb{T}^2$. So every interchange of sum and integral below is legal.

The kernel we will need is

$$K(\lambda, \mu) = 1 + s(\lambda, \mu) + \overline{s(\lambda, \mu)}. \quad (2.8)$$

Since s has only strictly mixed Fourier modes ($\lambda^{-k}\mu^{-j}$ with $k, j \geq 1$), neither s nor its conjugate contributes to the integral in λ alone or in μ alone. So

$$\int_{\mathbb{T}} K(\lambda, \mu) d\tau(\lambda) = 1 \quad (\text{for each fixed } \mu), \quad \int_{\mathbb{T}} K(\lambda, \mu) d\tau(\mu) = 1 \quad (\text{for each fixed } \lambda).$$

These are the *marginal identities*. They say: K is like the kernel of the identity operator, but with a real, symmetric, nonnegative correction.

The positive kernel theorem

Theorem 2.1 (Exterior Faber coefficients and the positive kernel). *Let Ω be a bounded convex domain with regular real-analytic Jordan boundary. Let h be as in (2.1), and let q , $(b_k)_{k \geq 0}$, and s_{kj} be defined by (2.2), (2.4), and (2.6). Then:*

1. $b_0 = 1$, and for each $k \geq 1$, b_k is a polynomial of degree k with leading coefficient c^{-k} in (2.1).
2. The mixed coefficients s_{kj} satisfy the geometric decay estimate (2.7).
3. The kernel K in (2.8) is continuous on $\mathbb{T}^2$, nonnegative everywhere, and has integral one in each variable with the other fixed.

Proof. Part 1 (degree and leading coefficient). Write $h(\lambda) - z = c\lambda \cdot (1 + (c_0 - z)/(c\lambda) + O(\lambda^{-2}))$. For large $|\lambda|$,

$$\frac{q(\lambda)}{h(\lambda) - z} = \frac{c\lambda + O(\lambda^{-1})}{c\lambda} \cdot \left(1 + \frac{c_0 - z}{c\lambda} + \dots\right)^{-1} = 1 + O(\lambda^{-1}).$$

Expanding the inverse as a geometric series in $(c_0 - z)/(c\lambda)$ produces, for the coefficient of λ^{-k} , a polynomial in z of degree k . Inspecting the largest power of z shows it is $c^{-k}z^k$. So $b_0 = 1$ and $b_k(z) = c^{-k}z^k + (\text{lower degree})$ for $k \geq 1$. In particular the b_k form a basis of the scalar polynomials.

Part 2 (geometric decay and joint extension). Consider

$$H(\lambda, \mu) = \frac{h(\lambda) - h(\mu)}{\lambda - \mu}. \quad (2.9)$$

Write $u = \lambda^{-1}$, $v = \mu^{-1}$, and $h(\lambda) = c\lambda + g(u)$ with g holomorphic on $|u| < r_0^{-1}$. Then

$$H(\lambda, \mu) = c - uv \frac{g(u) - g(v)}{u - v}.$$

The divided difference $(g(u) - g(v))/(u - v)$ is jointly holomorphic in (u, v) , with diagonal value $g'(u)$, so this formula covers the diagonal and infinities. Direct differentiation gives

$$s(\lambda, \mu) = \frac{\lambda \partial_\lambda H}{H} = -\frac{u \partial_u H}{H}.$$

Since H is nonzero on the reciprocal bidisk $|u|, |v| < r_0^{-1}$ (injectivity of h , and $H \rightarrow c$ at infinity), s is holomorphic there and vanishes on both axes $\{u = 0\}$ and $\{v = 0\}$. Its Taylor expansion therefore contains exactly the strictly mixed powers $\lambda^{-k}\mu^{-j}$ with $k, j \geq 1$.

For any $r_0 < \rho < 1$, apply Cauchy estimates on the closed disk $|u|, |v| \leq \rho^{-1}$: there is a constant C_ρ with $|s_{kj}| \leq C_\rho \rho^{k+j}$. Summing the resulting geometric series shows the Laurent expansion (2.6) converges normally (absolutely and uniformly on compacta, including those meeting infinity) throughout the product collar $|\lambda|, |\mu| > r_0$. Restricting to $|\lambda| = |\mu| = 1$ gives absolute uniform convergence on $\mathbb{T}^2$.

Comparing (2.4) and (2.6) coefficient-by-coefficient in λ^{-k} , with μ fixed in the exterior collar, yields (2.5). Indeed, $\lambda/(\lambda - \mu) = \sum_{k \geq 0} \mu^k \lambda^{-k}$, so the difference of the two sides matches $b_k(h(\mu)) - \mu^k$.

Part 3 (positivity of K and marginal identities). For distinct $\lambda, \mu \in \mathbb{T}$, split K as

$$K(\lambda, \mu) = 2 \Re \frac{q(\lambda)}{h(\lambda) - h(\mu)} + \left(1 - \frac{\lambda}{\lambda - \mu} - \frac{\bar{\lambda}}{\bar{\lambda} - \bar{\mu}} \right).$$

On $|\lambda| = |\mu| = 1$, we have $\bar{\lambda} = 1/\lambda$, $\bar{\mu} = 1/\mu$, and one verifies directly that $\lambda/(\lambda - \mu) + \bar{\lambda}/(\bar{\lambda} - \bar{\mu}) = 1$, so the bracketed term vanishes. Hence

$$K(\lambda, \mu) = 2 \Re \frac{q(\lambda)}{h(\lambda) - h(\mu)}.$$

Now $\Re(q/(h(\lambda) - h(\mu))) = |q|^2 / \Re(\bar{q}(h(\lambda) - h(\mu)))$ where $|q|^2 > 0$, and the denominator is nonnegative by (2.3) with $z = h(\mu)$. So $K(\lambda, \mu) \geq 0$ for distinct λ, μ , and continuity gives the same on the diagonal.

For the marginal identities: every term in s has Fourier mode $\lambda^{-k}\mu^{-j}$ with $k, j \geq 1$, so integrating against $d\tau(\lambda)$ picks out $k = 0$, which never occurs; and similarly for the conjugate. Hence $\int_{\mathbb{T}} s(\lambda, \mu) d\tau(\lambda) = \int_{\mathbb{T}} \overline{s(\lambda, \mu)} d\tau(\lambda) = 0$, so $\int_{\mathbb{T}} K(\lambda, \mu) d\tau(\lambda) = 1$. Symmetric argument in μ . $\square$

Identification with Faber polynomials

Since (2.5) gives $b_k(z) - h^{-1}(z)^k = O(z^{-1})$ as $z \rightarrow \infty$, the b_k are the classical *Faber polynomials* of the exterior coordinate h [4, 5].

3 The coefficient norm estimate

The positive kernel K from Theorem 2.1 is the entire source of the constant 2 in the main theorem. This section extracts one clean inequality from it. Fix any dimension m and any finite-family of coefficient matrices $C_0, C_1, \dots, C_N \in M_m(\mathbb{C})$, and put

$$G(z) = \sum_{k=0}^N b_k(z) C_k.$$

That is, G is an $M_m(\mathbb{C})$ -valued polynomial written in the Faber basis of Ω . We want to bound the Hilbert–Schmidt L^2 -norm of $G \circ h$ on the unit circle from above and below by the coefficient Hilbert–Schmidt norms $\|C_k\|_2$.

The operator T defined by K

Consider the bounded operator

$$(Tu)(\mu) = \int_{\mathbb{T}} K(\lambda, \mu) u(\lambda) d\tau(\lambda), \quad u \in L^2(\mathbb{T}; \mathcal{H}), \quad (3.1)$$

where $\mathcal{H}$ is any separable Hilbert space (in practice $\mathcal{H} = M_m(\mathbb{C})$ with the Hilbert–Schmidt inner product).

Lemma 3.1. *The operator T extends to a contraction of $L^2(\mathbb{T}; \mathcal{H})$.*

Proof. Nonnegativity $K \geq 0$ and the marginal identity $\int_{\mathbb{T}} K(\lambda, \mu) d\tau(\lambda) = 1$ together give, for each fixed μ , by Cauchy–Schwarz,

$$\|(Tu)(\mu)\|^2 \leq \left(\int_{\mathbb{T}} K(\lambda, \mu) d\tau(\lambda) \right) \left(\int_{\mathbb{T}} K(\lambda, \mu) \|u(\lambda)\|^2 d\tau(\lambda) \right) = \int_{\mathbb{T}} K(\lambda, \mu) \|u(\lambda)\|^2 d\tau(\lambda).$$

Integrate over μ and use the other marginal identity:

$$\|Tu\|_{L^2}^2 \leq \int_{\mathbb{T}} \left(\int_{\mathbb{T}} K(\lambda, \mu) d\tau(\mu) \right) \|u(\lambda)\|^2 d\tau(\lambda) = \|u\|_{L^2}^2.$$

□

What T does on monomials

For $u(\lambda) = \sum_{k=1}^N C_k \lambda^k$ (each $C_k \in M_m(\mathbb{C})$), only the s summand of K contributes to $(Tu)(\mu)$: the term 1 picks up $\int_{\mathbb{T}} \lambda^k d\tau = 0$, and the $\bar{s}$ term has only positive powers of λ . Hence

$$(Tu)(\mu) = \sum_{j \geq 1} \left(\sum_{k=1}^N s_{kj} C_k \right) \mu^{-j}. \quad (3.2)$$

(The normal convergence of Theorem 2.1 justifies every interchange.) Since the μ^{-j} are orthonormal in $L^2(\mathbb{T})$,

$$\|Tu\|_{L^2}^2 = \sum_{j \geq 1} \left\| \sum_{k=1}^N s_{kj} C_k \right\|_2^2,$$

where $\|\cdot\|_2$ denotes the Hilbert–Schmidt norm on $M_m(\mathbb{C})$. Combined with Lemma 3.1,

$$\sum_{j \geq 1} \left\| \sum_{k=1}^N s_{kj} C_k \right\|_2^2 \leq \sum_{k=1}^N \|C_k\|_2^2. \quad (3.3)$$

The two-sided coefficient inequality

By (2.5),

$$G(h(\mu)) = C_0 + \sum_{k=1}^N C_k \mu^k + \sum_{j \geq 1} \left(\sum_{k=1}^N s_{kj} C_k \right) \mu^{-j} = C_0 + u(\mu) + (Tu)(\mu).$$

The three summands have disjoint Fourier supports (constant, positive, negative), so Parseval and (3.3) give

$$\|C_0\|_2^2 + \sum_{k=1}^N \|C_k\|_2^2 \leq \int_{\mathbb{T}} \|G(h(\mu))\|_2^2 d\tau(\mu) \leq \|C_0\|_2^2 + 2 \sum_{k=1}^N \|C_k\|_2^2.$$

Passing from coefficients to boundary values never shrinks the coefficient Hilbert–Schmidt norm, and can enlarge it by at most a factor of 2.

Proposition 3.2 (The coefficient norm estimate). *For every positive integer m , every integer $N \geq 0$, and every $M_m(\mathbb{C})$ -valued polynomial $G(z) = \sum_{k=0}^N b_k(z) C_k$ in the Faber basis of Ω , one has*

$$\sum_{k=0}^N \|C_k\|_2^2 \leq \int_{\mathbb{T}} \|G(h(\mu))\|_2^2 d\tau(\mu) \leq \|C_0\|_2^2 + 2 \sum_{k=1}^N \|C_k\|_2^2. \quad (3.4)$$

Proof. This is exactly the combined identity above. The lower bound uses only $\|Tu\|^2 \geq 0$. The upper bound uses Lemma 3.1 to shoehorn $\|Tu\|_{L^2} \leq \sum_k \|C_k\|_2$ into the Parseval identity. $\square$

A note on the literature

In the notation of [15, Proposition 3.1], our h, λ, μ correspond to the companion's G, t, w , and $s(\lambda, \mu) = B(\mu, \lambda)$ (i.e., $s_{kj} = \beta_{jk}$). Equation (3.2) uses this orientation and does not require symmetry of the kernel.

4 The resolvent and the double-layer positivity

From now on the matrix $A \in M_n(\mathbb{C})$ is fixed, with $W(A) \subset \Omega$. We will also fix the exterior coordinate h of Ω from section 2. The purpose of this section is to define the *exterior resolvent* and show that it has a positive Hermitian part.

The resolvent function

Define, on the exterior circle $\mathbb{T}$,

$$R(\lambda) = q(\lambda) (h(\lambda)I - A)^{-1} \in M_n(\mathbb{C}). \quad (4.1)$$

Here $h(\lambda)I - A$ is invertible whenever $h(\lambda) \notin \sigma(A)$. Since $\sigma(A) \subset W(A) \subset \Omega$ (compact) and the circle $\mathbb{T}$ parametrises $\partial\Omega$ (compact, disjoint from $\sigma(A)$), the resolvent exists and is continuous on $\mathbb{T}$, and extends holomorphically to a collar $\{|\lambda| > r_A\}$ for some $r_0 < r_A < 1$ with r_A closer to one than r_0 .

Theorem 4.1 (Resolvent expansion and double-layer positivity). *Let $\Omega, h, q, (b_k)$ be as above. For every $n \geq 1$ and every $A \in M_n(\mathbb{C})$ with $W(A) \subset \Omega$, the following hold.*

1. Expansion. R has the Laurent series

$$R(\lambda) = \sum_{k \geq 0} b_k(A) \lambda^{-k}, \quad |\lambda| \geq 1, \quad (4.2)$$

converging absolutely and uniformly on $\mathbb{T}$. The coefficient matrices $b_k(A)$ decay geometrically: there are constants $C < \infty$ and $0 < \theta < 1$ (depending on A and h) with $\|b_k(A)\| \leq C\theta^k$ for all $k \geq 0$.

2. Means. $\int_{\mathbb{T}} R(\lambda) d\tau(\lambda) = I$ and $\int_{\mathbb{T}} (R(\lambda) + R(\lambda)^*) d\tau(\lambda) = 2I$.

3. Positivity. On the exterior circle, $R(\lambda) + R(\lambda)^* \succeq 0$.

Proof. Part 1 (expansion and decay). Every eigenvalue of A belongs to $W(A) \subset \Omega$, by testing the corresponding unit eigenvector. Hence the resolvent $(h(\lambda)I - A)^{-1}$ is holomorphic on and outside $\mathbb{T}$, including a small exterior collar. Expanding at ∞ in λ gives a Laurent series in λ^{-1} , and comparing with (2.4) with z replaced by A shows that the coefficient of λ^{-k} is exactly $b_k(A)$. The series converges absolutely and uniformly on $\mathbb{T}$ by compactness of the exterior collar. To get the geometric bound, pick $r_A < \theta < 1$ with R holomorphic on $|\lambda| > r_A$, apply Cauchy estimates on the circle $|\lambda| = \theta^{-1}$.

Part 2 (means). The constant coefficient $b_0(A) = b_0(z)|_{z=A} = I$ (Theorem 2.1, part 1). So by (4.2) and orthogonality of the modes λ^{-k} , $\int_{\mathbb{T}} R d\tau = b_0(A) = I$. Taking adjoints and adding gives $\int_{\mathbb{T}} (R + R^*) d\tau = 2I$.

Part 3 (positivity). Let $B(\lambda) = h(\lambda)I - A$ so that $R = qB^{-1}$. Then $R^* = \bar{q}(B^*)^{-1}$ and

$$R + R^* = qB^{-1} + \bar{q}(B^*)^{-1} = (B^*)^{-1}(\bar{q}B + qB^*)B^{-1}.$$

Since $(B^*)^{-1}$ and B^{-1} are adjoints of each other, they give a congruence that preserves positivity. So it suffices to show that $\overline{q(\lambda)}B(\lambda) + q(\lambda)B(\lambda)^* \succeq 0$.

For any unit vector $v \in \mathbb{C}^n$, the quadratic form on the left at v is

$$v^*(\bar{q}B + qB^*)v = \bar{q}v^*Bv + qv^*B^*v = 2\Re(\overline{q(\lambda)}v^*B(\lambda)v).$$

Substituting $B = h(\lambda)I - A$:

$$v^*Bv = h(\lambda) - v^*Av.$$

So the quadratic form is $\Re(\overline{q(\lambda)}(h(\lambda) - v^*Av))$. By the support inequality (2.3) with $z = v^*Av \in W(A) \subset \bar{\Omega}$, this is ≥ 0 . The congruence argument finishes the proof. $\square$

5 The main estimate on an analytic convex domain

Now the payoff. In this section $\Omega, h, q, (b_k)$ are as in Section 2, $A \in M_n(\mathbb{C})$ satisfies $W(A) \subset \Omega$, and $R(\lambda) = q(\lambda)(h(\lambda)I - A)^{-1}$ is the exterior resolvent from Theorem 4.1.

Statement and strategy

Theorem 5.1 (Bound on an analytic convex domain). *Let $n, m \geq 1$. Let $\Omega \subset \mathbb{C}$ be bounded, open, convex, with regular real-analytic Jordan boundary, and let $A \in M_n(\mathbb{C})$ with $W(A) \subset \Omega$. Let F be an $M_m(\mathbb{C})$ -valued polynomial with $\|F(z)\| \leq 1$ on $\partial\Omega$ (equivalently, on $\bar{\Omega}$, since F is a polynomial and the boundary controls the closure). Then*

$$\|F[A]\| \leq 2.$$

Three moves: top singular vectors x, y turn the norm into a scalar $y^*F[A]x = \|F[A]\|$; the two products FG and GF turn the diagonal functionals $x^*G[A]x$ and $y^*G[A]y$ into an off-diagonal one; and $R + R^* \succeq 0$ compares the resulting $P + P^*, Q + Q^*, M + L^*$ fields.

Setup: singular vectors and block notation

Put $\gamma = \|F[A]\|$. If $\gamma > 0$, choose unit vectors x, y in $\mathbb{C}^n \otimes \mathbb{C}^m$ forming a top singular pair:

$$F[A]x = \gamma y, \quad F[A]^*y = \gamma x. \quad (5.1)$$

If $\gamma = 0$, there is nothing to prove.

Write $x = \sum_{i=1}^m X e_i \otimes e_i$ and $y = \sum_{i=1}^m Y e_i \otimes e_i$ for $n \times m$ matrices X, Y , where $e_1, \dots, e_m$ is the standard basis of $\mathbb{C}^m$. Then $\|X\|_2^2 = \|Y\|_2^2 = 1$ (with $\|X\|_2^2 = \text{tr}(X^*X)$).

Define the $m \times m$ matrix functions

$$P = X^*RX, \quad Q = Y^*RY, \quad M = Y^*RX, \quad L = X^*RY. \quad (5.2)$$

Their Laurent coefficients in the variable λ^{-1} will be denoted with subscripts: $M_k = Y^*b_k(A)X$, $P_k = X^*b_k(A)X$, $Q_k = Y^*b_k(A)Y$, $L_k = X^*b_k(A)Y$. Since $\|b_k(A)\| \leq C\theta^k$ (Theorem 4.1), all four series converge absolutely and uniformly on $\mathbb{T}$, and all four are square-summable in the Hilbert-Schmidt norm.

Important bookkeeping at $k = 0$. Since $b_0(A) = I$,

$$P_0 = X^*X, \quad Q_0 = Y^*Y, \quad L_0^* = M_0. \quad (5.3)$$

In particular P_0 and Q_0 are Hermitian and positive semidefinite.

Finally, set

$$a^2 = \sum_{k \geq 0} \|M_k\|_2^2, \quad \text{finite and } a > 0. \quad (5.4)$$

The finiteness is from the geometric decay of $b_k(A)$. The positivity $a > 0$ follows from applying the cross-functional bound below to $G = F$: the value $y^*F[A]x = \gamma > 0$ cannot be achieved by a Cauchy-Schwarz bound with a zero right-hand side.

Step A. The off-diagonal functional

For any $M_m(\mathbb{C})$ -valued polynomial $G(z) = \sum_{k=0}^N b_k(z)C_k$, direct expansion in the tensor product gives

$$y^*G[A]x = \sum_{k=0}^N \sum_{i,j=1}^m (M_k)_{ij} (C_k)_{ij}. \quad (5.5)$$

Indeed, $G[A]x = \sum_{k,j} (b_k(A)X e_j) \otimes (C_k e_j)$, so

$$y^*G[A]x = \sum_{k,i,j} (Y e_i)^* b_k(A) (X e_j) \cdot e_i^* C_k e_j,$$

which is (5.5).

Cauchy-Schwarz on (5.5) gives

$$|y^*G[A]x| \leq \left(\sum_{k \geq 0} \|M_k\|_2^2 \right)^{1/2} \left(\sum_{k=0}^N \|C_k\|_2^2 \right)^{1/2} = a \left(\sum_{k=0}^N \|C_k\|_2^2 \right)^{1/2}. \quad (5.6)$$

Plugging in $G = F$ and using $y^*F[A]x = \gamma$ from (5.1) shows that $a > 0$.

Step B. Two ordered tests

Polynomial evaluation with matrix coefficients respects multiplication in the given order. If $F(z) = \sum_i D_i z^i$ and $G(z) = \sum_j E_j z^j$, then

$$(FG)[A] = \sum_{i,j} A^{i+j} \otimes D_i E_j = F[A]G[A].$$

Similarly $(GF)[A] = G[A]F[A]$ when factors are in the reversed order. Multiplying (5.1) on the left by suitable pieces,

$$y^*(FG)[A]x = y^*F[A]G[A]x = \gamma x^*G[A]x, \quad y^*(GF)[A]x = \gamma y^*G[A]y. \quad (5.7)$$

So the diagonal functionals $x^*G[A]x$ and $y^*G[A]y$ can be turned into the off-diagonal functional by ordering the two factors differently.

Write the two products in the Faber basis:

$$FG = \sum_{k \geq 0} b_k U_k, \quad GF = \sum_{k \geq 0} b_k V_k,$$

with finitely many nonzero coefficient matrices $U_k, V_k \in M_m(\mathbb{C})$. Define the boundary norm

$$\|H\|_{\partial}^2 = \int_{\mathbb{T}} \|H(h(\mu))\|_2^2 d\tau(\mu), \quad H \in M_m(\mathbb{C})[z].$$

Because $\|F(z)\| \leq 1$ on the boundary, submultiplicativity of the pointwise operator norm gives

$$\|FG\|_{\partial} \leq \|F\|_{\partial} \|G\|_{\partial} \leq \|G\|_{\partial}, \quad \|GF\|_{\partial} \leq \|G\|_{\partial}.$$

(Using $\|F(z)G(z)\|_2 \leq \|F(z)\| \|G(z)\|_2$ and its reverse-order analog.)

Apply (5.6) to FG and GF , and use (5.7), the lower coefficient estimate from Proposition 3.2, and the boundary multiplier bounds:

$$\begin{aligned} \gamma |x^*G[A]x| &= |y^*(FG)[A]x| \leq a \left(\sum_k \|U_k\|_2^2 \right)^{1/2} \leq a \|FG\|_{\partial} \leq a \|G\|_{\partial}, \\ \gamma |y^*G[A]y| &= |y^*(GF)[A]x| \leq a \left(\sum_k \|V_k\|_2^2 \right)^{1/2} \leq a \|GF\|_{\partial} \leq a \|G\|_{\partial}. \end{aligned}$$

The upper coefficient estimate for G in Proposition 3.2 turns $\|G\|_{\partial}$ into a norm of the coefficient array: $\|G\|_{\partial} \leq (\|C_0\|_2^2 + 2 \sum_{k \geq 1} \|C_k\|_2^2)^{1/2}$. Therefore

$$\max\{|x^*G[A]x|, |y^*G[A]y|\} \leq \frac{a}{\gamma} \left(\|C_0\|_2^2 + 2 \sum_{k \geq 1} \|C_k\|_2^2 \right)^{1/2}. \quad (5.8)$$

This is the crucial estimate that produces the weight of two.

Step C. Weighted dual norms

The functional $x^*G[A]x$ pairs C_k against P_k (repeat the calculation of (5.5), but with y replaced by x ; the coefficient array is $P_k = X^*b_k(A)X$). Similarly $y^*G[A]y$ pairs C_k against Q_k . So (5.8) constrains the size of the sequences (P_k) and (Q_k) , with weights that reverse the ones in (5.8).

Concretely, for each $N \geq 0$ set

$$d_N^{(P)} = \|P_0\|_2^2 + \frac{1}{2} \sum_{k=1}^N \|P_k\|_2^2.$$

Test (5.8) with the coefficient sequence

$$C_0 = \overline{P_0}, \quad C_k = \overline{P_k}/2 \quad (1 \leq k \leq N), \quad C_k = 0 \quad (k > N),$$

where the bar denotes entrywise complex conjugation. The scalar paired value is exactly $d_N^{(P)}$. The weighted norm on the right is $\|C_0\|_2^2 + 2 \sum_{k=1}^N \|C_k\|_2^2 = \|P_0\|_2^2 + \frac{1}{2} \sum_{k=1}^N \|P_k\|_2^2 = d_N^{(P)}$. So

$$d_N^{(P)} \leq \frac{a}{\gamma} (d_N^{(P)})^{1/2}.$$

If $d_N^{(P)} = 0$ this is trivial. Otherwise divide and square: $d_N^{(P)} \leq a^2/\gamma^2$. Letting $N \rightarrow \infty$:

$$\|P_0\|_2^2 + \frac{1}{2} \sum_{k \geq 1} \|P_k\|_2^2 \leq \frac{a^2}{\gamma^2}, \quad \|Q_0\|_2^2 + \frac{1}{2} \sum_{k \geq 1} \|Q_k\|_2^2 \leq \frac{a^2}{\gamma^2}. \quad (5.9)$$

The same finite test with Q_k in place of P_k gives the second bound.

Step D. Boundary L^2 norms of the blocks

Now compute the square-integrated Hilbert-Schmidt norms of the three extreme matrix fields $P + P^*$, $Q + Q^*$, and $M + L^*$ on the circle.

Diagonal blocks. The function $P + P^*$ has Laurent expansion

$$(P + P^*)(\lambda) = 2P_0 + \sum_{k \geq 1} P_k \lambda^{-k} + \sum_{k \geq 1} P_k^* \lambda^k$$

where we used $P_0^* = P_0$ and $\overline{\lambda}^k = \lambda^{-k}$ on $|\lambda| = 1$. The three groups (constant, negative powers, positive powers) have disjoint Fourier support, so Parseval gives

$$\int_{\mathbb{T}} \|P + P^*\|_2^2 d\tau(\lambda) = \|2P_0\|_2^2 + \sum_{k \geq 1} \|P_k\|_2^2 + \sum_{k \geq 1} \|P_k^*\|_2^2 = 4\|P_0\|_2^2 + 2 \sum_{k \geq 1} \|P_k\|_2^2 = 4d_\infty^{(P)},$$

where $d_\infty^{(P)} = \|P_0\|_2^2 + \frac{1}{2} \sum_{k \geq 1} \|P_k\|_2^2$ is the limit of the sequence $d_N^{(P)}$ from Step C. Combined with (5.9),

$$\int_{\mathbb{T}} \|P + P^*\|_2^2 d\tau \leq 4a^2/\gamma^2, \quad \int_{\mathbb{T}} \|Q + Q^*\|_2^2 d\tau \leq 4a^2/\gamma^2. \quad (5.10)$$

The second inequality follows from the second half of (5.9) by the same calculation with Q in place of P .

Cross block. The function $M + L^*$ has Laurent expansion

$$(M + L^*)(\lambda) = \sum_{k \geq 0} M_k \lambda^{-k} + \sum_{k \geq 0} L_k^* \lambda^k.$$

The constant coefficient is $M_0 + L_0^* = 2M_0$, using $L_0^* = M_0$ from (5.3). Grouping by Fourier mode,

$$\int_{\mathbb{T}} \|M + L^*\|_2^2 d\tau = \|2M_0\|_2^2 + \sum_{k \geq 1} \|M_k\|_2^2 + \sum_{k \geq 1} \|L_k^*\|_2^2.$$

Since $\|L_k^*\|_2 = \|L_k\|_2$ (the Hilbert-Schmidt norm is invariant under adjoint) and every term is nonnegative,

$$\int_{\mathbb{T}} \|M + L^*\|_2^2 d\tau \geq 4\|M_0\|_2^2 + \sum_{k \geq 1} \|M_k\|_2^2 \geq \sum_{k \geq 0} \|M_k\|_2^2 = a^2. \quad (5.11)$$

The final inequality holds because the coefficient of $4\|M_0\|_2^2$ exceeds the coefficient of $\|M_0\|_2^2$; dropping the $\|L_k\|_2^2$ terms only makes the left smaller.

Step E. Positive block comparison

The proof now closes with linear algebra.

From Theorem 4.1, $R(\lambda) + R(\lambda)^* \succeq 0$ for every $\lambda \in \mathbb{T}$. Compress by the $n \times 2m$ matrix $[X \ Y]$. The $2m \times 2m$ matrix of blocks

$$\Delta(\lambda) = \begin{pmatrix} P + P^* & L + M^* \\ M + L^* & Q + Q^* \end{pmatrix}(\lambda) = \begin{pmatrix} X \\ Y \end{pmatrix}^* (R + R^*)(\lambda) \begin{pmatrix} X & Y \end{pmatrix} \succeq 0$$

is positive semidefinite for every λ . A compression of a positive semidefinite matrix is positive semidefinite, so this is immediate.

Let $J = \text{diag}(I_m, -I_m)$. If $\Delta \succeq 0$, then $J\Delta J \succeq 0$ too (the conjugation by J keeps positivity). If two matrices S and T are positive semidefinite, then $\text{tr}(ST) \geq 0$: writing $S = AA^*$ and $T = BB^*$, $\text{tr}(ST) = \text{tr}(AA^*BB^*) = \text{tr}((A^*B)^*(A^*B)) = \|A^*B\|_2^2 \geq 0$. So

$$0 \leq \text{tr}(\Delta(J\Delta J)) = \text{tr}((\Delta J)^2).$$

Compute $(\Delta J)^2$ directly. $\Delta J = \begin{pmatrix} P + P^* & -(L + M^*) \\ M + L^* & -(Q + Q^*) \end{pmatrix}$. Then

$$\text{tr}((\Delta J)^2) = \text{tr}((P + P^*)^2) - \text{tr}((L + M^*)(M + L^*)) - \text{tr}((M + L^*)(L + M^*)) + \text{tr}((Q + Q^*)^2).$$

Both $(L + M^*)(M + L^*) = AA^*$ and $(M + L^*)(L + M^*) = A^*A$ (with $A = L + M^*$) have the same trace $\|A\|_2^2 = \|M + L^*\|_2^2 \geq 0$, so the two middle terms sum to

$$-\text{tr}((L + M^*)(M + L^*) + (M + L^*)(L + M^*)) = -2\text{tr}(L + M^*)(L + M^*) = -2\|M + L^*\|_2^2.$$

Putting it all together:

$$0 \leq \|P + P^*\|_2^2 + \|Q + Q^*\|_2^2 - 2\|M + L^*\|_2^2. \quad (5.12)$$

This holds for every $\lambda \in \mathbb{T}$.

Integrate (5.12) over $\mu \in \mathbb{T}$, use (5.10) for the two diagonal terms and (5.11) for the cross term:

$$0 \leq 4a^2/\gamma^2 + 4a^2/\gamma^2 - 2a^2 = 8a^2/\gamma^2 - 2a^2.$$

So

$$2a^2 \leq 8a^2/\gamma^2.$$

Since $a > 0$, divide by a^2 : $\gamma^2 \leq 4$, hence $\gamma \leq 2$.

This finishes the proof that $\|F[A]\| \leq 2$. $\square$

Remarks on the argument

The factor of two in (3.4) is sharp (Section 6). Convexity enters only through (2.3). For scalar $m = 1$ one could use $FG = GF$; here the two products are genuinely different, and using them separately is the point of Step B.

6 Enclosing domains and the general case

We have proved $\|F[A]\| \leq 2$ under the hypothesis $W(A) \subset \Omega$ with Ω an analytic convex domain. We now remove that hypothesis: for an arbitrary compact convex $W(A)$ we construct analytic convex neighborhoods Ω_j shrinking to $W(A)$, apply the previous section, and pass to the limit.

Analytic convex outer approximation

Lemma 6.1 (Analytic convex outer approximation). *Let $K \subset \mathbb{C}$ be nonempty, compact, and convex. There exist bounded open convex domains $\Omega_j \subset \mathbb{C}$ with regular real-analytic Jordan boundaries such that $K \subset \Omega_j$ for every j and*

$$\sup_{z \in \bar{\Omega}_j} \text{dist}(z, K) \xrightarrow{j \rightarrow \infty} 0.$$

Proof. The construction is a smooth convex interpolation between K and a small thickening. Fix $0 < \delta < 1$. Choose a finite δ -net D_δ of unit directions that includes $1, -1, i, -i$. Let $s_K(\nu) = \max_{z \in K} \Re(\bar{\nu}z)$ be the support function of K . Choose $t_\delta > 0$ small enough that $|D_\delta|e^{-t_\delta\delta} < 1$.

Define the smooth, strictly convex function

$$\Phi_\delta(z) = \sum_{\nu \in D_\delta} \exp(t_\delta (\Re(\bar{\nu}z) - s_K(\nu) - \delta))$$

and set $\Omega_\delta = \{z \in \mathbb{C} : \Phi_\delta(z) < 1\}$.

(i) $K \subset \Omega_\delta$. For $z \in K$ and any ν , $\Re(\bar{\nu}z) \leq s_K(\nu)$, so each exponential is at most $e^{-t_\delta\delta}$, and the sum is at most $|D_\delta|e^{-t_\delta\delta} < 1$.

(ii) $\bar{\Omega}_\delta$ lies in a fixed rectangle. If $\Phi_\delta(z) \leq 1$, then every summand is ≤ 1 (since each is nonnegative), hence $\Re(\bar{\nu}z) \leq s_K(\nu) + \delta$ for every $\nu \in D_\delta$. In particular, for $\nu = 1$ and $\nu = -1$, we get $\Re z$ in $[\min_{w \in K} \Re w - \delta, \max_{w \in K} \Re w + \delta]$; and similarly the imaginary part. So all $\bar{\Omega}_\delta$ for $0 < \delta < 1$ sit inside one fixed bounded rectangle $R \subset \mathbb{C}$.

(iii) Φ_δ is strictly convex with positive definite Hessian. The Hessian is a sum of positive multiples of $\nu\nu^\top$ (outer products), where ν ranges over D_δ . Each $\nu\nu^\top$ is positive semidefinite; the four directions $1, -1, i, -i$ together span $\mathbb{R}^2$, so the sum is strictly positive definite. Hence Φ_δ is strictly convex and real analytic.

(iv) *Level set is regular and a Jordan curve.* If z_0 were a critical point of Φ_δ with $\Phi_\delta(z_0) = 1$, then by convexity Φ_δ would attain a global minimum at z_0 , contradicting $\Phi_\delta(\text{any } z \in K) < 1$. So $\nabla \Phi_\delta \neq 0$ on $\Phi_\delta^{-1}(1)$. The implicit function theorem makes that level a regular real-analytic curve. Pick an interior point $w_0 \in \Omega_\delta$ (any point of K will do). Each ray from w_0 crosses $\partial\Omega_\delta$ exactly once, because Φ_δ is strictly convex and bounded. Radial projection gives a continuous bijection $\partial\Omega_\delta \rightarrow \mathbf{S}^1$, so $\partial\Omega_\delta$ is a Jordan curve and $\overline{\Omega}_\delta = \{\Phi_\delta \leq 1\}$.

(v) *Distance to K .* For $z \in \overline{\Omega}_\delta$ let $w \in K$ be the nearest point and $d = |z - w|$. If $d > 0$, set $\nu = (z - w)/d$, pick $\mu \in D_\delta$ with $|\mu - \nu| \leq \delta$, and let $v \in K$ attain $s_K(\mu) = \Re(\overline{\mu}v)$.

By minimality of w , $\Re(\overline{\nu}(w' - w)) \leq 0$ for every $w' \in K$; in particular $\Re(\overline{\nu}(v - w)) \leq 0$. By the support inequality at μ , $\Re(\overline{\mu}(v - w)) \geq 0$.

Then

- $\Re(\overline{\mu}(z - w)) = d\Re(\overline{\mu}\nu) \geq (1 - \delta)d$ (since $|\mu - \nu| \leq \delta$).
- $\Re(\overline{\mu}(z - v)) = \Re(\overline{\mu}z - s_K(\mu)) \leq \delta$ (from $\Phi_\delta(z) \leq 1$ and $s_K(\mu) = \Re(\overline{\mu}v)$).
- $\Re(\overline{\mu}(v - w)) \leq \delta|v - w|$ (from $\Re(\overline{\nu}(v - w)) \leq 0$ and $|\mu - \nu| \leq \delta$).

So

$$(1 - \delta)d \leq \Re(\overline{\mu}(z - w)) \leq \delta + \delta|v - w| \leq \delta(1 + \text{diam } K),$$

i.e., $\text{dist}(z, K) = d \leq \delta(1 + \text{diam } K)/(1 - \delta)$. Points of K have distance zero, so the same bound holds throughout. Taking $\delta_j = 1/(j + 1)$ gives the Ω_j . $\square$

Proof of the main theorem

Proof of Theorem 1.1. First recall the Toeplitz–Hausdorff convexity theorem [17, 11]: $W(A)$ is convex.

Compactly. $W(A) = \{u^*Au : u^*u = 1\}$ is the image of the unit sphere (compact) under a continuous map, hence is compact.

Convexly. Take any $u, v \in \mathbb{C}^n$ with $|u| = |v| = 1$. We want to show that the segment $[u^*Au, v^*Av]$ lies in $W(A)$. Let $V = \text{span}\{u, v\}$, of complex dimension 1 or 2.

If $\dim V = 1$, then $v = e^{i\theta}u$ and $v^*Av = u^*Au$, so the segment is trivial.

If $\dim V = 2$, work in an orthonormal basis $\{e_1, e_2\}$ of V , and let B be the 2×2 matrix of $P_V A|_V$ in this basis. For a unit vector $x = x_1e_1 + x_2e_2$, writing $X = 2\Re(\overline{x_1}x_2)$, $Y = 2\Im(\overline{x_1}x_2)$, $Z = |x_1|^2 - |x_2|^2$, one has $X^2 + Y^2 + Z^2 = 1$ (because $4|x_1x_2|^2 + (|x_1|^2 - |x_2|^2)^2 = (|x_1|^2 + |x_2|^2)^2 = 1$), so (X, Y, Z) ranges over the unit sphere $\mathbf{S}^2 \subset \mathbb{R}^3$, and x^*Bx is a real-linear function $f(X, Y, Z)$ into $\mathbf{C} \cong \mathbb{R}^2$. The linear part of f has a one-dimensional kernel. Every point of the unit ball can be pushed along the kernel direction onto the sphere without changing f , so $f(\mathbf{S}^2) = f(\mathbb{B}^3)$. The image of a ball under a linear map is convex, so $f(\mathbf{S}^2)$ is convex. In particular, the segment $[u^*Au, v^*Av] \subset f(\mathbf{S}^2) \subset W(A)$.

So $W(A)$ is compact and convex.

Enclosing. Apply Lemma 6.1 with $K = W(A)$. Get analytic convex domains Ω_j with $W(A) \subset \Omega_j$ and $\sup_{\overline{\Omega}_j} \text{dist}(\cdot, W(A)) \rightarrow 0$. In particular every $\overline{\Omega}_j$ lies inside the fixed bounded rectangle from Lemma 6.1(ii).

Uniform bound. Fix one Ω_j . Let $M_j = \sup_{z \in \overline{\Omega}_j} \|P(z)\|$. Two cases.

If $M_j = 0$, then P vanishes on Ω_j (open), so P is the zero polynomial and $P[A] = 0$.

If $M_j > 0$, multiply P by $1/M_j$. The normalized polynomial satisfies $\|P/M_j\| \leq 1$ on $\partial\Omega_j$, so Theorem 5.1 applies and gives $\|(P/M_j)[A]\| \leq 2$, i.e., $\|P[A]\| \leq 2M_j$.

In either case, $\|P[A]\| \leq 2M_j$.

Shrinking. Since $\bar{\Omega}_j$ converges to $W(A)$ in Hausdorff distance and all $\bar{\Omega}_j$ lie in a fixed compact set, uniform continuity of the continuous function $z \mapsto \|P(z)\|$ implies $M_j \rightarrow \sup_{z \in W(A)} \|P(z)\| = \max_{W(A)} \|P\|$. Passing to the limit in $\|P[A]\| \leq 2M_j$,

$$\|P[A]\| \leq 2 \max_{z \in W(A)} \|P(z)\|.$$

This is exactly (1.3). □

Sharpness of the constant 2

Sharpness in Theorem 1.1. Take $n = 2$, $m = 1$, $P(z) = z$, and $A = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}$. Then $W(A)$ is the closed unit disk and $\|A\| = 2$, so $\|P[A]\| = 2$ while $\max_{W(A)} |P| = 1$: the ratio is exactly 2. □

A The exterior conformal collar

This appendix supplies the classical conformal input used in Section 2. Everything here is covered by Ahlfors [2]; we spell out the two ingredients that are rarely stated explicitly: real-analytic boundary continuation and global injectivity on a small exterior collar.

Lemma A.1 (Exterior coordinate and a univalent collar). *Let $\Omega \subset \mathbb{C}$ be a bounded convex domain with regular real-analytic Jordan boundary. Then there are $0 < r_0 < 1$ and a map h which is meromorphic and one-to-one on $\{\lambda : |\lambda| > r_0\} \cup \{\infty\}$, has a simple pole at ∞ , and maps $|\lambda| > 1$ conformally onto $\mathbb{C} \setminus \bar{\Omega}$, and $|\lambda| = 1$ onto $\partial\Omega$. The expansion at infinity is*

$$h(\lambda) = c\lambda + c_0 + \sum_{j \geq 1} c_j \lambda^{-j}, \quad c \neq 0.$$

For λ on the unit circle, $q(\lambda) = \lambda h'(\lambda)$ is a nonzero outward normal vector at $h(\lambda)$, and the support inequality (2.3) holds.

Proof. Step 1: A bounded simply connected model. Choose a point $z_* \in \Omega$ and apply $z \mapsto 1/(z - z_*)$ (with ∞ mapped to 0). The spherical exterior $(\mathbb{C} \setminus \bar{\Omega}) \cup \{\infty\}$ maps onto a bounded simply connected domain U containing 0. Because Ω is convex, each ray from z_* meets $\partial\Omega$ in a unique point; after inversion, the exterior portions of those rays become a domain that is star-shaped about 0. Its boundary is a regular real-analytic Jordan curve: the inversion of a regular real-analytic curve is again one.

Step 2: Riemann mapping. By the Riemann mapping theorem [2, Ch. 6, §1.1, Thm. 1], there is a conformal bijection $g : \mathbb{D} \rightarrow U$ with $g(0) = 0$. By Carathéodory's Jordan-domain theorem [2, Ch. 6, §1.2], g extends to a homeomorphism $\bar{\mathbb{D}} \rightarrow \bar{U}$.

Step 3: Schwarz reflection. At each boundary point, Schwarz reflection [2, Ch. 4, §6.5; Ch. 6, §§1.3–1.4] extends g holomorphically across $\partial\mathbb{D}$; local continuations patch by the identity theorem to a single holomorphic function on a neighborhood of $\bar{\mathbb{D}}$.

Step 4: Injectivity on a neighborhood. Otherwise, distinct w_n, w'_n with $\tilde{g}(w_n) = \tilde{g}(w'_n)$ would converge (after subsequences) to a common w^* in $\bar{\mathbb{D}}$. Since g is a homeomorphism on $\bar{\mathbb{D}}$, both sequences eventually lie in a small neighborhood of w^* on which $\tilde{g}$ is injective, contradiction. So $\tilde{g}$ is injective on an open neighborhood of $\bar{\mathbb{D}}$, which contains $\{w : |w| < R_0\}$ for some $R_0 > 1$.

Step 5: Assembly. The function g is univalent on $\{w : |w| < R_0\} \supset \overline{\mathbb{D}}$. Its only zero there is 0, and that zero is simple ($g'(0) \neq 0$ by conformality). Set

$$h(\lambda) = z_* + \frac{1}{g(1/\lambda)}, \quad r_0 = R_0^{-1}.$$

Then h is meromorphic on $\{|\lambda| > r_0\} \cup \{\infty\}$, has a simple pole at ∞ (from the simple zero of g at 0), maps $|\lambda| > 1$ conformally onto $\mathbb{C} \setminus \overline{\Omega}$ (from g mapping $\mathbb{D}$ to U and $z \mapsto z_* + 1/z$ undoing the inversion), and maps $|\lambda| = 1$ onto $\partial\Omega$. The expansion $h(\lambda) = c\lambda + c_0 + O(1/\lambda)$ has $c = 1/g'(0) \neq 0$.

Step 6: The support inequality. At $\lambda = e^{i\theta} \in \mathbb{T}$, the radial and angular derivatives of h are $q(\lambda) = \lambda h'(\lambda)$ and $iq(\lambda)$, respectively. They are perpendicular and nonzero. Increasing the radial parameter $|\lambda|$ traces out the exterior of $\overline{\Omega}$, so $q(\lambda)$ points outward. Since Ω is convex with a smooth boundary, the supporting-line property holds at $h(\lambda)$: for every $z \in \overline{\Omega}$,

$$\Re(\overline{q(\lambda)}(h(\lambda) - z)) \geq 0.$$

This is exactly (2.3). □

References

- [1] P. Åhag, R. Czyż, and J. Virtanen, *The complete Crouzeix conjecture in dimension three and the Clouâtre–Ostermann–Ransford conjecture*, [arXiv:2608.27346v4](#), September 21, 2026.
- [2] L. V. Ahlfors, *Complex Analysis*, third ed., McGraw–Hill, New York, 1979.
- [3] C. Badea, M. Crouzeix, and B. Delyon, Convex domains and K -spectral sets, *Math. Z.* **252** (2006), 345–365. [doi:10.1007/s00209-005-0857-y](#).
- [4] B. Beckermann, Image numérique, GMRES et polynômes de Faber, *C. R. Math.* **340** (2005), 855–860. [doi:10.1016/j.crma.2005.04.027](#).
- [5] B. Beckermann and M. Crouzeix, Faber polynomials of matrices for non-convex sets, *Jaen J. Approx.* **6** (2014), 219–231. [arXiv:1310.1356v1](#).
- [6] B. Beckermann and N. Stylianopoulos, Bergman orthogonal polynomials and the Grunsky matrix, *Constr. Approx.* **47** (2018), 211–235. [doi:10.1007/s00365-017-9381-7](#).
- [7] M. Crouzeix, Bounds for analytical functions of matrices, *Integral Equations Operator Theory* **48** (2004), 461–477. [doi:10.1007/s00020-002-1188-6](#).
- [8] M. Crouzeix, Numerical range and functional calculus in Hilbert space, *J. Funct. Anal.* **244** (2007), 668–690. [doi:10.1016/j.jfa.2006.10.013](#).
- [9] M. Crouzeix and C. Palencia, The numerical range is a $(1 + \sqrt{2})$ -spectral set, *SIAM J. Matrix Anal. Appl.* **38** (2017), 649–655. [doi:10.1137/17M1116672](#).
- [10] B. Delyon and F. Delyon, Generalization of von Neumann’s spectral sets and integral representation of operators, *Bull. Soc. Math. France* **127** (1999), 25–41. [doi:10.24033/bsmf.2340](#).
- [11] F. Hausdorff, Der Wertvorrat einer Bilinearform, *Math. Z.* **3** (1919), 314–316. [doi:10.1007/BF01292610](#).

- [12] S. Jin, *The numerical range is a 2-spectral set*, [Preprints.org 202607.1919](#), version 4, August 7, 2026.
- [13] E. Lorist and F. L. Schwenninger, *A solution to Crouzeix’s conjecture*, [arXiv:2608.03841v2](#), August 17, 2026.
- [14] Q. Luo, *Two boundary-certificate proofs of the Crouzeix theorem: Fiberwise Fourier recurrence and defect-Gram telescoping*, [Preprints.org 202608.1661](#), version 2, August 28, 2026.
- [15] OpenAI, *The complete Crouzeix theorem: optimal similarity and a common positive boundary representation*, OpenAI Math Release preprint, [OAI:The-complete-Crouzeix-theorem-September-23-2026](#), 2026.
- [16] T. Ransford and F. L. Schwenninger, Remarks on the Crouzeix–Palencia proof that the numerical range is a $(1 + \sqrt{2})$ -spectral set, *SIAM J. Matrix Anal. Appl.* **39** (2018), 342–345. [doi:10.1137/17M1143757](#).
- [17] O. Toeplitz, Das algebraische Analogon zu einem Satze von Fejér, *Math. Z.* **2** (1918), 187–197. [doi:10.1007/BF01212904](#).