On the spectral properties of certain matrices generated by involutive automorphisms

This talk is based on a work in collaboration with A. Iserles.

Antonella Zanna
University of Bergen, Norway

email: anto@ii.uib.no
http://www.ii.uib.no/~anto

Motivations

- Integration methods using exponentials in GI need fast algorithms that approximate the matrix exponential.
- Exact computation is not an issue – but it is crucial that the exponential approximation is in the Lie group.
- Matrix exponential $\exp(tZ)$ computed by GPDs: based on a splitting

$$Z = P + K$$

where P and K are obtained by projecting A onto the $+$ and $-$ eigenspaces of an involutive automorphism in the algebra $\mathfrak{g}$.

- Numerical stability: If Z is 'nice', it is important that also K and P are 'nice', otherwise we might generate very large matrices in the course of computation.

Relate the eigenvalues of the matrices Z with those of P and K

Generalized polar decompositions

Ingredients:

- A Lie group G
- An involutive automorphism σ

What we can do with them:

- We can factorize

$$z = xy \qquad \text{GPD}$$

where

$$\sigma(x) = x^{-1}$$

and

$$\sigma(y) = y$$

where x, y are appropriate group elements which are determined by z and σ.

In particular, if $z = \exp(tZ)$, it is true that

$$\exp(tZ) = \exp(X(t)) \exp(Y(t))$$

and for t sufficiently small, the functions $X(t)$ and $Y(t)$ are uniquely determined.

$$
\begin{aligned}
X &= Pt - \tfrac{1}{2}[P,K]t^2 - \tfrac{1}{6}[K,[P,K]]t^3 \\
&\quad + \Big(\tfrac{1}{24}[P,[P,[P,K]]] - \tfrac{1}{24}[K,[K,[P,K]]]\Big)t^4 \\
&\quad + \Big(\tfrac{7}{360}[K,[P,[P,[P,K]]]] - \tfrac{1}{120}[K,[K,[K,[P,K]]]] - \tfrac{1}{180}[[P,K],[P,[P,K]]]\Big)t^5 \\
&\quad + \mathcal{O}(t^6),
\end{aligned}
$$

$$
\begin{aligned}
Y &= Kt - \tfrac{1}{12}[P,[P,K]]t^3 + \Big(\tfrac{1}{120}[P,[P,[P,[P,K]]]] \\
&\qquad + \tfrac{1}{720}[K,[K,[P,[P,K]]]] - \tfrac{1}{240}[[P,K],[K,[P,K]]]\Big)t^5 + \mathcal{O}(t^7).
\end{aligned}
$$

In general, all the terms in the expansion of $X(t)$ and $Y(t)$ are obtained by explicit recurrence relations in terms of P and K.

At the algebra level

Assume $Z \in \mathfrak{g}$. A Lie-algebra automorphism is a one-to-one map $\mathrm{d}\sigma : \mathfrak{g} \times \mathfrak{g}$ such that

$$\mathrm{d}\sigma([A, B]) = [\,\mathrm{d}\sigma(A), \mathrm{d}\sigma(B)], \quad A, B \in \mathfrak{g},$$

moreove it is involutive if $\mathrm{d}\sigma^2 = \mathrm{id}$.
We are interested in *inner automorphisms*, i.e. automorphisms of the form

$$\mathrm{d}\sigma(Z) = HZH^{-1},$$

for some involutive matrix H ($H^2 = I$).
For reasons of stability, we consider involutory matrices $H \in \mathbb{C}$ that are unitary, i.e. $H^{-1} = H^T$.

Theorem 1 *Every involution in* $\mathrm{U}(n)$ *is of the type*

$$H = I - 2\sum_{k=1}^{s} \mathbf{u}_k \mathbf{u}_k^*,$$

where $\{\mathbf{u}_1, \dots, \mathbf{u}_s\}$ *is an orthonormal basis of a subspace of* $\mathbb{C}^n$*, for* $s = 1, 2, \dots, n$*.*

Proof. H is unitary hence normal, therefore the matrix $I - H$ is normal and has a full set of orthonormal eigenvectors $\mathbf{u}_1, \ldots, \mathbf{u}_n$,

$$I - H = \sum_{k=1}^{n} \alpha_k \mathbf{u}_k \mathbf{u}_k^*,$$

the α_k being the respective eigenvalues. Hence, $H = I - \sum_{k=1}^{n} \alpha_k \mathbf{u}_k \mathbf{u}_k^*$.
From $H^2 = I$,

$$\begin{aligned} O = H^2 - I &= -2 \sum_{k=1}^{n} \alpha_k \mathbf{u}_k \mathbf{u}_k^* + \sum_{k,l=1}^{n} \alpha_k \alpha_j \mathbf{u}_k \mathbf{u}_k^* \mathbf{u}_j \mathbf{u}_j^* \\ &= \sum_{k=1}^{n} (\alpha_k - 2) \alpha_k \mathbf{u}_k \mathbf{u}_k^* \end{aligned}$$

because of the orthonormality of the eigenvectors $\mathbf{u}_k$. Since the $\mathbf{u}_k$ are also linearly independent, it must be

$$\begin{aligned} \alpha_j &= 2 \text{ for } j = 1, \ldots, s, \\ \alpha_l &= 0 \text{ for } l = s+1, \ldots, n. \end{aligned}$$

□

Some background theory: The field of values of a matrix

Let $A \in M_n[\mathbb{C}]$. The *field of values* or *numerical range* of a matrix A is the subset of the complex plane

$$F(A) = \{\mathbf{x}^* A\mathbf{x} \ : \ \mathbf{x}^*\mathbf{x} = 1\}.$$

Some properties:

- For every A, $F(A)$ is a convex set.
- For every A,

  $$\sigma(A) \subset F(A).$$

  From $A\mathbf{x} = \lambda\mathbf{x}$ we have $\lambda = \mathbf{x}^* A\mathbf{x} \in F(A)$.
- If D diagonal,

  $$F(D) = \text{conv } \sigma(D)$$

  Trivially, conv $\sigma(D) \subseteq F(D)$. If $\kappa \in F(D)$ then $\kappa = \mathbf{x}^* D\mathbf{x}$ with $\mathbf{x} = \sum_{k=1}^n x_k \mathbf{e}_k$ and $\sum_{k=1}^n |x_k|^2 = 1$.

  Then, $\kappa = \left(\sum_{k=1}^n \bar{x_k}\mathbf{e}_k^T\right) D \left(\sum_{j=1}^n x_j \mathbf{e}_j\right) = \sum_{k=1}^n |x_k|^2 d_{k,k} \in \text{conv } \sigma(D)$.
- If U is a unitary matrix,

  $$F(UAU^*) = F(A)$$
- If A is normal, i.e. $A^*A = AA^*$, then

  $$F(A) = \text{conv } \sigma(A)$$

- If $A = B + C$, then

$$\sigma(A) \subseteq F(A) \subseteq F(B) + F(C)$$

(subadditive property)

Moreover, if B, C are normal with eigenvalues $\{\beta_i\}$ and $\{\gamma_i\}$,

$$\sigma(A) \subseteq \text{conv } \sigma(B) + \text{conv } \sigma(C) = \text{conv } \{(\beta_i + \gamma_j),\ i, j = 1, 2, \ldots, n\}$$

On the spectrum of the matrix P

Recall that

$$P = \tfrac{1}{2}(Z - HZH).$$

Proposition 2 *If κ is an eigenvalue of P with eigenvector $\mathbf{y}$, then $-\kappa$ is also eigenvalue of P with eigenvector $H\mathbf{y}$.*

Proof. Since by construction $-P = \mathrm{d}\sigma(P) = HPH$ and H is an involution, it follows that $PH = -HP$.
We have

$$P(H\mathbf{y}) = -HP\mathbf{y} = -\kappa(H\mathbf{y}).$$

□

Proposition 3 *Assume that $H = I - 2UU^*$, where $U = [\mathbf{u}_1, \ldots \mathbf{u}_s]$. Then the matrix P has rank at most $2s$ and if Z is normal with eigenvalues λ_i, then*

$$\sigma(P) \subseteq \tfrac{1}{2}\textit{conv}\{(\lambda_i - \lambda_j),\ i, j = 1, \ldots, n\}$$

and, consequently,

$$\rho(P) \leq \frac{1}{2}\textit{diam conv}\,\sigma(Z).$$

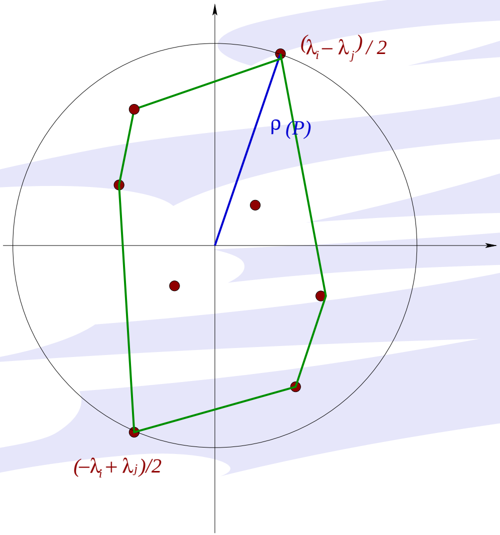
Proof. We verify the last two assertions.
Let us study $\sigma(P)$.

$$P = \frac{1}{2}(Z - HZH)$$

Recall: if B and C are normal, with eigenvalues β_i and γ_i respectively, then

$$\sigma(B + C) \subseteq \text{conv}\{(\beta_i + \gamma_j), i, j = 1, \ldots, n\}.$$

$$\sigma(P) \subseteq \tfrac{1}{2}\text{conv}\{(\lambda_i - \lambda_j),\ i, j = 1, \ldots, n\}$$



Hence

$$\begin{aligned} \rho(P) &\leq \frac{1}{2} \max |\lambda_i - \lambda_j| \\ &= \frac{1}{2} \text{diam conv } \sigma(Z) \end{aligned}$$

□

On the spectrum of the matrix K

We denote by $\mu_1, \ldots, \mu_n$ the eigenvalues of K with eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$. By construction,

$$K = \frac{1}{2}(Z + \mathrm{d}\sigma(Z)) = \frac{1}{2}(Z + HZH).$$

Lemma 4 *Assume that Z is normal. Then,*

$$\sigma(K) \subset \textit{conv}\, \sigma(Z).$$

Proof. First note that, if $B = C + D$ then

$$F(B) \subseteq F(C) + F(D), \qquad F(\alpha B) = \alpha F(B), \quad \alpha \in \mathbb{C}.$$

We have:

$$\begin{aligned} \sigma(K) \subseteq F(K) &\subseteq \frac{1}{2} F(Z + HZH) \\ &\subseteq \frac{1}{2}\left(F(Z) + F(HZH)\right) = \frac{1}{2}\left(\text{conv}\ \sigma(Z) + \text{conv}\ \sigma(Z)\right) \\ &= \text{conv}\ \sigma(Z). \end{aligned}$$

(Note that K need not be normal). □

Theorem 5 *Assume that Z is normal and that $H = I - 2\sum_{k=1}^{s} \mathbf{u}_k \mathbf{u}_k^*$, where $\mathbf{u}_1, \ldots, \mathbf{u}_s$ is a set of orthonormal complex vectors. Then,*

$$\|Z - K\|_F \leq \frac{\sqrt{2s}}{2} \textit{diam conv}\, \sigma(Z)$$

where diam Ω is the diameter of the set Ω.

Proof. From $Z = K + P$ we deduce that

$$\|Z - K\|_F^2 = \|P\|_F^2 = \sum_{i=1}^{2s} |\kappa_i|^2 = 2\sum_{k=1}^{s} |\kappa_i|^2,$$

where $\kappa_1, \ldots, \kappa_s, \kappa_{s+1} = -\kappa_1, \ldots, \kappa_{2s} = -\kappa_s$ are the eigenvalues of P.

$$\|Z - K\|_F^2 \leq 2s\rho(P)^2, \qquad \rho(P) = \max|\kappa_i|.$$

From the characterization of P,

$$\rho(P) \leq \frac{1}{2} \text{diam conv} \sigma(Z),$$

hence

$$\|Z - K\|_F^2 \leq \frac{s}{2} (\text{diam conv} \sigma(Z))^2,$$

from which the result follows by taking the square root of both sides. □

Remark: The spectra of P and K depend on the distribution of the eigenvalues of Z on the complex plane.

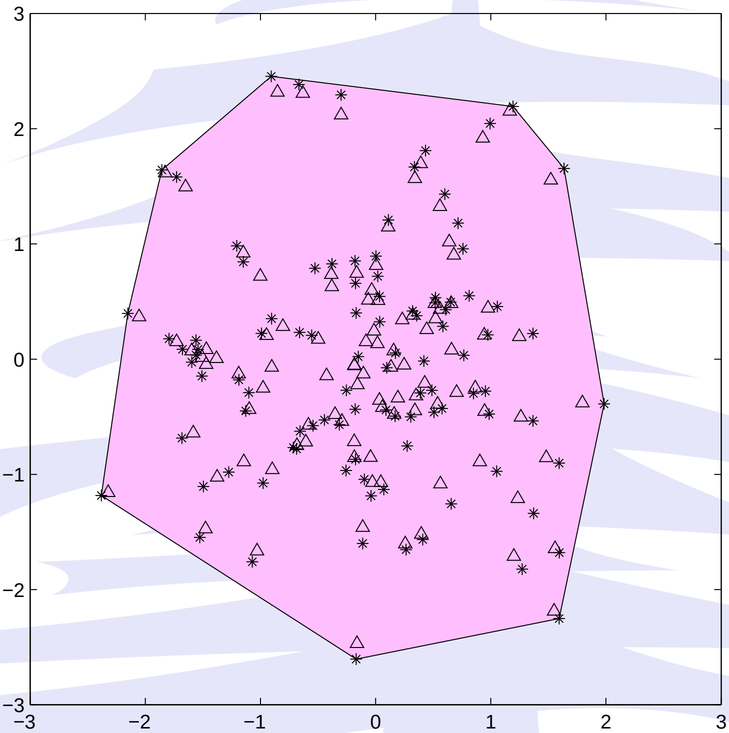


Figure 1: Eigenvalues of K (triangles) and of Z (asterisks) in $M_{100}[\mathbb{C}]$ for $s = 1$.

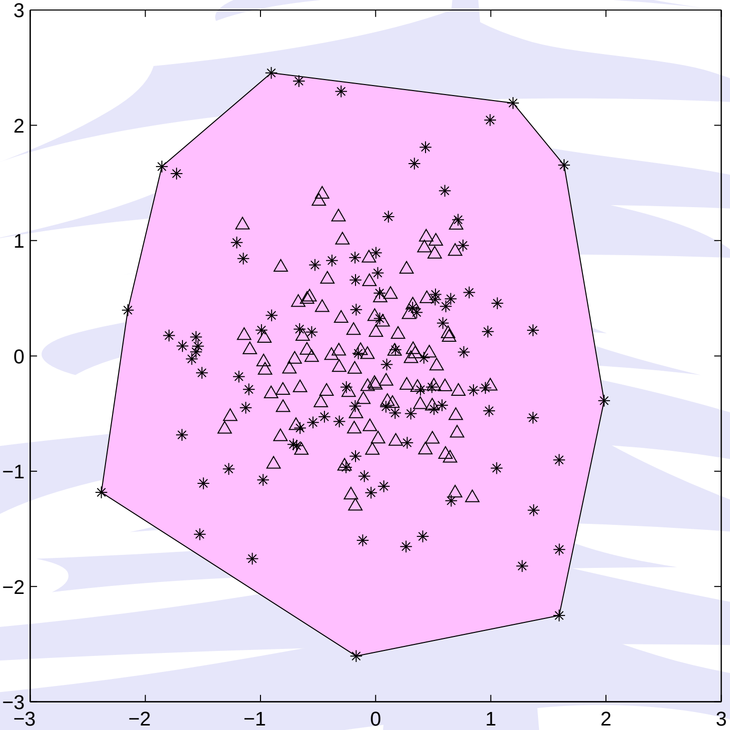


Figure 2: Eigenvalues of K (triangles) and of Z (asterisks) in $M_{100}[\mathbb{C}]$ for $s = 50$.

Back
Close

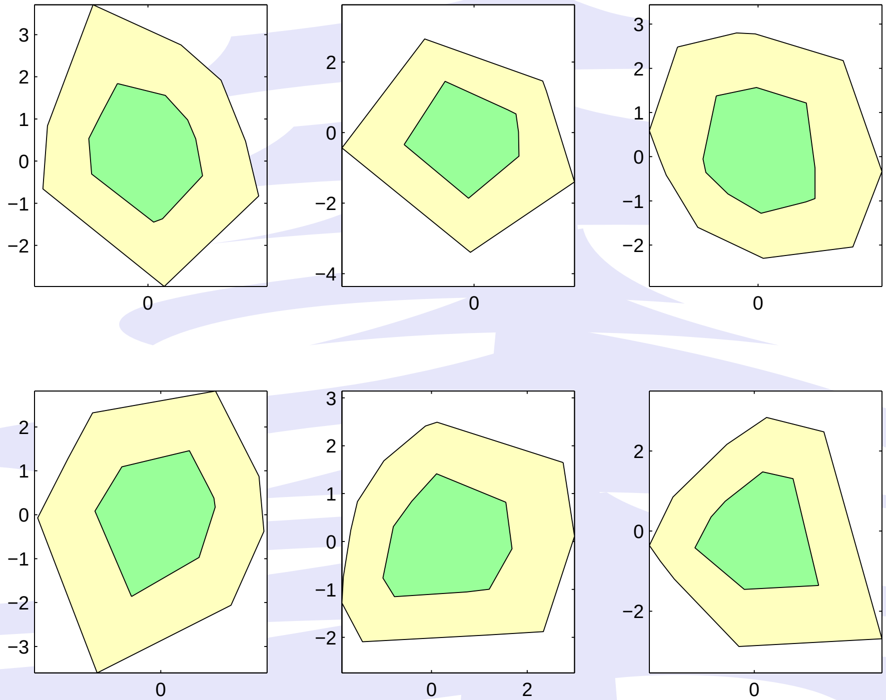


Figure 3: Convex hulls of $\sigma(Z)$ and $\sigma(K)$ for various normal Z, $n = 200$, $s = 100$.

The eigenvalues of K as zeros of a rational function

Assume that Z is normal and denote by λ_k the eigenvalues of Z and by μ_k the eigenvalues of K.

Theorem 6 *Either $\mu \in \sigma(Z)$ or it is a zero of the rational function*

$$\psi(x) = \sum_{k=1}^{n} \frac{|\zeta_k|^2}{\lambda_k - x},$$

where $\sum_{k=1}^{n} \zeta_k \mathbf{x}_k = \sum_{l=1}^{s} (\mathbf{u}_l^* A \mathbf{v}) \mathbf{u}_l$, where $K\mathbf{v} = \mu \mathbf{v}$ and the $\mathbf{x}_k$ are eigenvectors of Z.

Interlacing properties of the eigenvalues of K

Assume that Z is Hermitian with distinct eigenvalues and let $s = 1$, thus $\mathbf{u}_1 = \mathbf{u}$.
Assume that $\mathbf{u}$ is not orthogonal to any eigenvalue of Z.

$$\psi(x) = \sum_{k=1}^{n} \frac{|\zeta_k|^2}{\lambda_k - x},$$

Then function ψ is of type $(n-1)/n$ therefore it has precisely $n-1$ real zeros.
Observe that these zeros occur in intervals of the type $(\lambda_k, \lambda_{k+1})$ because of the change of sign, and since K has n eigenvalues, a trivial counting argument tells us that there is one such interval in which there are two μ_k.
In other words, $\exists p \in \{1, 2, \ldots, n\}$ such that

$$\begin{aligned} \mu_k &\in (\lambda_k, \lambda_{k+1}), \qquad k = 1, 2, \ldots, p, \\ \mu_k &\in (\lambda_{k-1}, \lambda_k), \qquad k = p+1, p+2, \ldots, n. \end{aligned}$$

This property is also known as *interlacing property of the eigenvalues of Hermitian matrices* or *Weyl's theorem* [1].

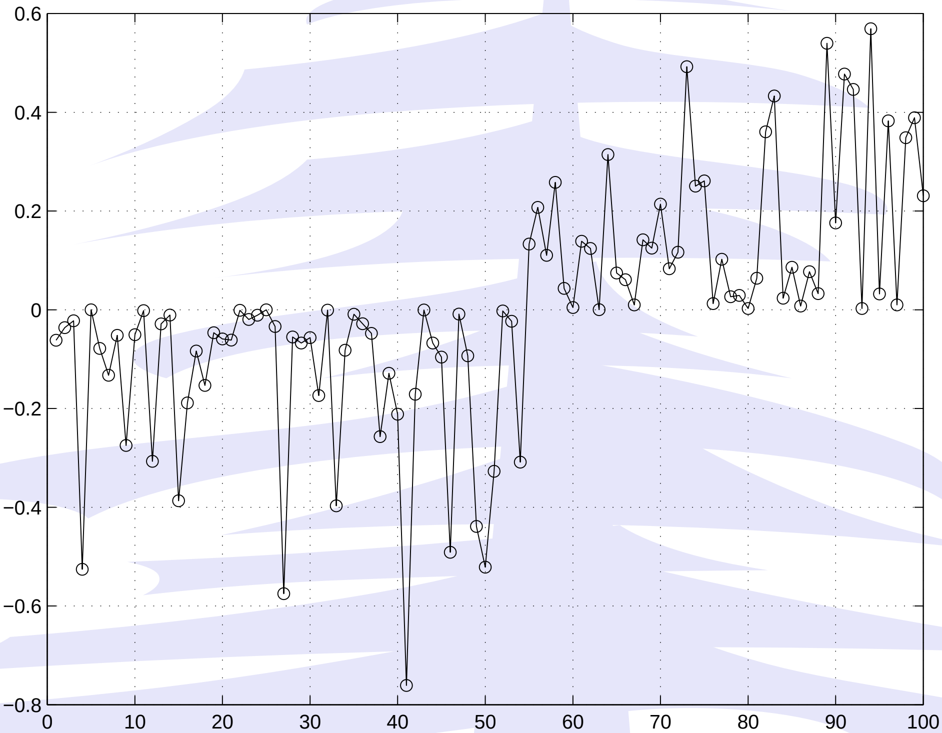


Figure 4: The difference $\lambda - \mu$ for a random Hermitian matrix Z of dimension $n = 100$

The non-normal case

The results need not be true in the case when Z is not normal. In fact, it is possible to find non-normal Z and U such that $\sigma(K) \notin \operatorname{conv} \sigma(Z)$.

References

[1] Horn and Johnson, Matrix Analysis, CUP, 1985
[2] Horn and Johnson, Topics in Matrix Analysis, CUP 1991
[3] Iserles and Zanna, In preparation, 2003
[4] Munthe-Kaas, Quispel and Zanna, Generalized polar decompositions on Lie groups, J. FoCM, 2001.