The Abdus Salam
International Centre for Theoretical Physics

United Nations
Educational, Scientific and
Cultural Organization

IAEA
International Atomic Energy Agency

2044-5

# Summer School and Advanced Workshop on Trends and Developments in Linear Algebra

*22 June - 10 July, 2009*

## Numerical ranges (classical and higher-rank) with applications to QIT (quantum information theory)

John Holbrook
*University of GuelphGuelph ON*
*Canada*

Summer School and Advanced Workshop on Trends and Developments in Linear Algebra

International Centre for Theoretical Physics, Trieste, Italy

22 June – 10 July, 2009

Version 1 of lecture notes related to my course

Numerical ranges (classical and higher-rank) with applications to QIT (quantum information theory)

John Holbrook, University of Guelph, Canada

Remarks: Concerning the versions: these notes are a work in progress and early versions will at some points be merely outlines; later versions are intended to fill in the details.
Concerning the exercises: these may be helpful in grasping concepts and techniques that might be new to some; the exercises in some cases introduce auxiliary material; if there is interest we may discuss solutions to some of the exercises during the problem/tutorial sessions. The exercises vary widely in difficulty, from the routine to some that are considerably more challenging (meaning I don't know good solutions myself).
Concerning prerequisites: I try to make the lectures accessible to those familiar with the following parts of a standard undergraduate math program: linear algebra and matrices, metric spaces and multivariate analysis, plus a little Fourier analysis and complex function theory.

## 1. The setting

We work in a complex Hilbert space $\mathcal{H}$, which we may usually take to be finite–dimensional, and so represented as the $n \times 1$ column space $\mathbb{C}^n$. The inner product $(x, y) = \sum_1^n x_k \overline{y_k} = y^* x$ as a matrix product. For $M \in M_{n,m} = n \times m$ complex matrices, $M^*$ is the conjugate transpose:

$$M^* = \overline{M^t} = [\overline{m_{ji}}] \text{ where } M = [m_{ij}].$$

Key property: for a square matrix $M$ (ie $M \in M_n = M_{n,n}$) we have

$$(Mx, y) = (x, M^* y).$$

Exercise: Verify this property.

Remark: We tend to use the notations most familiar to mathematicians. In physics and QIT the notation varies: eg inner products are often conjugate linear in the first variable ($< x, y >= x^*y$) and $M^*$ is replaced by $M^\dagger$.

The Euclidian norm on $\mathcal{H}$ is given by $\|x\| = \sqrt{(x,x)}$ and the operator norm of a matrix $M$ is given by $\max\{\|Mu\| : \|u\| = 1\}$, so that $\|M\|$ is the Lipschitz constant of $M$ as a mapping on $(\mathcal{H}, \|\cdot\|)$.

2. Numerical range and radius (classical)

For $M \in M_n$ we define the numerical range of $M$ as

$$W(M) = \{(Mu, u) : \|u\| = 1\}.$$

Clearly $W(M)$ is a compact subset of the complex plane $\mathbb{C}$ (for an operator $T$ on an infinite–dimensional $\mathcal{H}$, however, $W(T)$ might not be closed).

Examples:

If $M \in M_2$, then $W(M)$ is a (filled–in) ellipse (related, as we'll see, to the qubit and Bloch sphere in QIT).

If $M \in M_3$, then $W(M)$ may have a greater variety of shapes, typically three–lobed.

Figure: typical $W(M)$ for $M$ in $M_2$ and $M_3$ ...

Proposition: If $M \in M_n$ is a normal matrix (ie $MM^* = M^*M$), then

$$W(M) = \text{conv}\{\lambda_j : j = 1, 2, \ldots, n\},$$

the convex hull of the eigenvalues $\lambda_j$ of $M$.
Proof: The spectral theorem for normal $M$ tells us that we may choose an orthonormal basis of eigenvectors $u_j$: $Mu_j = \lambda_j u_j$ and $(u_j, u_k) = \delta_{jk}$ (ie $= 1$ if $j = k$, $= 0$ otherwise). Thus $\|u\| = 1$ implies that $u = \sum z_j u_j$ with $\sum |z_j|^2 = 1$ and

$$(Mu, u) = (\sum z_j \lambda_j u_j, \sum z_k u_k) = \sum |z_j|^2 \lambda_j$$

(cross–terms are 0 by orthogonality of the eigenvectors), and this is a convex combination of the $\lambda_j$. QED

Remark: For any $M$, normal or otherwise, each eigenvalue $\lambda$ lies in $W(M)$ (consider a unit eigenvector $u$ corresponding to $\lambda$).

Examples:

If $M$ is Hermitian, ie $M^* = M$, the eigenvalues $\lambda_j$ are real and we may assume that $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$. Thus $W(M) = \text{conv}\{\lambda_j : j = 1, 2, \ldots, n\} = [\lambda_1, \lambda_n]$, an interval in the real line $\mathbb{R}$.

If $M$ is unitary, ie $M^* = M^{-1}$, the eigenvalues $\lambda_j$ lie on the unit circle (ie $|\lambda_j| = 1$) and $W(M)$ is a (filled-in) polygon inscribed in the unit circle.

Figure: typical $W(M)$ for a $5 \times 5$ unitary $M$ ...

Exercise: Sketch $W(M)$ where $M$ is the permutation matrix

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Exercise: Show that $W(\cdot)$ is "weakly unitarily invariant", ie for any unitary $U$ we have $W(UMU^*) = W(M)$.

Exercise: Show that if $M$ is block–diagonal with $M = M_1 \oplus M_2$ then $W(M) = \text{conv}\{W(M_1) \cup W(M_2)\}$.

The numerical radius $w(M)$ is defined as $\max\{|(Mu, u)| : \|u\| = 1\}$.

Figure: relationship between $W(M)$ and $w(M)$ ...

Proposition: The numerical radius $w(\cdot)$ is a norm on $M_n$, equivalent to the operator norm as follows:

$$w(M) \le \|M\| \le 2w(M).$$

Remark: Here the fact that we work in complex $\mathcal{H}$ is essential; for example, if on $\mathbb{R}^2$ the matrix $M$ represents rotation by 90 degrees, ie $M = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, then $(Mu, u) = 0$ for all $u$, so that $W(M) = \{0\}$ and $w(M) = 0$, whereas $\|M\| = 1$.

Proof: It is easy to see that $w(\cdot)$ is a seminorm, ie that $w(X+Y) \le w(X) + w(Y)$ and $w(zX) = |z|w(X)$. If $\|u\| = 1$ the Cauchy–Schwarz inequality tells us that $|(Mu, u)| \le \|Mu\| \, \|u\| \le \|M\| \, \|u\|^2 = \|M\|$. Hence $W(M)$ lies inside the ball of radius $\|M\|$ centred at 0 in $\mathbb{C}$; clearly then $w(M) \le \|M\|$.

On the other hand, $\|M\| = \max\{|(Mu, w)| : \|u\| = \|w\| = 1\}$ and we can use a "polarization" identity such as

$$|(Mu, w)| = |\frac{1}{2\pi} \int_0^{2\pi} (M(u + e^{i\theta}w), (u + e^{i\theta}w))e^{i\theta} \, d\theta|;$$

here we rely simply on the orthogonality of the complex harmonics:

$$\frac{1}{2\pi} \int_0^{2\pi} e^{i(j-k)\theta} \, d\theta = \delta_{jk}.$$

It follows that $|(Mu, w)|$ is no greater than

$$\frac{1}{2\pi} \int_0^{2\pi} |(M(u+e^{i\theta}w), (u+e^{i\theta}w))e^{i\theta}| \, d\theta \le \frac{1}{2\pi} \int_0^{2\pi} w(M)((u+e^{i\theta}w), (u+e^{i\theta}w)) \, d\theta$$

$$= w(M)((u, u) + (w, w)) = 2w(M).$$

QED

3. Some historical highlights

19th century: the range of the "Rayleigh quotient" $(Mv, v)/(v, v)$ was used in the Stürm–Liouville theory of 2nd order PDEs. In effect, $W(M)$ played a role, but usually in a context where $M$ was Hermitian.

c. 1918: Toeplitz and Hausdorff considered $W(M)$ for arbitrary $M \in M_n$ and proved that it is always a convex subset of $\mathbb{C}$ (see the Toeplitz–Hausdorff Theorem, below).

c. 1960: Lax and Wendroff studied iterative methods for the numerical solution of certain PDEs. These were modeled by computing powers of a matrix $M \in M_n$ and stability of the method depended on the power–bounded property of $M$: $\sup_k \|M^k\| < \infty$. Lax and Wendroff noted that $w(M) \leq 1$ is a sufficient condition for power–boundedness and established that

$$w(M) \leq 1 \quad \Rightarrow \quad \sup_k \|M^k\| \leq a_n$$

for all $M \in M_n$, where the $a_n$ are constants depending on $n$ and tending to infinity as $n$ increases. For more details, consult the book of Gustafson and Rao [G–R1997], pp 98 and following.

c. 1965 Halmos proposed that such inequalities should be independent of the size $n$ of the matrix, and Berger responded by proving that $w(M^k) \leq w^k(M)$. Thus

$$w(M) \leq 1 \quad \Rightarrow \quad \|M^k\| \leq 2w(M^k) \leq 2w^k(M) \leq 2(!)$$

(see Berger's power inequality, below).

c. 2006: Choi, Kribs, and Życzkowski introduced the higher-rank numerical ranges of matrices related to the noise operators for a quantum channel and to the problem of finding correctable codes for such a channel.

c. 2007: Woerdeman proved the convexity of the higher–rank numerical ranges (a striking extension of the Toeplitz–Hausdorff Theorem) by applying the theory of the algebraic Riccati equation to earlier results of Choi, Giesinger, Holbrook, Kribs, and Życzkowski.

c. 2008: Li and Sze revealed important aspects of the higher–rank numerical range structure by representing it as an intersection of directly computable half–planes.

Let's put some flesh on these bones ...

4. Toeplitz–Hausdorff and the Bloch sphere of quantum mechanics

Our Hilbert space $\mathcal{H} = \mathbb{C}^n$ may represent the state space of a quantum system, where the "pure states" of the system are identified with unit vectors $u \in \mathbb{C}^n$. Since $u$ and $e^{i\theta}u$ represent the same state, it is convenient to identify the state with the matrix $uu^*$ ($e^{i\theta}$ "drops out"). As a mapping $uu^*$ is a rank–one projection: $(uu^*)v = u(u^*v) = (v, u)u$, the orthogonal projection of $v$ onto the one–dimensional subspace $\mathrm{span}\{u\}$.

With this point of view, we can see the numerical range $W(M)$ as the image of the set of pure states via a certain linear functional $\varphi_M$:

$W(M) = \{(Mu, u) : \|u\| = 1\} = \{u^*Mu : \|u\| = 1\} = \{\mathrm{trace}(u^*Mu) : \|u\| = 1\}$,

where $\mathrm{trace}(X)$ is the trace of the square matrix $X \in M_n$: $\mathrm{trace}(X) = \sum_1^n x_{kk}$.

Recall the "commutativity of the trace":

Proposition: If $A$ is $n \times m$ and $B$ is $m \times n$ then $\mathrm{trace}(AB) = \mathrm{trace}(BA)$.
Exercise: Verify this proposition.

Here $u^*(Mu)$ is $1 \times 1$ whereas $(Mu)u^*$ is $n \times n$:

$$(Mu, u) = \mathrm{trace}(u^*Mu) = \mathrm{trace}(Mu)u^* = \mathrm{trace}M(uu^*),$$

so that $W(M) = \varphi_M(\{uu^* : \|u\| = 1\}) = \varphi_M(\text{pure states})$, where $\varphi_M$ is the linear functional defined on $M_n$ by $\varphi_M(X) = \mathrm{trace}(MX)$.

When $n = 2$ the quantum system is known as the qubit, the most elementary form of quantum information. The pure states correspond to unit $u \in \mathbb{C}^2$ and these are (in various ways) represented by the "Bloch sphere", a geometrically faithful copy of the ordinary sphere $S^2$ in $\mathbb{R}^3$. For example, if $u = \begin{bmatrix} a \\ b + ci \end{bmatrix}$ is a unit vector in $\mathbb{C}^2$ with $a, b, c \in \mathbb{R}$ (since $e^{i\theta}u$ and $u$ are identified this can represent any qubit pure state), then

$$uu^* = \begin{bmatrix} a^2 & ab - aci \\ ab + aci & b^2 + c^2 \end{bmatrix}$$

with $a^2 + b^2 + c^2 = 1$. Thus

$$uu^* = \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 2a^2 - 1 & 2ab - 2aci \\ 2ab + 2aci & 2(b^2 + c^2) - 1 \end{bmatrix}.$$

Note that $2a^2 - 1 = 1 - 2(b^2 + c^2) = x$ and setting also $2ab = y$ and $2ac = z$ we may write

$$uu^* = \frac{1}{2}\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} + \frac{1}{2}(x\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix} + y\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix} + z\begin{bmatrix}0 & -i\\ i & 0\end{bmatrix}). \tag{1}$$

It is easy to verify that $x^2 + y^2 + z^2 = 1$:

$$1 - x^2 = (1 - x)(1 + x) = 2(b^2 + c^2)2a^2 = 4(a^2b^2 + a^2c^2) = y^2 + z^2.$$

Conversely, starting from any $(x, y, z) \in \mathbb{R}^3$ such that $x^2 + y^2 + z^2 = 1$ (ie $(x, y, z) \in S^2$) we can reverse the steps above (set $a = \pm\sqrt{(1 + x)/2}$, etc). Thus the pure states of a qubit are identified with (an affine copy of) $S^2$, and we obtain one form of the Bloch sphere representation of the qubit.

We also see that when $n = 2$ the numerical range $W(M)$ is the image of $S^2$ in $\mathbb{R}^3$ under a linear functional $\varphi_M$ mapping to $\mathbb{C} = \mathbb{R}^2$. Such an image must be a (filled–in) ellipse in $\mathbb{C}$; consequently, $W(M)$ is convex for any $M \in M_2$. The convexity of $W(M)$ for $n \times n$ matrices $M$ is only a step away.

Toeplitz–Hausdorff Theorem: For any $M \in M_n$, the numerical range $W(M)$ is a convex subset of $\mathbb{C}$.
Proof: (Except for our explicit mention of quantum systems and Bloch spheres, this is the approach of Davis in [D1971].) Consider distinct points $(Mu, u)$ and $(Mw, w)$ in $W(M)$; here $\|u\| = \|w\| = 1$ and $u, w$ are linearly independent. The pure states $uu^*$ and $ww^*$ are part of the qubit represented by the two–dimensional subspace span$\{u, w\}$. Thus $W(M)$ includes trace$(Mvv^*)$ for every $vv^*$ in the corresponding Bloch sphere; these constitute a (filled–in) ellipse in $\mathbb{C}$ containing $(Mu, u)$ and $(Mw, w)$. QED

Exercise: We have seen that $W(M)$ is an ellipse (filled–in) for every $M \in M_2$. The form of (1) makes it clear that the centre of the ellipse must be trace$(M(1/2)I_2)$ = trace$(M)/2$. Find the foci of $M(W)$ and the lengths of the major and minor axes. Hint: By choosing an orthonormal basis $u_1, u_2$ where $u_1$ is an eigenvector of $M$ we may put $M$ in upper triangular form $\begin{bmatrix}\lambda & \beta\\ 0 & \mu\end{bmatrix}$. (This is a very special case of the Schur upper triangular form for families of commuting matrices.) Note that $\lambda$ and $\mu$ are the eigenvalues of $M$.

Exercise: In our proof of the Toeplitz–Hausdorff Theorem we have viewed the set of pure states on $\mathbb{C}^n$ as a union

$$\bigcup\{S^2(u_1,u_2) : u_1, u_2 \text{ orthonormal}\},$$

where $S^2(u_1,u_2)$ is the Bloch sphere corresponding to the two–dimensional subspace span$\{u_1,u_2\}$. Show that the pure states on $\mathbb{C}^3$ may be regarded as the Cartesian product $S^2 \times S^5$ or, more economically, as $S^2 \times CP^2$ where $CP^2$ is the complex projective plane.

Exercise: Pauli matrices in (1) (see Petz [P2008] p 6); Frobenius inner product; ...

Exercise: $W(M) = \varphi_M(\text{conv}\{\text{pure states}\})$; conv{pure states} = all positive semidefinite matrices with trace 1 = mixed states (representing a statistical distribution of pure states) = density matrices ...

5. Berger's power inequality

Berger's original approach used a sort of "unitary dilation" for $M$ such that $w(M) \leq 1$: a unitary operator $U$ on a larger (usually infinite–dimensional) Hilbert space $\mathcal{K}$ containing $\mathcal{H}$ such that on $\mathcal{H}$

$$M^k = 2P_{\mathcal{H}}U^k$$

for $k = 1, 2, \ldots$. Here we use a related technique that confines the action to $\mathcal{H}$ itself.

Lemma: For any $M \in M_n$, $w(M) < 1$ iff for all $\theta \in \mathbb{R}$ and $h \in \mathcal{H} \setminus \{\vec{0}\}$

$$\sum_{k=1}^{\infty} e^{-ik\theta}((M^*)^k h, h) + 2\|h\|^2 + \sum_{k=1}^{\infty} e^{ik\theta}(M^k h, h) > 0. \tag{2}$$

Proof: Since each eigenvalue $\lambda$ of $M$ is in $W(M)$, $w(M) < 1$ implies that $|\lambda^k| \leq w^k(M) \to_k 0$ geometrically. Suppose $M$ has $n$ distinct eigenvalues (the "generic" case); then $M = S\text{diag}(\lambda_1, \ldots, \lambda_n)S^{-1}$ for some nonsingular $S$ and $\|M^k\| \leq \|S\|\,\|S^{-1}\|w^k(M)$. If the Jordan canonical form of $M$ has some blocks larger than $1 \times 1$, a somewhat more complicated argument shows that $\|M^k\| \leq$ Const.$r^k$ for any $r < 1$ such that $r > w(M)\,(\geq \max|\lambda_j|)$. In fact, though, in our applications of this lemma, we may assume $M$ is generic since any $M$ is arbitrarily close to generic matrices.

In any case, we see that the LHS of (2) makes sense as an absolutely convergent series. Now $w(M) < 1$ also implies that $\mathrm{Re}\, e^{i\theta}(Mg, g) < \|g\|^2$ for all $g \in \mathcal{H} \setminus \{\vec{0}\}$ and $\theta \in \mathbb{R}$; thus

$$\mathrm{Re}\,((I - e^{i\theta}M)g, g) > 0.$$

With $h = (I - e^{i\theta}M)g$ we have $\mathrm{Re}\,(h, (I - e^{i\theta}M)^{-1}h) > 0$ so that

$$2\mathrm{Re}\,((I - e^{i\theta}M)^{-1}h, h) > 0.$$

Using the "Neumann" series $(I - e^{i\theta}M)^{-1} = \sum_{k=0}^{\infty}(e^{i\theta}M)^k$, we obtain (2). These steps may be reversed. QED

Note: Recall that (by virtue of the orthogonality of the complex harmonics $e^{ik\varphi}$) the convolution product

$$f * g(\theta) = \frac{1}{2\pi}\int_0^{2\pi} f(\theta - \varphi)g(\varphi)\,d\varphi = \sum_{k=-\infty}^{\infty} \hat{f}(k)\hat{g}(k)e^{ik\theta},$$

where $f(\theta) = \sum_{-\infty}^{\infty} \hat{f}(k)e^{ik\theta}$ and $g(\theta) = \sum_{-\infty}^{\infty} \hat{g}(k)e^{ik\theta}$; we may assume here that $\sum_{-\infty}^{\infty} |\hat{f}(k)| < \infty$ and $\sum_{-\infty}^{\infty} |\hat{g}(k)| < \infty$.

To prove Berger's power inequality $w(M^k) \leq w^k(M)$ $(k = 1, 2, \dots)$ it is sufficient to show that $w(M) \leq 1$ implies $w(M^k) \leq 1$ (because $w(\cdot)$ is homogeneous). In fact, we may as well prove a more general result (compare Kato [K1965], Berger and Stampfli [BS1967]):

Proposition: If $w(M) \leq 1$ then $w(\varphi(M)) \leq 1$ for any function $\varphi$ analytic on $\mathbb{D}$ (the closed unit disc) such that $\varphi : \mathbb{D} \to \mathbb{D}$ and $\varphi(0) = 0$.
Proof: By means of appropriate approximations of $M$ and $\varphi$ we may assume that $w(M) < 1$ and that $\varphi$ is a polynomial with $\varphi(z) = \sum_{j=1}^{m} c_j z^j$ and

$$\|\varphi\|_\infty = \max\{|\varphi(z)| : z \in \mathbb{D}\} < 1.$$

It follows (spectral mapping theorem) that for each eigenvalue $\lambda$ of $Q = \varphi(M)$ we have $|\lambda| < 1$; hence $\|Q^k\| \to_k 0$ geometrically (recall the discussion of this point in the proof of the lemma above). We'll show that $w(Q) < 1$ using the lemma, ie by verifying that for all $h \in \mathcal{H} \setminus \{\vec{0}\}$ and $\theta \in \mathbb{R}$

$$\sum_{k=1}^{\infty} e^{-ik\theta}((Q^*)^k h, h) + 2\|h\|^2 + \sum_{k=1}^{\infty} e^{ik\theta}(Q^k h, h) > 0.$$

The LHS above is the convolution product $f * g(\theta)$ where

$$f(\theta) = \sum_{k=1}^{\infty} \overline{(e^{i\theta}\varphi(e^{i\theta}))^k} + 1 + \sum_{k=1}^{\infty} (e^{i\theta}\varphi(e^{i\theta}))^k$$

and

$$g(\theta) = \sum_{k=1}^{\infty} e^{-ik\theta}((M^*)^k h, h) + 2\|h\|^2 + \sum_{k=1}^{\infty} e^{ik\theta}(M^k h, h).$$

The lemma tells us that $g(\theta) > 0$. We also have $f(\theta) > 0$, since $f(\theta) = 1 + 2\mathrm{Re}(w/(1-w))$ where $w = e^{i\theta}\varphi(e^{i\theta})$: $|w| < 1$ and

$$f(\theta) = \mathrm{Re}(1 + 2w/(1-w)) = \mathrm{Re}(\frac{1+w}{1-w}) = \frac{1-|w|^2}{|1-w|^2} > 0.$$

It follows that the integral $f * g(\theta) > 0$. QED

6. Multiplicative properties of the numerical radius

Berger's inequality leads naturally to the question: when do we have $w(AB) \leq w(A)w(B)$? The answer is "not usually", even when $A$ and $B$ commute. However, Berger's inequality does tell us what we can expect along these lines:

Proposition: If $AB = BA$ then

$$w(AB) \leq 2w(A)w(B). \tag{3}$$

Proof: Since $w(\cdot)$ is homogeneous, it is enough to show that $w(A), w(B) \leq 1$ implies $w(AB) \leq 2$. Since $AB = BA$ we have

$$AB = \frac{1}{4}((A+B)^2 - (A-B)^2)$$

and

$$w(AB) \leq \frac{1}{4}(w((A+B)^2) + w((A-B)^2)).$$

A special case of Berger's inequality then shows that

$$w(AB) \leq \frac{1}{4}((w(A+B))^2 + (w(A-B))^2),$$

and this is no greater than

$$\frac{1}{4}((w(A)+w(B))^2 + (w(A)+w(B))^2) \leq \frac{1}{4}((1+1)^2 + (1+1)^2) = 2.$$

QED

Exercise: Let $J_2$ denote the $2 \times 2$ Jordan nilpotent $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. Show that $w(J_2) = 1/2$ and, more generally, that $w(X) = \|X\|/2$ whenever $X^2 = 0_n$.

Exercise: Show that the inequality (3) is best possible even for (commuting) $4 \times 4$ matrices by verifying that $w(AB) = 1/2 = w(A) = w(B)$ for $A = I_2 \otimes J_2$ and $B = J_2 \otimes I_2$. Recall that the tensor product (or Kronecker product) $X \otimes Y$ of matrices $X$ and $Y$ is the block matrix $[x_{ij}Y]$.

Exercise: Show that $w(X \otimes Y) = w(Y \otimes X) \leq w(X)\|Y\|$. Hint: $X \otimes Y$ and $Y \otimes X$ are unitarily similar (via a permutation matrix); we may assume $\|Y\| = 1$ so that $Y = (U+W)/2$ where $U$ and $W$ are unitary and for the inequality we may take $Y$ to be a unitary $U$; by a unitary similarity we may assume $U$ is diagonal so that $U \otimes X$ is block diagonal.

Exercise: Complete the following proof (due to Ando) that $XY = YX$ implies $w(XY) \leq \sqrt{2}w(Y)\|X\|$:

$$(XYu, u)^2 = ((XY \otimes XY)u \otimes u, u \otimes u) = ((X \otimes Y)(Y \otimes X)u \otimes u, u \otimes u);$$

use (3) and the exercise above.
Remark: Better inequalities are known: if

$$c = \max\{w(XY) : XY = YX, \|X\| = w(Y) = 1\},$$

it can be shown that $1.06 < c < 1.17$.

In certain rare situations $w(\cdot)$ does turn out to be submultiplicative. For example:

Proposition: If $B/w(B) = \psi(A/w(A))$ for some analytic $\psi : \mathbb{D} \to \mathbb{D}$, then $w(AB) \leq w(A)w(B)$.
Proof: We must show that $w(C\psi(C)) \leq 1$, where $C = A/w(A)$. Note that $w(C) = 1$ so that the eigenvalues of $C$ are in $\mathbb{D}$ and $\psi(C)$ is well–defined. Now $C\psi(C) = \varphi(C)$ where $\varphi(x) = z\psi(z)$, and $w(\varphi(C)) \leq 1$ by virtue of the proposition of section 5. QED

Remark: The question of existence of such a mapping $\psi$ is a Pick-Nevanlinna interpolation problem and depends on whether or not the Pick matrix relating matching eigenvalues of $A/w(A)$ and $B/w(B)$ is positive semi–definite. In the following application we essentially use a special case of this procedure.

Proposition: If $A$ and $B$ are commuting $2 \times 2$ matrices, then $w(AB) \leq w(A)w(B)$.
Proof: Let $A' = A/w(A)$ and $B' = B/w(B)$. By the proposition above it is sufficient to find analytic $\psi : \mathbb{D} \to \mathbb{D}$ such that $\psi(A') = B'$ or $\psi(B') = A'$. We know about the elliptical form of $W(X)$ for $2 \times 2$ $X$, allowing us to assume that the eigenvalues $\alpha_1, \alpha_2, \beta_1, \beta_2$ of $A'$ and $B'$ are strictly inside $\mathbb{D}$ (otherwise $A'$ or $B'$ is diagonal, etc). Number the eigenvalues so that $\alpha_k, \beta_k$ belong to common eigenvectors of $A', B'$.

Recall the Möbius functions $\mu_\alpha : \mathbb{D} \to \mathbb{D}$ defined, when $|\alpha| < 1$, by $\mu_\alpha(z) = (z - \alpha)/(1 - \overline{\alpha}z)$. These are automorphisms of $\mathbb{D}$, ie analytic injective maps from $\mathbb{D}$ onto $\mathbb{D}$. Let $\alpha_3 = \mu_{\alpha_1}(\alpha_2)$ and $\beta_3 = \mu_{\beta_1}(\beta_2)$. We may assume $|\alpha_3| \geq |\beta_3|$ (otherwise reverse the roles of $A$ and $B$), so that $\beta_3 = \gamma\alpha_3$ for some $|\gamma| \leq 1$. Let $\psi(z) = \mu_{\beta_1}^{-1}(\gamma\mu_{\alpha_1}(z))$. Then

$$\psi(\alpha_1) = \mu_{\beta_1}^{-1}(\gamma \cdot 0) = \mu_{\beta_1}^{-1}(0) = \beta_1,$$

and

$$\psi(\alpha_2) = \mu_{\beta_1}^{-1}(\gamma\alpha_3) = \mu_{\beta_1}^{-1}(\beta_3) = \beta_2.$$

Hence $\psi(A') = B'$. QED

Exercise: Note that $w(A) = \|A\|$ if $A$ is normal. Show also that $w(AB) \leq w(A)w(B)$ if $AB = BA$ and at least one of $A, B$ is normal. Hint: If $A$ is normal then we may assume $A$ is diagonal and $B$ is block–diagonal.

Exercise: If we trace back through the ingredients used in our proof that $w(\cdot)$ is submultiplicative for commuting $2 \times 2$ matrices, we see that the proof is rather elaborate. Since we have a simple description of the numerical ranges of $2 \times 2$ matrices, it is natural to ask for a more elementary proof. Can you find one?

Exercise: We have seen that

$$\max\{w(AB) : AB = BA, w(A) = w(B) = 1, A, B \in M_n\} = 2$$

if $n \geq 4$. We have also seen that the corresponding constant is 1 for $2 \times 2$ matrices. What about $3 \times 3$ matrices? It can be shown that

$$c_3 = \max\{w(AB) : AB = BA, w(A) = w(B) = 1, A, B \in M_3\}$$

is strictly between 1 and 2. Can you find a better upper bound? Experiments reveal that the lower bound is at least 1.1.

## 7. Quantum channels, correctable codes, and the higher–rank numerical ranges

Communication within and between quantum systems is modeled by the action of a quantum channel $\mathcal{E}$ on density matrices. Recall that a density matrix $\rho$ is a positive semi-definite matrix with trace 1, and it models a statistical distribution of possible pure states. A quantum channel is determined by certain matrices $E_j$, known as noise, error, Kraus, or Choi operators for the channel; it acts as follows:

$$\mathcal{E}(\rho) = \sum_j E_j \rho E_j^*,$$

where $\sum_j E_j^* E_j = I$.

Exercise: Show that this last condition is necessary and sufficient for the channel to be trace–preserving.

Note that such a channel $\mathcal{E}$ preserves positivity, ie $\rho$ positive semi-definite (psd) implies $\mathcal{E}(\rho)$ also psd. In fact such a map $\mathcal{E}$ is completely positive; Choi derived the operator sum format from this abstract condition, with at most $n^2$ noise operators $E_j$ required for an $n$–dimensional system ($\rho \in M_n$).

Typically information is lost when $\rho$ passes through such a quantum channel, but there may be correctable subspaces $S \subseteq \mathcal{H}$, ie $S$ such that another quantum channel $\mathcal{R}$ may be designed so that it recovers $\rho$ from $\mathcal{E}(\rho)$ for all $\rho$ supported on $S$:

$$\mathcal{R}(\mathcal{E}(\rho)) = \rho$$

provided all eigenvectors of $\rho$ corresponding to positive eigenvalues lie in $S$.

Exercise: Show that $\rho$ is supported on $S$ iff it has the form $\rho_1 \oplus 0$ with respect to the orthogonal decomposition $\mathcal{H} = S \oplus S^\perp$.

Bennett, DiVincenzo, Smolin, and Wooters, and independently Knill and Laflamme discovered (c. 1996) a criterion for correctable subspace: a subspace $S$ is correctable for $\mathcal{E}$ iff there are $\lambda_{ij} \in \mathbb{C}$ such that, for all $i, j$,

$$PE_i^* E_j P = \lambda_{ij} P, \quad \text{(BDKLSW)}$$

where $P = P_S$, orthogonal projection onto $S$. We examine the justification of the criterion (BRKLSW) in the next section.

Exercise: Show that the matrix $[\lambda_{ij}]$ occurring in (BDKLSW) must be psd.

Since, for quantum error correction, it is better to have a larger, rather than smaller, correctable "code space" $S$, Choi, Kribs, and Życzkowski had the happy idea (c. 2006) of stratifying the problem of finding correctable $S$. They defined the rank–$k$ numerical range of a matrix $M$ as

$$\Lambda_k(M) = \{\lambda \in \mathbb{C} : \text{ for some rank–}k \text{ projection } P,\ PMP = \lambda P\}.$$

Clearly in applying the BDKLSW criterion we examine $\Lambda_k(M)$ where $M = E_i^* E_j$ so that a key question about higher-rank numerical ranges is: when are they nonempty? In fact, in (BDKLSW) we are looking for some sort of "joint" higher–rank numerical range (discussed in a later section) because the same $P$ must work for all $i, j$; however, the notion has a rich theory even for a single $M$, and this case already has some applications in QIT.

Remark: The equation $PMP = \lambda P$ is equivalent to $(M - \lambda I)S \perp S$ and so to the existence of orthonormal $u_1, \ldots, u_k$ (any o.n. basis for $S$) such that $(M - \lambda I)u_i \perp u_j$, ie $(Mu_i, u_j) = \lambda \delta_{ij}$. Clearly then, the classical numerical range $W(M)$ coincides with $\Lambda_1(M)$, and the terminology "rank–$k$ numerical range" makes sense.

In fact $\Lambda_k(M)$ corresponds to the scalar matrices in the $k$–th spatial numerical range of $M$, a notion that has been studied for some time. However, the remarkable properties of this scalar part, and the relations with QIT, have come to light only recently.

8. Verification of the BDKLSW criterion

to be completed ...

9. Elementary geometry of the higher–rank numerical ranges and the CKŻ conjecture

Consider first a normal matrix $M \in M_n$. Let the eigenvalues of $M$ be $\lambda_j$. Define

$$\Delta_k(M) = \{\lambda : \text{there exist } k \text{ disjoint index sets } J_1, J_2, \ldots, J_k \subseteq \{1, 2, \ldots, n\}$$

$$\text{such that } \lambda \in \bigcap_{i=1}^{k} \operatorname{conv}\{\lambda_j : j \in J_i\}\ \}.$$

We use the delta notation as a reminder of the crucial disjointness of the index sets.

Figure: $\Delta_2(U)$ for a typical $7 \times 7$ unitary $U$ ...

By an explicit construction we can show that each point $\lambda$ of $\Delta_k(M)$ is also in $\Lambda_k(M)$:

Proposition: Let $M$ be a normal matrix as above. Then $\Delta_k(M) \subseteq \Lambda_k(M)$.
Proof: Let $u_j$ $(j = 1, 2, \ldots, n)$ be an o.n. basis of eigenvectors with $Mu_j = \lambda_j u_j$. Consider any $\lambda \in \Delta_k(M)$. For each $i$ we can express $\lambda$ as a convex

combination $\sum_{j \in J_i} t_{ij}\lambda_j$. Let

$$w_i = \sum_{j \in J_i} \sqrt{t_{ij}} u_j.$$

It is easy to see that $w_1, \ldots, w_k$ are o.n. Note that $Mw_i = \sum_{j \in J_i} \sqrt{t_{ij}}\lambda_j u_j$ so that

$$(Mw_i, w_i) = (\sum_{j \in J_i} \sqrt{t_{ij}}\lambda_j u_j, \sum_{j \in J_i} \sqrt{t_{ij}} u_j) = \sum_{j \in J_i} t_{ij}\lambda_j = \lambda.$$

Disjointness of the index sets ensures also that $(Mw_i, w_j) = 0$ if $i \neq j$.

It follows that $PMP = \lambda P$ where $P$ is orthogonal projection onto $S = \text{span}\{w_1, \ldots, w_k\}$. QED

The proposition above gives a simple "lower bound" for $\Lambda_k(M)$ in the normal case. We also have a simple "upper bound" that applies to any $M$, normal or not, in terms of classical numerical ranges of "compressions" $P_L M|_L$ of $M$ to certain subspaces $L$.

Proposition: For any $M \in M_n$,

$$\Lambda_k(M) \subseteq \bigcap \{W(P_L M|_L) : L \text{ is a subspace of dimension } \geq n - k + 1\}.$$

Proof: For $\lambda \in \Lambda_k(M)$ we have a $k$-dimensional subspace $S$ such that $(Mu, u) = \lambda$ for any unit vector $u \in S$. If the subspace $L$ has dimension no less than $n - k + 1$, then $S$ and $L$ must meet nontrivially. Consider unit $u \in S \cap L$:

$$(P_L Mu, u) = (Mu, P_L u) = (Mu, u) = \lambda,$$

so that $\lambda \in W(P_L M|_L)$. QED

Choi, Kribs, and Życzkowski used similar ideas to identify the higher–rank numerical ranges of Hermitian matrices. Let $M \in M_n$ be Hermitian with (real) eigenvalues ordered as follows: $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$.

Proposition: Let $M$ be a Hermitian matrix as above. Then $\Lambda_k(M) = [\lambda_k, \lambda_{n-k+1}]$ when $k \leq (n+1)/2$ and otherwise $\Lambda_k(M)$ is empty.
Proof: If $k \leq (n+1)/2$, the disjoint index sets

$$J_1 = \{1, n\}, J_2 = \{2, n-1\}, \ldots, J_k = \{k, n-k+1\}$$

show that

$$\Delta_k(M) \supseteq \bigcap_{i=1}^{k} [\lambda_i, \lambda_{n-i+1}] = [\lambda_k, \lambda_{n-k+1}].$$

Hence $\Lambda_k(M) \supseteq [\lambda_k, \lambda_{n-k+1}]$. On the other hand, if $u_j$ are o.n. eigenvalues with $Mu_j = \lambda_j u_j$ and $L = \mathrm{span}\{u_k, \ldots, u_n\}$ we have $W(P_L M|_L) = [\lambda_k, \lambda_n]$ while with $L' = \mathrm{span}\{u_1, \ldots, u_{n-k+1}\}$ we have $W(P_{L'} M|_{L'}) = [\lambda_1, \lambda_{n-k+1}]$. The last proposition shows that

$$\Lambda_k(M) \subseteq [\lambda_k, \lambda_n] \cap [\lambda_1, \lambda_{n-k+1}] = [\lambda_k, \lambda_{n-k+1}].$$

The relation above also shows that $\Lambda_k(M) = \emptyset$ if $k > (n+1)/2$. QED

More generally, for any normal $M \in M_n$ with eigenvalues $\lambda_j$ and corresponding o.n. eigenvectors $u_j$ we see that, if $J$ is any index set and $L = \mathrm{span}\{u_j : j \in J\}$, then $W(P_L M|_L) = \mathrm{conv}\{\lambda_j : j \in J\}$. Hence $\Lambda_k(M) \subseteq \Omega_k(M)$ where

$$\Omega_k(M) = \bigcap \{\mathrm{conv}\{\lambda_j : j \in J\} : \#(J) = n - k + 1\}.$$

Choi, Kribs, and Życzkowski made the conjecture that we actually have equality: $\Lambda_k(M) = \Omega_k(M)$ for every normal $M$. A related conjecture would be that for any $M \in M_n$, normal or not, we have

$$\Lambda_k(M) = \bigcap \{W(P_L M|_L) : L \text{ is a subspace of dimension } n - k + 1\}.$$

As we'll see in later sections, both these conjectures are now theorems.

If, for normal $M$, we have $\Delta_k(M) = \Omega_k(M)$ (as in the case of Hermitian $M$), we obtain an easy proof of the CKŻ conjecture for that $M$. Before the work on convexity and/or the Li–Sze half–spaces was available, it had been noted that, for example, this method suffices to prove the CKŻ conjecture for normal $M$ when $n \geq 3k$.

Exercise: Verify that $\Delta_k(M) = \Omega_k(M)$ when $M$ is a $7 \times 7$ unitary and $k = 2$, but that this is not usually the case when $M$ is a $5 \times 5$ unitary (and $k = 2$).

## 10. Convexity via Riccati equations

Although convexity of the higher–rank numerical ranges $\Lambda_k(M)$ is just one consequence of the Li–Sze representation as the intersection of half–planes (discussed in a later section), it is instructive to see how convexity may be obtained from the theory of the algebraic Riccati equation. Thus our goal in this section is to follow one approach to the following striking extension of the Toeplitz–Hausdorff Theorem:

Proposition: For any $M \in M_n$ each of the higher–rank numerical ranges $\Lambda_k(M)$ is convex (although some may be empty).

First note that to show $\Lambda_k(M)$ convex it's enough to consider the case where $n = 2k$: if $\lambda_1$ and $\lambda_2$ are distinct points in $\Lambda_k(M)$ there are $k$–dimensional subspaces $L_1$ and $L_2$ such that

$$(M - \lambda_1)L_1 \perp L_1 \text{ and } (M - \lambda_2)L_2 \perp L_2,$$

with $L_1 \cap L_2 = \{\vec{0}\}$. Let $S = \text{span}\{L_1 \cup L_2\}$, a subspace of dimension $2k$. To show that $(\lambda_1 + \lambda_2)/2 \in \Lambda_k(M)$ it is enough to know that $(\lambda_1 + \lambda_2)/2 \in \Lambda_k(P_S M|_S)$, and $P_S M|_S$ may be represented by a matrix in $M_{2k}$.

Note also that $\Lambda_k(\cdot)$ respects affine transformations:

$$\Lambda_k(\alpha M + \beta I) = \alpha \Lambda_k(M) + \beta$$

for all complex $\alpha, \beta$.

Exercise: Verify this property.

By the appropriate affine transformation we may replace $\lambda_1, \lambda_2$ by $\pm 1$ and we reduce our problem to showing that $\pm 1 \in \Lambda_k(M)$ implies $0 \in \Lambda_k(M)$. Let $S_\pm$ have dimension $k$ be such that

$$(M \mp I)S_\pm \perp S_\pm.$$

Let $V$ be isometric on each of $S_\pm$, and let $V$ also "straighten" the subspaces: $VS_+ \perp VS_-$. Then

$$((V^{-1})^* M V^{-1} \mp I) VS_\pm \perp VS_\pm,$$

so that with respect to $VS_+ \oplus VS_-$ we have $(V^{-1})^* M V^{-1}$ represented by the matrix

$$\begin{bmatrix} I_k & X \\ Y & -I_k \end{bmatrix}$$

for some $X, Y \in M_k$.

Note also that the property $0 \in \Lambda_k(M)$ is indifferent to conjugation:

Lemma: If $Q$ is nonsingular then $0 \in \Lambda_k(M)$ iff $0 \in \Lambda_k(Q^*MQ)$.
Proof: If $0 \in \Lambda_k(Q^*MQ)$ then there is a $k$–dimensional subspace $S$ such that $Q^*MQS \perp S$. It follows that $M(QS) \perp QS$, and $QS$ is also $k$–dimensional. QED

Thus it suffices to show that $0 \in \Lambda_k(\begin{bmatrix} I & X \\ Y & -I \end{bmatrix})$. To this end we seek $Z \in M_k$ such that

$$\begin{bmatrix} I & Z^* \end{bmatrix} \begin{bmatrix} I & X \\ Y & -I \end{bmatrix} \begin{bmatrix} I \\ Z \end{bmatrix} = 0_k; \tag{4}$$

the column space $S$ of $\begin{bmatrix} I \\ Z \end{bmatrix}$ is $k$–dimensional and shows that $0 \in \Lambda_k(\begin{bmatrix} I & X \\ Y & -I \end{bmatrix})$.

Now (4) is equivalent to

$$I + XZ + Z^*Y - Z^*Z = 0_k,$$

so it's enough to show this equation has a solution $Z \in M_k$ for any given $X, Y \in M_k$. We manipulate the equation into a more convenient form by observing that we'd have $XZ + Z^*Y = Z^*Z - I$ Hermitian, ie

$$XZ + Z^*Y = Z^*X^* + Y^*Z,$$

so that $H = (X - Y^*)Z = Z^*(X^* - Y)$ is Hermitian. Then $Z = (X - Y^*)^{-1}H$ puts the equation in the form

$$I + AH + HB - HRH = 0_k,$$

where $A = X(X - Y^*)^{-1}$, $B = (X^* - Y)^{-1}Y$, and $R = (X^* - Y)^{-1}(X - Y^*)^{-1}$; note that $R$ is positive definite, which we write $R > 0$. It is easy to check that $A = B^* + I$, so that with $Q = B + (1/2)I$ the equation to be solved becomes

$$HRH - HQ - Q^*H - I = 0_k \text{ and } H \text{ Hermitian.} \tag{5}$$

After earlier work by Choi, Giesinger, Holbrook, Kribs, and Życzkowski had reached this point, Woerdeman (c. 2007) made the key step by recognizing (5) as a type of Riccati equation that does always have a (Hermitian) solution. This is by no means obvious, but follows from techniques well–known in the CARE community: CARE=continuous algebraic Riccati equation.

In the interests of completeness we present one approach to this result on the solutions of (certain) matrix quadratic equations. We'll show a little more: given $R, C > 0$ and arbitrary $Q \in M_k$ there exists $H \geq 0$ (ie psd $H$) such that

$$HRH - HQ - Q^*H - C = 0_k \tag{6}$$

(in our case $C = I_k$). We construct $H$ as the limit of a decreasing sequence $X_\nu$ of psd matrices.

Choose $X_0 \geq 0_k$ such that $Q - RX_0$ is "stable", ie all of its eigenvalues lie in the $LHP^\circ$ (the open left half–plane); this ensures that $e^{(Q-RX_0)t}$ decays exponentially as $t \to \infty$. A natural choice is $X_0 = rR^{-1}$ for large $r$: then $Q - RX_0 = Q - rI$ and large $r$ shifts the eigenvalues of $Q$ into $LHP^\circ$.

Define inductively Hermitian $X_1, X_2, \ldots$ by:

$$X_{\nu+1}(Q - RX_\nu) + (Q - RX_\nu)^* X_{\nu+1} = -X_\nu R X_\nu - C. \tag{7}$$

We'll see that $X_1 \geq X_2 \geq X_3 \geq \cdots \geq 0$ so that $H = \lim_\nu X_\nu \geq 0$ and (7) in the limit says

$$H(Q - RH) + (Q^* - HR)H = -HRH - C,$$

ie $0 = HRH - HQ - Q^*H - C$, as required.

Helpful fact: if $S$ is stable and $W \geq 0$ then

$$X = \int_0^\infty e^{S^*t} W e^{St}\, dt \quad (\geq 0) \tag{8}$$

is the unique solution of $XS + S^*X = -W$ (the Lyapunov equation). Thus (7) is a good definition of $X_{\nu+1}$ (and $X_{\nu+1} \geq 0$), provided we know that at each step of the induction $Q - RX_\nu$ remains stable.

First we check that (8) does give a solution to the Lyapunov equation: we need

$$\int_0^\infty (S^* e^{S^*t} W e^{St} + e^{S^*t} W e^{St} S)\, dt = -W;$$

the integrand is $d/dt(e^{S^*t} W e^{St})$ ("product rule") so that the integral is

$$e^{S^*\infty} W e^{s\infty} - e^0 W e^0 = -W.$$

Next we check uniqueness of the solution: one instructive method is to consider the linear maps $R_S$ and $L_{S^*}$ defined on $M_k$ by

$$R_S(X) = XS, \quad L_{S^*}(X) = S^* X.$$

We need to know that $R_S + L_{S^*}$ is invertible. It's easy to check that $R_S$ and $S$ have the same spectrum: $\sigma(R_S) = \sigma(S)$ (although the eigenvalues will have different multiplicities), so that $\sigma(R_S) \subseteq LHP^\circ$. similarly,

$$\sigma(L_{S^*}) = \sigma(S^*) = \overline{\sigma(S)} \subseteq LHP^\circ.$$

Moreover, $R_S$ and $L_{S^*}$ commute so that

$$\sigma(R_S + L_{S^*}) \subseteq \sigma(R_S) + \sigma(L_{S^*}) \subseteq LHP^\circ + LHP^\circ = LHP^\circ.$$

In particular, $0 \notin \sigma(R_S + L_{S^*})$.

To see that $Q - RX_{\nu+1}$ remains stable, note that (7) may be written as

$$X_{\nu+1}(Q - RX_{\nu+1}) + (Q - RX_{\nu+1})^* X_{\nu+1} = \tag{9}$$

$$-C - X_{\nu+1} R X_{\nu+1} - (X_{\nu+1} - X_\nu) R (X_{\nu+1} - X_\nu) \quad (< 0).$$

Thus $X_{\nu+1} T + T^* X_{\nu+1} < 0$ with $T = Q - RX_{\nu+1}$. To show $T$ is stable consider $\lambda \in \sigma(T)$ and corresponding eigenvector $x$. Then

$$0 > ((X_{\nu+1} T + T^* X_{\nu+1})x, x) = (\lambda + \overline{\lambda})(X_{\nu+1} x, x)$$

and since $(X_{\nu+1} x, x) \geq 0$ we must have $\mathrm{Re}\lambda < 0$.

Finally, consider (9) with $\nu$ ($\geq 1$) replaced by $\nu - 1$ and subtract (7): we obtain

$$(X_\nu - X_{\nu+1})(Q - RX_\nu) + (Q - TX_\nu)^*(X_\nu - X_{\nu+1}) \quad = \qquad (10)$$
$$-(X_\nu - X_{\nu+1})R(x_\nu - X_{\nu+1}) = -W$$

with $W \geq 0$. We know this has a unique solution $X_\nu - X_{\nu+1} \geq 0$ so indeed

$$X_1 \geq X_2 \geq X_3 \geq \dots \quad (\geq 0)$$

and $H = \lim_{\nu\to\infty} X_\nu$ exists.

Exercise: In an earlier exercise we saw that for a typical $5 \times 5$ unitary $M$, $\Omega_2(M)$ is a filled–in pentagon while $\Delta_2(M)$ is the boundary of that pentagon. In particular the extreme points (vertices of the pentagon) are in $\Lambda_2(M)$. Now that we know $\Lambda_2(M)$ is convex we can conclude that $\Lambda_2(M) = \Omega_2(M)$, ie that the CKŻ conjecture holds for such $M$. Can you find a general proof of the CKŻ conjecture on the same basis, ie by showing that for any normal $M$ the extreme points of $\Omega_k(M)$ are included in $\Delta_k(M)$ (therefore in $\Lambda_k(M)$) and applying convexity?

## 11. Converse to Cauchy interlacing theorem

One ingredient in the Li-Sze representation of $\Lambda_k(M)$ is a beautiful application of the Cauchy interlacing theorem, or rather of its converse. We therefore include this section as preparation.

Cauchy's interlacing theorem along with its converse (see for example Fan and Pall [F–P1957]) says that for real sequences

$$a_1 \leq a_2 \leq \cdots \leq a_n$$

and

$$b_1 \leq b_2 \leq \cdots \leq b_m$$

we have

$$a_k \leq b_k \leq a_{k+(n-m)} \quad (k = 1, 2, \dots, m)$$

$\Longleftrightarrow$

the $a_k$'s are eigenvalues of a Hermitian $n \times n$ matrix $A$ and the $b_k$'s are eigenvalues of an $m \times m$ principal submatrix of $A$.

There are many proofs of $\Longleftarrow$ (Cauchy's interlacing theorem), including Cauchy's own(?); for example, one can use the Courant–Fischer–Weyl min-max principle. Here is an approach based simply on the Intermediate Value Theorem, taken more–or–less from Hwang [H2004]. I like it also because it leads naturally into a proof of $\Longrightarrow$.

By a straightforward induction we need only consider the case where $m = n - 1$.

Exercise: Check out that "straightforward induction".

Thus (permuting the basis if necessary) our Hermitian $A$ has the form $\begin{bmatrix} a & y^* \\ y & B \end{bmatrix}$, where $B$ is an $(n-1) \times (n-1)$ matrix, $y$ is an $(n-1)$–dimensional column vector, and $a$ is a real scalar.

Since $B$ is also Hermitian, the spectral theorem finds a unitary $U$ such that $U^*BU = \mathrm{diag}(b_1, \ldots, b_{n-1}) = D$ where the $b_k$ are the eigenvalues of $B$. Hence

$$(1 \oplus U)^* A (1 \oplus U) = \begin{bmatrix} a & z^* \\ z & D \end{bmatrix}.$$

From this form one easily computes the characteristic polynomial $p(\lambda)$ of $A$:

$$p(\lambda) = (\lambda - a) \prod_{k=1}^{n-1} (\lambda - b_k) - \sum_{k=1}^{n-1} |z_k|^2 \prod_{j \neq k} (\lambda - b_j).$$

Thus $p(b_i)$ has only the single term

$$-|z_i|^2 \prod_{j \neq i} (b_i - b_j).$$

By jiggling (small perturbations) we may assume the $b_j$ are distinct and each $z_j \neq 0$. Thus, in view of the ordering of the $b_j$, the sign of $p(b_i)$ is $(-1)(-1)^{n-1-i} = (-1)^{n-i}$. The alternation of these signs plus the IVT (Intermediate Value Theorem) ensure the existence of zeros $a_k$ of $p$ (ie eigenvalues of $A$) such that

$$a_1 < b_1 < a_2 < b_2 < \cdots < b_{n-1} < a_n,$$

as claimed. For example, $p(b_{n-1})$ has sign $-1$, ie $p(b_{n-1}) < 0$ and since $p(\lambda)$ has leading term $\lambda^n$ it is eventually positive so that the IVT finds $a_n > b_{n-1}$ with $p(a_n) = 0$.

The history of $\Longrightarrow$ (the converse of Cauchy's interlacing theorem) seems obscure; [F-P1957] and Mirsky [M1958] are the earliest references I have found explicitly but there are reports of earlier work, even 19th century work, along these lines ... there is even the suggestion that Cauchy himself may have been aware of the converse. Here we provide a proof using simply the form of $p(\lambda)$ noted above.

Given

$$a_1 \leq b_1 \leq a_2 \leq b_2 \leq \cdots \leq b_{n-1} \leq a_n,$$

we first jiggle the values so as to deal with the more convenient

$$a_1 < b_1 < a_2 < b_2 < \cdots < b_{n-1} < a_n$$

with all $b_k \neq 0$. We construct $A$ in the form $\begin{bmatrix} a & z^* \\ z & D \end{bmatrix}$ with $D = \mathrm{diag}(b_1, \ldots, b_{n-1})$ and $a (\in \mathbb{R}), z$ to be chosen so that the eigenvalues of $A$ are $a_1, \ldots, a_n$. Again the characteristic polynomial of $A$ is

$$p(\lambda) = (\lambda - a) \prod_k (\lambda - b_k) - \sum_k |z_k|^2 \prod_{j \neq k} (\lambda - b_j)$$

and we need this to coincide with

$$q(\lambda) = \prod_{k=1}^{n} (\lambda - a_k).$$

Since both $p$ and $q$ have $\lambda^n$ as leading term, we need only ensure that $p = q$ at $n$ distinct points; let's use $b_1, \ldots, b_{n-1}$ and 0. Now

$$p(b_i) = -|z_i|^2 \prod_{j \neq i} (b_i - b_j)$$

while

$$q(b_i) = \prod_k (b_i - a_k),$$

so we choose $z_i$ so that

$$|z_i|^2 = -\prod_k (b_i - a_k) / \prod_{j \neq i} (b_i - b_j);$$

the given ordering of the $a_k$ and $b_k$ ensures that this expression is positive; count the negative factors:

$$1 + (n - i) + (n - 1 - i) = 2(n - i).$$

Finally, $p(0) = q(0)$ requires

$$-a \prod_k (-b_k) - \sum_k |z_k|^2 \prod_{j \neq k} (-b_j) = \prod_k (-a_k),$$

and since the $b_k$ are nonzero this equation has a (real) solution $a$. QED

12. The Li–Sze half-planes

The Li–Sze half–planes subtly extend (to higher–rank numerical ranges) the following construction for the classical numerical ranges $W(M)$. Given $M \in M_n$ and $\theta \in \mathbb{R}$ define $\mathrm{Re}(e^{-i\theta}M)$ as $(e^{-i\theta}M + e^{i\theta}M^*)/2$. Note that $\mathrm{Re}(e^{-i\theta}M)$ is Hermitian, so that

$$\max\{(\mathrm{Re}(e^{-i\theta}M)u, u) : \|u\| = 1\} = \lambda_1(\theta),$$

the largest eigenvalue of $\mathrm{Re}(e^{-i\theta}M)$. Consider the half–plane

$$H_1(M, \theta) = e^{i\theta}\{z \in \mathbb{C} : \mathrm{Re}\, z \leq \lambda_1(\theta)\}.$$

For all $\theta$, $W(M) \subseteq H_1(M, \theta)$: let $\|u\| = 1$; then $(Mu, u) = e^{i\theta}(e^{-i\theta}Mu, u)$ and $z = (e^{-i\theta}Mu, u)$ satisfies

$$\mathrm{Re}\, z = ((e^{-i\theta}Mu, u) + (u, e^{-i\theta}Mu))/2 = (\mathrm{Re}(e^{-i\theta}M)u, u) \leq \lambda_1(\theta).$$

Let $u_1(\theta)$ be a unit eigenvector of $\mathrm{Re}(e^{-i\theta}M)$ corresponding to $\lambda_1(\theta)$. Then of course $(Mu_1(\theta), u_1(\theta)) \in W(M)$, but note also that $w = (Mu_1(\theta), u_1(\theta))$

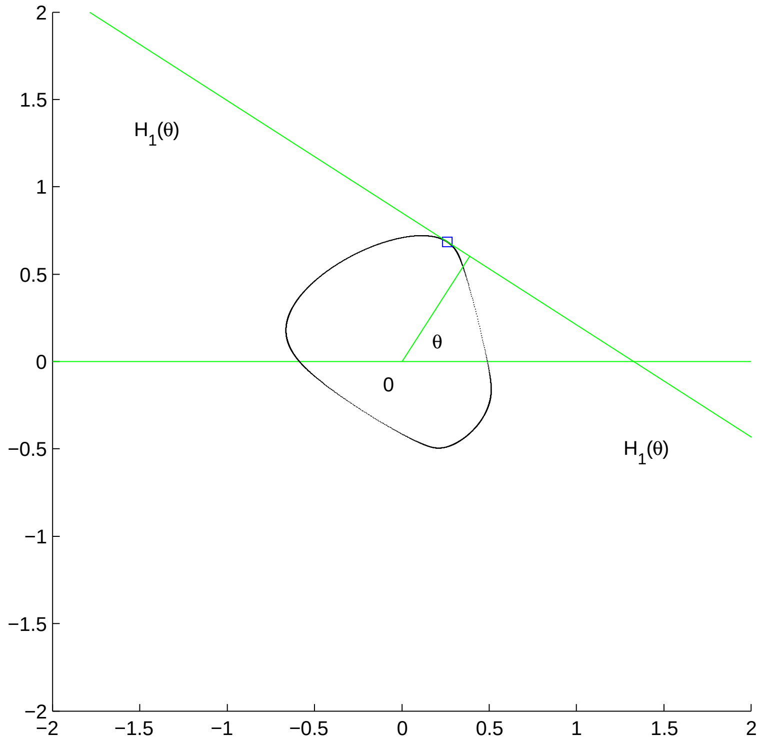


Figure 1: Typical relation between $W(M)$ and $H_1(M, \theta) = H_1(\theta)$, showing also $(Mu_1(\theta), u_1(\theta)) \in W(M) \cap \partial H_1(M, \theta)$

lies in the boundary of $H_1(M, \theta)$, namely the line $e^{i\theta}\{z : \operatorname{Re} z = \lambda_1(\theta)\}$: $z = e^{-i\theta}w = (e^{-i\theta}Mu_1(\theta), u_1(\theta))$ and

$$\operatorname{Re} z = (\operatorname{Re}(e^{-i\theta}M)u_1(\theta), u_1(\theta)) = (\lambda_1(\theta)u_1(\theta), u_1(\theta)) = \lambda_1(\theta).$$

Thus each $H_1(M, \theta)$ is a supporting half–plane for $W(M)$ and the convexity of $W(M)$ (Toeplitz–Hausdorff Theorem) implies that

$$W(M) \quad (= \Lambda_1(M)) \quad = \bigcap_{\theta \in \mathbb{R}} H_1(M, \theta).$$

Moreover, plotting the curve $\{(Mu_1(\theta), u_1(\theta)) : 0 \leq \theta \leq 2\pi\}$ is a well–known method for sketching (the boundary of) $W(M)$. See Figure 1.

Let $\lambda_k(\theta)$ denote the $k$–th largest eigenvalue of $\operatorname{Re}(e^{-i\theta}M)$. Li and Sze found a remarkable extension of the phenomena described above for $W(M)$: for all $M$ and $k$

$$\Lambda_k(M) = \bigcap_{\theta \in \mathbb{R}} H_k(M, \theta), \tag{11}$$

where $H_k(M, \theta)$ is the half–plane $e^{i\theta}\{z \in \mathbb{C} : \operatorname{Re} z \leq \lambda_k(\theta)\}$.

Li and Sze first proved (11) for normal $M$ and it is the purpose of this section to explain their argument for normal $M$, showing how it relates to the Cauchy interlacing theorem (converse) discussed in section 11.

Note first that (11) for normal $M \in M_n$ implies the CKŻ conjecture that

$$\Lambda_k(M) = \Omega_k(M) = \bigcap \{\operatorname{conv}\{\lambda_j : j \in J\} : \#(J) = n - k + 1\},$$

where $\lambda_j$ are the eigenvalues of (normal) $M$. We have seen that $\Lambda_k(M) \subseteq \Omega_k(M)$ on general grounds (recall section 9). On the other hand, (11) tells us that for any $\lambda \notin \Lambda_K(M)$ there is some $\theta$ such that $\lambda \notin H_k(M, \theta)$. Since $M$ is normal each of the eigenvalues

$$\lambda_k(\theta), \lambda_{k+1}(\theta), \ldots, \lambda_n(\theta)$$

of $\operatorname{Re}(e^{-i\theta} M)$ is $\operatorname{Re}(e^{-i\theta}\lambda_j)$ for some eigenvalue $\lambda_j$ of $M$. Hence there is a set $J$ of $n - k + 1$ indices such that for each $j \in J$ we have $\operatorname{Re}(e^{-i\theta}\lambda_j) \leq \lambda_k(\theta)$, ie $\lambda_j \in H_k(M, \theta)$. Thus $\operatorname{conv}\{\lambda_j : j \in J\} \subseteq H_k(M, \theta)$ and

$$\lambda \notin \operatorname{conv}\{\lambda_j : j \in J\} \supseteq \Omega_k(M).$$

Exercise: In the discussion above we have used the common interpretation of $\operatorname{Re} X$ as the Hermitian $(X + X^*)/2$. Another possible interpretation would be "elementwise", as $[\operatorname{Re} x_{ij}]$ (not usually Hermitian). When do these two interpretations agree?

We have seen that "easy" proofs of the CKŻ conjecture fail when $\Delta_k(M) \neq \Omega_k(M)$. Perhaps the simplest "hard" case of the CKŻ conjecture: $M$ is a $5 \times 5$ unitary $U$ and $k = 2$; here

$$\Omega_2(U) = \bigcap_{\#(J)=4} \operatorname{conv}\{\lambda_j : j \in J\}$$

is a filled–in pentagon, while $\Delta_2(U)$ is just its boundary. We have

$$\Delta_2(U) \subseteq \Lambda_2(U) \subseteq \Omega_2(U)$$

but in order to show $\Lambda_2(U) = \Omega_2(U)$ (the CKŻ conjecture for this case) we must find interior points in $\Lambda_2(U)$.

There are at least three quite different ways to achieve this; none of them is straightforward.
1. [CHKŻ] gave a topological argument (involving winding numbers etc) filling the pentagon with certain specially contrived elements of $\Lambda_2(U)$ (seems hard to generalize, although it suffices for other cases where $n = 5k/2$).
2. [CGHK] + [Woerdeman] proved convexity for all $\Lambda_k(M)$ (see section 10). Given this, for situations like the pentagon we only need to capture the boundary – or even just vertices of $\Lambda_k$ (as suggested in an earlier exercise, it is likely that this approach could be extended to cover all normal $M$).
3. Li and Sze half–planes, illustrated below, provide a very general approach (ultimately extending to all $M$).

We illustrate the Li–Sze approach for the pentagon $\Omega_2(U)$. It will be rather clear that it extends to all normal $M$. Using other powerful ideas (congruence canonical forms, etc) Li and Sze obtain the corresponding result also for non–normal $M$ (see sections 13 and 14).

Note that it is clear for any $M \in M_n$ that $\Lambda_k(M) \subseteq H_k(M, \theta)$ for each $\theta$, since $W(P_L M|_L) \subseteq H_k(M, \theta)$ where $L$ is the $n-k+1$–dimensional subspace spanned by eigenvectors of $\mathrm{Re}(e^{-i\theta}M)$ corresponding to $\lambda_k(\theta), \ldots, \lambda_n(\theta)$ and we may invoke the second proposition of section 9.

To show that $\lambda \in \bigcap_\theta H_2(U)$ implies $\lambda \in \Lambda_2(U)$, note that by considering $N = U - \lambda I_5$ in place of $U$ we may assume $\lambda = 0$ and although the $\lambda_j$ will have shifted we still have at least two $\lambda_j$ in every open half–plane determined by a line through 0: otherwise $0 \notin H_2(N, \theta)$ for a $\theta$ perpendicular to the line. See Figure 2. Rotate so that 3 eigenvalues are in the right half–plane and two in the left. With $N$ in diagonal form let diagonal $R$ have eigenvalues $r_k > 0$ such that $Q = RNR$ has eigenvalues $1 + ia_j$ and $-1 + i(-b_j)$ as shown in Figure 3; eg

$$r_1 = \sqrt{\frac{\sqrt{1 + a_1^2}}{|\lambda_1|}}.$$

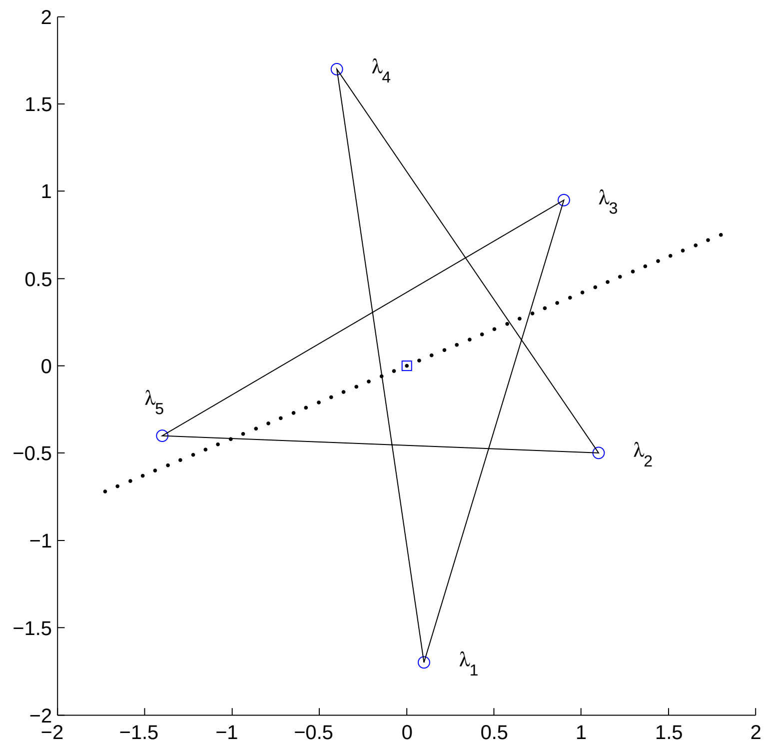


Figure 2: Typical line through 0 in the pentacle determined by the eigenvalues of $N$

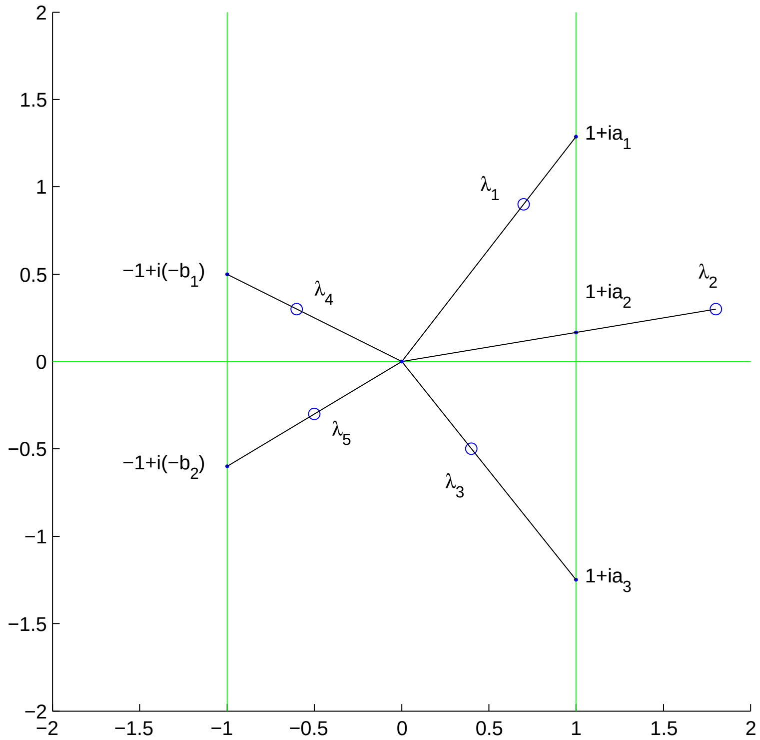


Figure 3: Eigenvalues (circles) for a typical $5 \times 5$ normal $N$ versus those for the modified $Q = RNR$

Note that $a_1 < a_2 < a_3$ and $b_1 < b_2$. By virtue of the lemma of section 10 we need only show that $0 \in \Lambda_2(Q)$. Now $Q = (I_3 + iA) \oplus (-I_2 + i(-B))$ where $A, B$ are Hermitian with eigenvalues $a_1 < a_2 < a_3$ and $b_1 < b_2$. We claim that the interlacing condition

$$a_1 \leq b_1 \leq a_2 \leq b_2 \leq a_3$$

is satisfied.

Suppose for example that $b_2 < a_2$: Figure 4 illustrates the resulting contradiction.

What if $b_2 > a_3$? Figure 5 illustrates the contradiction that would ensue.

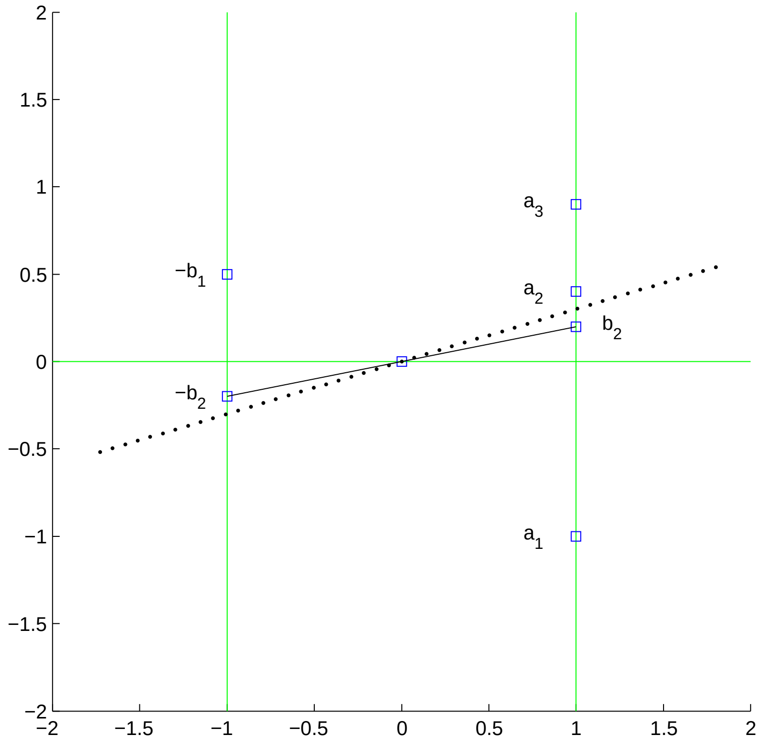


Figure 4: Only $1 + ia_1$ (and therefore $\lambda_1$) is below the dotted line: contradiction

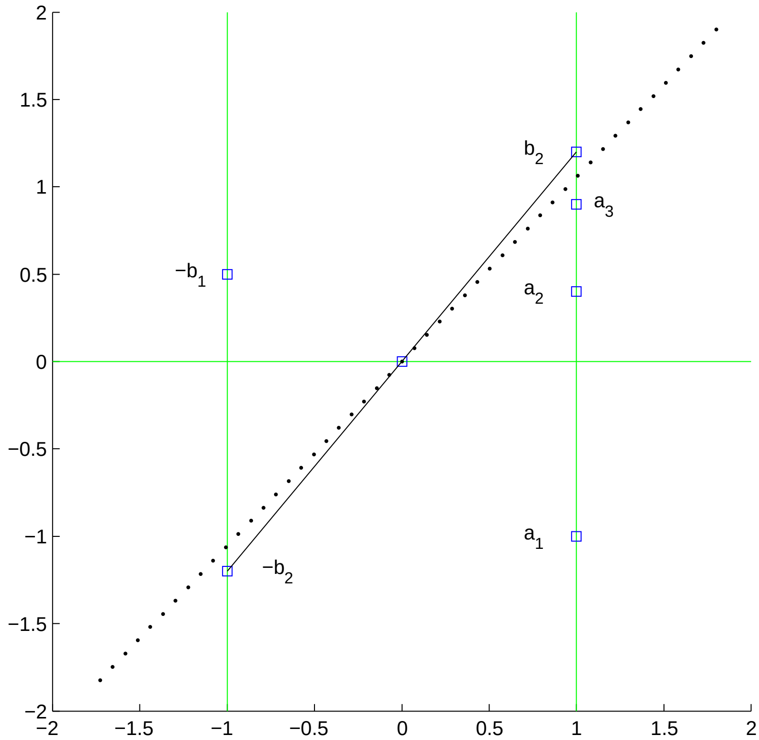


Figure 5: Only $-1 + i(-b_1)$ (and therefore $\lambda_4$) is above the dotted line: contradiction

Thus the converse to Cauchy's interlacing theorem finds a unitary $U$ (forgive the change in notation) such that $U^*AU = \begin{bmatrix} a & y \\ y^* & B \end{bmatrix}$, so that

$$(U \oplus I_2)^* Q (U \oplus I_2) = \begin{bmatrix} 1 + ia & y \\ y^* & 1 + iB \end{bmatrix} \oplus (-I_2 + i(-B))$$

has $4 \times 4$ submatrix $T = \begin{bmatrix} C & 0 \\ 0 & -C \end{bmatrix}$ where $C = I + iB$ is normal with eigenvalues $\gamma_1, \gamma_2$. Finally, $0 \in \Lambda_2(T)$ because $0 \in \Delta_2(T)$:

$$0 \in \operatorname{conv}\{\gamma_1, -\gamma_1\} \cap \{\gamma_2, -\gamma_2\},$$

and these index sets are disjoint. QED

sections to be completed:

13. Canonical forms

14. The Li-Sze representation of higher–rank numerical ranges

15. When are the ranges nonempty?

16. Joint numerical ranges

17. Normal compressions

References: (incomplete)


[BS1967] C. A. Berger and J. G. Stampfli, Mapping theorems for the numerical radius, Amer. J. Math. 89, 1047–1055, 1967

[D1971] C. Davis, The Toeplitz–Hausdorff theorem explained, Canadian Math. Bull. 14, 245–246, 1971

[F–P1957] Ky Fan and Gordon Pall, Imbedding conditions for Hermitian and normal matrices, Canadian J. Math. 9, 298–304, 1957

[G–R1997] K. E. Gustafson and D. K. M. Rao, Numerical Range, Springer 1997

[H2004] Suk–Geun Hwang, Cauchy's interlace theorem for eigenvalues of Hermitian matrices, Amer. Math. Monthly 111(2), 157–159, 2004(Feb)

[K1965] T. Kato, Some mapping theorems for the numerical range, Proc. Japan Acad. 41, 652–655, 1965

[M1958] L. Mirsky, Matrices with prescribed characteristic roots and diagonal elements, J. London Math. Soc. 33, 14–21, 1958

[P2008] D. Petz, Quantum Information Theory and Quantum Statistics, Springer 2008