Where in the complex plane does a matrix live?

(A question of L. N. Trefethen)

Connections Between Matrix Theory and Complex Analysis

What can eigenvalues do?

- If A is **normal** (e.g., real symmetric) or **near normal** (well-conditioned eigenvectors) then eigenvalues describe behavior in spectral norm perfectly or almost perfectly:

$$\|f(A)\| \approx \max_{\lambda \in \sigma(A)} |f(\lambda)|.$$

- Even if A is highly **nonnormal** (e.g., not diagonalizable, or diagonalizable but with eigenvectors that are almost linearly dependent), eigenvalues determine the *asymptotic* behavior of many functions of A:

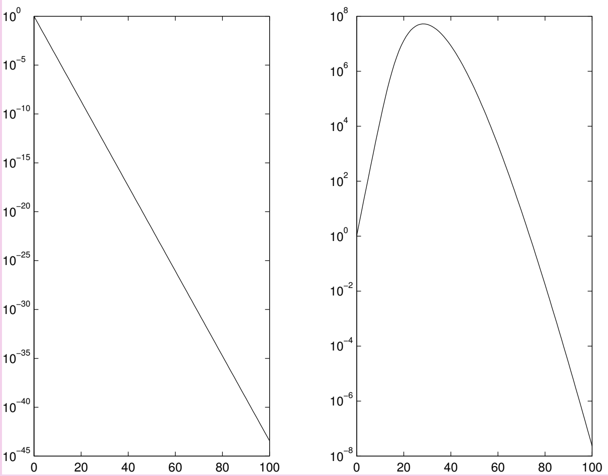
$$\|A^k\| \to 0 \quad \text{as } k \to \infty \quad \text{iff } \rho(A) < 1.$$

$$\|e^{tA}\| \to 0 \quad \text{as } t \to \infty \quad \text{iff } \mathrm{Re}(\sigma(A)) < 0.$$

What can eigenvalues NOT do?

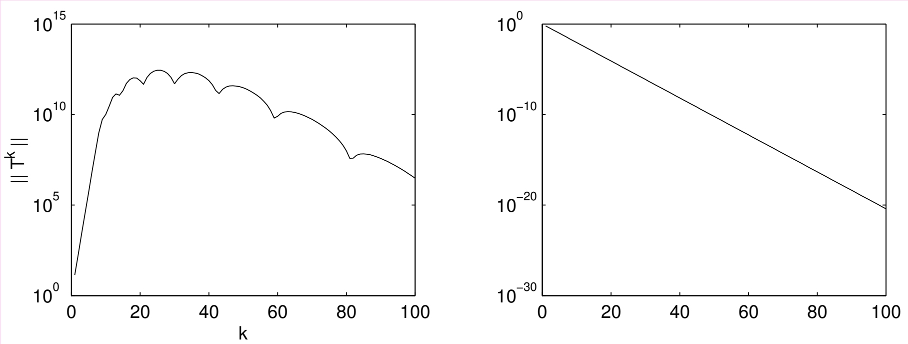
- e^{tA}: Determines the stability of $y' = Ay$.

  $\lim_{t\to\infty} \|e^{tA}\| = 0$ if and only if the eigenvalues of A have negative real parts. But eigenvalues alone cannot distinguish:



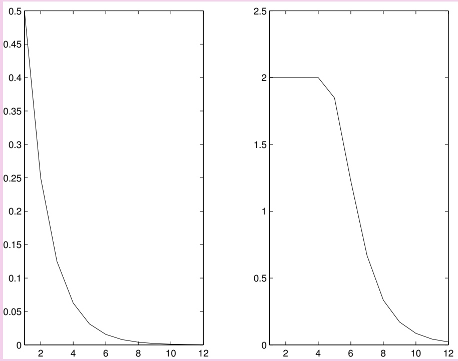
- A^k: Determines stability of finite difference schemes; determines the convergence of stationary iterative methods for linear systems.

  $\lim_{k\to\infty} \|A^k\| = 0$ if and only if $\rho(A) < 1$. But eigenvalues alone cannot distinguish:



- A^k: Markov chains.

  y_0 = initial state; $A^k y_0$ = state after k steps. $A^k y_0 \to v$ = eigenvector corresponding to eigenvalue 1. For k large, convergence rate is determined by second largest eigenvalue. But eigenvalues cannot distinguish:



- $\min_{\substack{p \in \mathcal{P}_k \\ p(0)=1}} \|p(A)\|$: Residual norm in ideal GMRES.

  Any possible convergence behavior of GMRES can be attained with a matrix having any given eigenvalues. (G., Pták, Strakoš, '96)

Given an n by n matrix A, find a set $S \subset \mathbf{C}$ that can be associated with A to give more information than the spectrum alone can provide about the 2-norm of functions of A.

- Field of values:

$$W(A) = \{\langle Aq, q\rangle : \ \langle q, q\rangle = 1\}.$$

- ϵ-pseudospectrum:

$$\sigma_\epsilon(A) = \{z \in \mathbf{C} : \ z \text{ is an eigenvalue of } A + E \text{ for some } E \text{ with } \|E\| < \epsilon\}.$$

- Polynomial numerical hull of degree k:

$$\mathcal{H}_k(A) = \{z \in \mathbf{C} : \ \|p(A)\| \geq |p(z)| \ \forall p \in \mathcal{P}_k\}.$$

Find a set S and scalars m and M with M/m of moderate size such that for all polynomials (or analytic functions) p:

$$m \cdot \sup_{z \in S} |p(z)| \le \|p(A)\| \le M \cdot \sup_{z \in S} |p(z)|.$$

- $S = \sigma(A)$, $m = 1$, $M = \kappa(V)$.

  If A is normal then $m = M = 1$, but if A is nonnormal then $\kappa(V)$ may be huge. Moreover, if columns of V have norm 1, then $\kappa(V)$ is close to smallest value that can be used for M.

- If A is nonnormal, might want S to contain more than the spectrum. BUT...

  If S contains more than $\sigma(A)$, must take $m = 0$ since if p is minimal polynomial of A then $p(A) = 0$ but $p(z) = 0$ only if $z \in \sigma(A)$.

- How to modify the problem?

$$m \cdot \sup_{z \in S} |p_{r-1}(z)| \leq \|p(A)\| \leq M \cdot \sup_{z \in S} |p(z)|$$

- If degree of minimal polynomial is r, then any $p(A) = p_{r-1}(A)$ for a certain $(r-1)$st degree polynomial – the one that matches p at the eigenvalues, and whose derivatives of order up through $t-1$ match those of p at an eigenvalue corresponding to a t by t Jordan block.

- The largest set S where above holds with $m = 1$ is called the **polynomial numerical hull of degree $\mathbf{r-1}$**. In general, however, we do not know good values for M ($<< \kappa(V)$).

Given a set S, for each p consider

$$\inf\{\|f\|_{\mathcal{L}^\infty(S)} : \ f(A) = p(A)\}. \qquad (*)$$

Find scalars m and M such that for all p:

$$m \cdot (*) \leq \|p(A)\| \leq M \cdot (*).$$

- $f(A) = p(A)$ if $f(z) = p_{r-1}(z) + \chi(z)h(z)$ for some $h \in H^\infty(S)$. Here χ is the minimal polynomial (of degree r) and p_{r-1} is the polynomial of degree $r-1$ satisfying $p_{r-1}(A) = p(A)$.

- (*) is a **Pick-Nevanlinna interpolation problem**.

Given $S \subset \mathbf{C}$, $\lambda_1, \ldots, \lambda_n \in S$, and $w_1, \ldots, w_n$, find

$$\inf\{\|f\|_{\mathcal{L}^\infty(S)} : \ f(\lambda_j) = w_j, \ j = 1, \ldots, n\}.$$

- If S is the open unit disk, then infimum is achieved by a function $\tilde{f}$ that is a scalar multiple of a finite **Blaschke product**:

$$\begin{aligned} \tilde{f}(z) &= \mu \prod_{k=0}^{n-1} \frac{z - \alpha_k}{1 - \bar{\alpha}_k z}, \quad |\alpha_k| < 1 \\ &= \mu \frac{\gamma_0 + \gamma_1 z + \ldots + \gamma_{n-1} z^{n-1}}{\bar{\gamma}_{n-1} + \bar{\gamma}_{n-2} z + \ldots + \bar{\gamma}_0 z^{n-1}}. \end{aligned}$$

- Using second representation, Glader and Lindström showed how to compute $\tilde{f}$ and $\|\tilde{f}\|_{\mathcal{L}^\infty(\mathcal{D})}$ by solving a simple eigenvalue problem.

Given $S \subset \mathbf{C}$, $\lambda_1, \ldots, \lambda_n \in S$, and $w_1, \ldots, w_n$, find

$$\inf\{\|f\|_{\mathcal{L}^\infty(S)} : \ f(\lambda_j) = w_j, \ j = 1, \ldots, n\}.$$

- If S is a simply-connected open set, it can be mapped onto the open unit disk $\mathcal{D}$ via a one-to-one analytic mapping g.

$$\inf\{\|F\|_{\mathcal{L}^\infty(S)} : \ F(\lambda_j) = w_j\} =$$

$$\inf\{\|f \circ g\|_{\mathcal{L}^\infty(S)} : \ (f \circ g)(\lambda_j) = w_j\} =$$

$$\inf\{\|f\|_{\mathcal{L}^\infty(\mathcal{D})} : \ f(g(\lambda_j)) = w_j\}.$$

- Some results also known when S is multiply-connected.

The Field of Values and 2 by 2 Matrices

- Suppose $S = W(A)$. Crouzeix showed that $M_{opt}(A, W(A)) \leq 11.08$ and he conjectures that $M_{opt}(A, W(A)) \leq 2$. (He proved this if A is 2 by 2 or if $W(A)$ is a disk.) In most cases, do not have good estimates for $m_{opt}(A, W(A))$, but...

- If A is a 2 by 2 matrix, since $W(A) = \mathcal{H}_1(A)$, the polynomial numerical hull of degree 1, and since any function of a 2 by 2 matrix A can be written as a first degree polynomial in A,

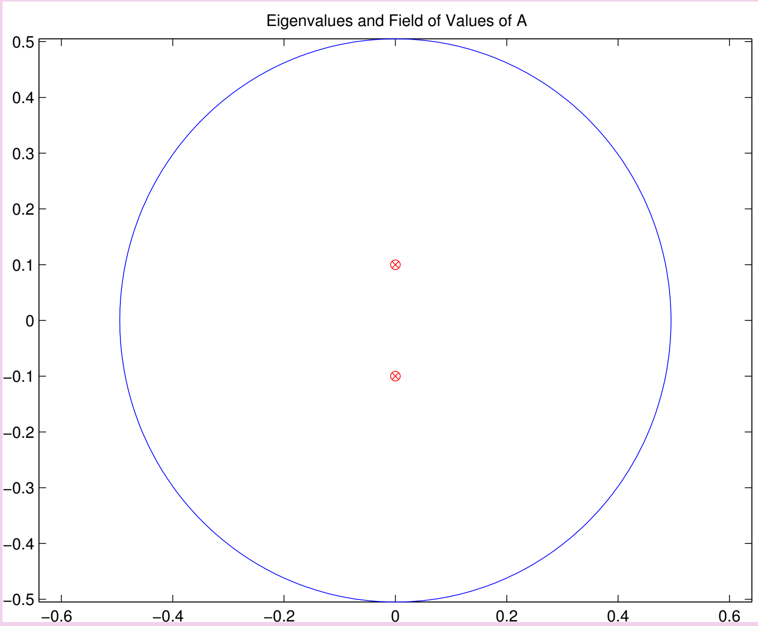
  $$\begin{aligned} \|p(A)\| &\geq \|p_1\|_{\mathcal{L}^\infty(W(A))} \\ &\geq \inf\{\|f\|_{\mathcal{L}^\infty(W(A))} : \ f(A) = p(A)\}. \end{aligned}$$

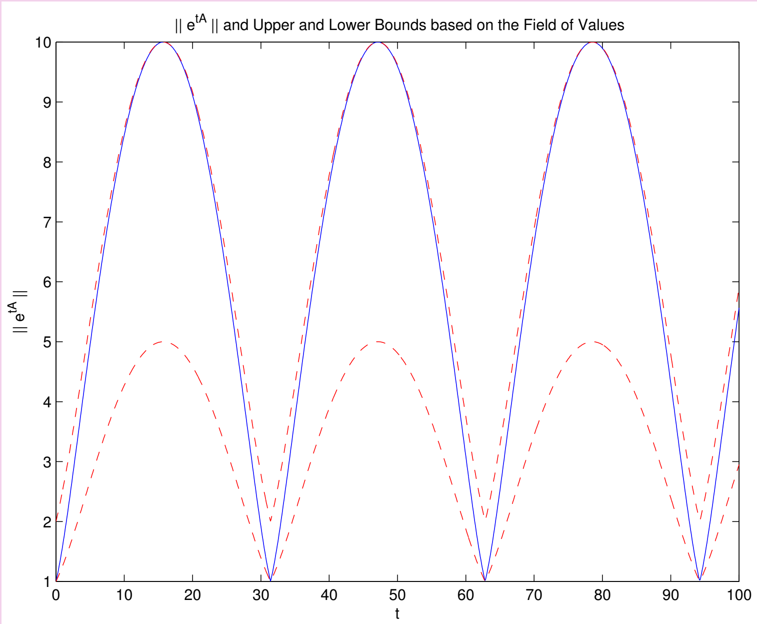
  Hence for 2 by 2 matrices:

  $m_{opt}(A, W(A)) = 1$ and $M_{opt}(A, W(A)) \leq 2$.

Example:

$$A = \begin{pmatrix} 0 & 1 \\ -.01 & 0 \end{pmatrix}$$





The Unit Disk and Perturbed Jordan Blocks

$$A = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ \nu & & & 0 \end{pmatrix}, \quad \nu \in (0,1).$$

- The eigenvalues of A are the nth roots of ν: $\lambda_j = \nu^{(1/n)} e^{2\pi i j/n}$.

- For $\nu = 1$, A is a normal matrix with eigenvalues uniformly distributed about the unit circle. $W(A)$ is the convex hull of the eigenvalues. $\mathcal{H}_{n-1}(A)$ consists of the eigenvalues and the origin. The ϵ-pseudospectrum consists of disks about the eigenvalues of radius ϵ.

- For $\nu = 0$, A is a Jordan block with eigenvalue 0. $W(A)$ is a disk about the origin of radius $\cos(\pi/(n+1))$. $\mathcal{H}_{n-1}(A)$ is a disk of radius $1 - \log(2n)/n + \log(\log(2n))/n + o(1/n)$, and this is equal to the ϵ-pseudospectrum for $\epsilon \approx \log(2n)/(2n) - \log(\log(2n))/(2n)$.

So where in $\mathbb{C}$ does A live?

$$A = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ \nu & & & 0 \end{pmatrix}, \quad \nu \in (0,1).$$

Theorem. For any polynomial p,

$$\|p(A)\| = \inf\{\|f\|_{\mathcal{L}^\infty(\mathcal{D})} : f(A) = p(A)\}.$$

Thus $M_{opt}(A,\mathcal{D}) = m_{opt}(A,\mathcal{D}) = 1$.

Proof: $A = V\Lambda V^{-1}$, where V^T is the Vandermonde matrix for the eigenvalues:

$$V^T = \begin{pmatrix} 1 & \lambda_1 & \dots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \dots & \lambda_2^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & \lambda_n & \dots & \lambda_n^{n-1} \end{pmatrix}$$

How do we compute the minimal-norm interpolating function $\tilde{f}$?

As noted earlier, it has the form

$$\tilde{f}(z) = \mu \frac{\gamma_0 + \gamma_1 z + \ldots + \gamma_{n-1} z^{n-1}}{\bar{\gamma}_{n-1} + \bar{\gamma}_{n-2} z + \ldots + \bar{\gamma}_0 z^{n-1}},$$

and satisfies $\tilde{f}(\lambda_j) = p(\lambda_j)$, $j = 1, \ldots, n$.

If $\gamma = (\gamma_0, \ldots, \gamma_{n-1})^T$, and Π is the permutation matrix with 1's on its skew diagonal, then these conditions are:

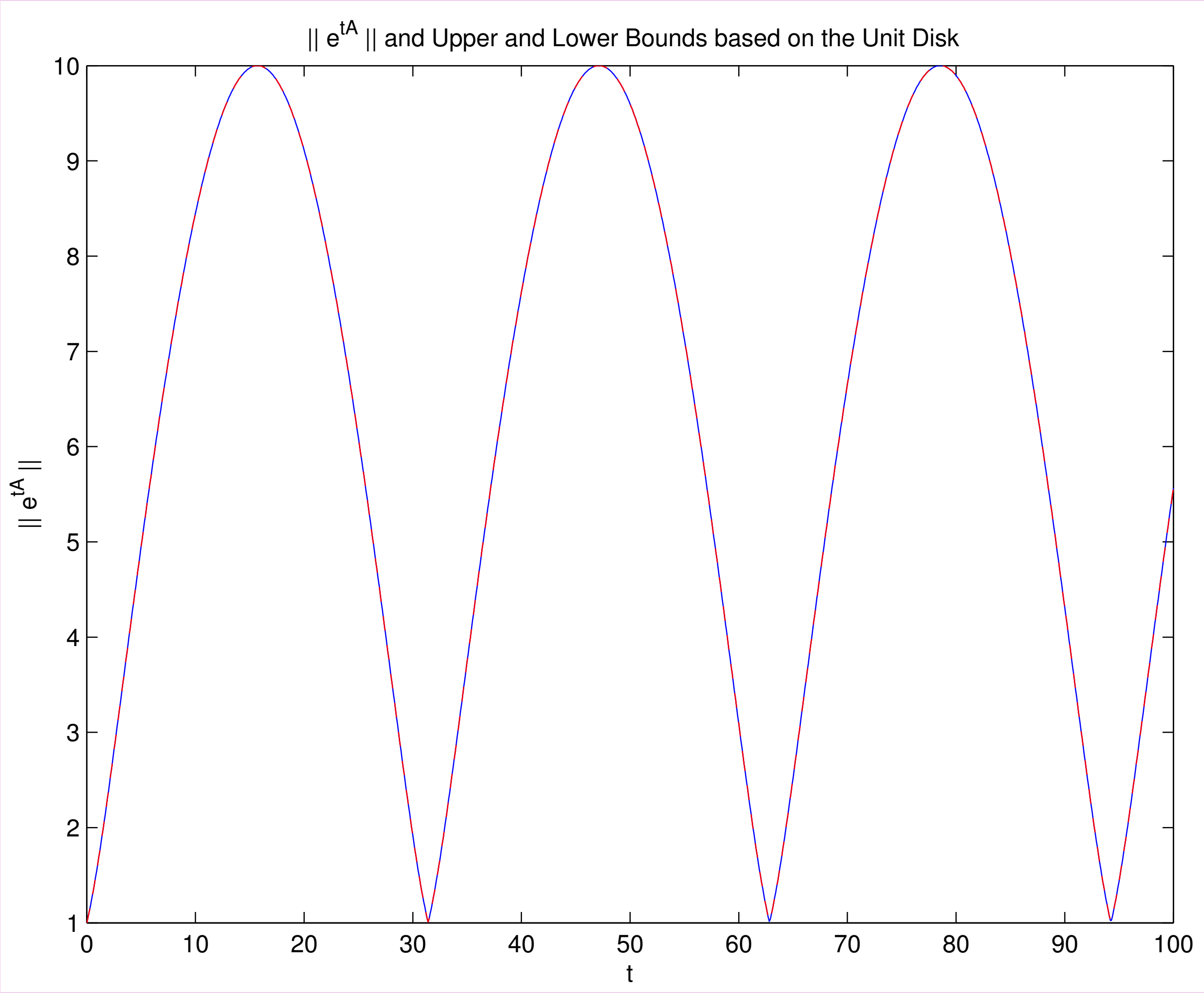
$$V^{-T} p(\Lambda) V^T \Pi \bar{\gamma} = (p(A))^T \Pi \bar{\gamma} = \mu \gamma.$$

Glader and Lindström showed that there is a real scalar μ for which this equation has a nonzero solution vector γ and that the largest such μ is $\|\tilde{f}\|_{\mathcal{L}^\infty(\mathcal{D})}$.

Since $(p(A))^T \Pi$ is (complex) symmetric, it has an SVD of the form $X \Sigma X^T$. The solutions to above equation are: $\gamma = \mathbf{x}_j$, $\mu = \sigma_j$; and $\gamma = i\mathbf{x}_j$, $\mu = -\sigma_j$. $\square$

Example:

$$A = \begin{pmatrix} 0 & 1 \\ -.01 & 0 \end{pmatrix}$$



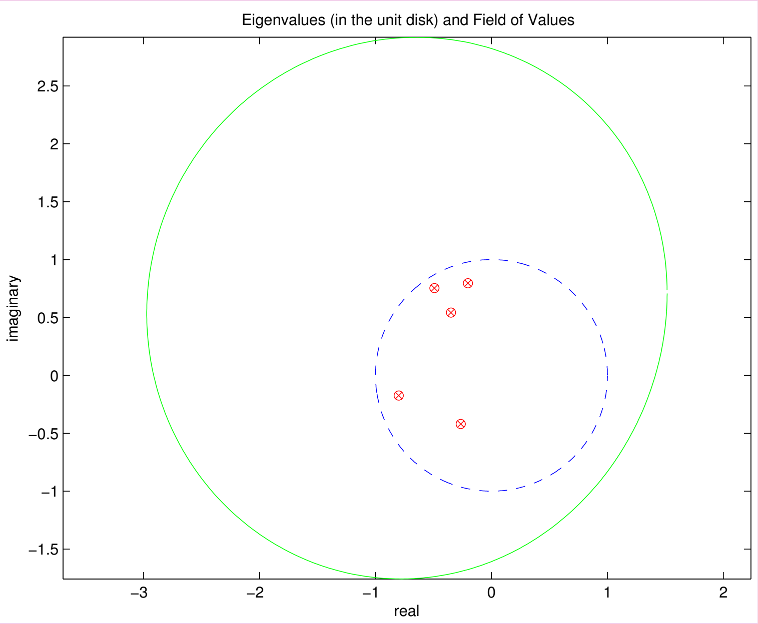
Corollary. If $A = V\Lambda V^{-1}$ where V^T is the Vandermonde matrix for Λ; i.e., if A is a companion matrix with eigenvalues in $\mathcal{D}$, then $m_{opt}(A, \mathcal{D}) = 1$.

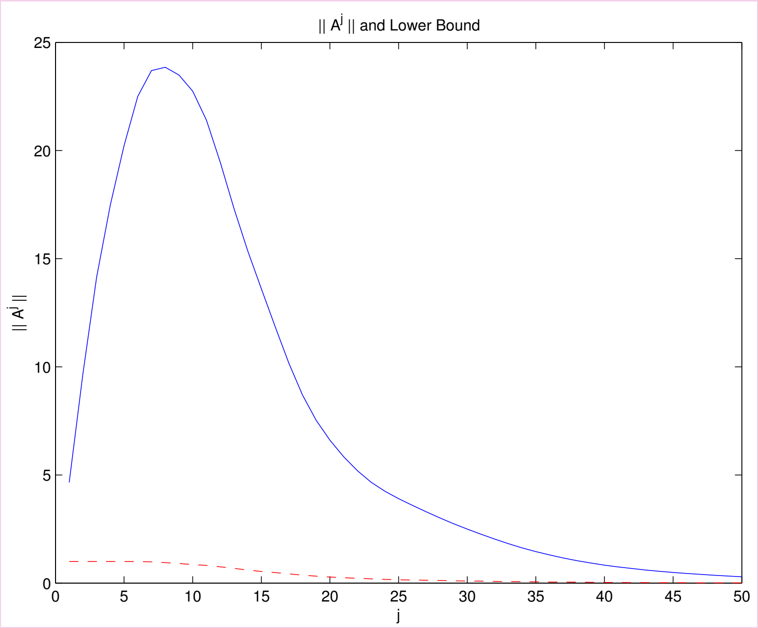
Proof:

$$V^{-T} p(\Lambda) V^T \Pi \bar{\gamma} = (p(A))^T \Pi \bar{\gamma} = \mu\gamma,$$

so $\|p(A)\| \geq |\mu|$. $\square$

Example: Companion matrix with 5 random eigenvalues in the unit disk. $\|A^j\|$ and lower bound.





Given an n by n matrix A, we looked for a set $S \subset \mathbf{C}$ and scalars m and M with M/m of moderate size ($<< \kappa(V)$ if $\kappa(V)$ is large) such that for all polynomials p:

$$m \cdot \inf\{\|f\|_{\mathcal{L}^\infty(S)} : f(A) = p(A)\} \leq \|p(A)\|$$

$$\leq M \cdot \inf\{\|f\|_{\mathcal{L}^\infty(S)} : f(A) = p(A)\}.$$

- In a few exceptional cases (2 by 2 matrices, and perturbed Jordan blocks), we found such a set.

- In general, it seems difficult, perhaps impossible, to find such a set. The problem is that interpolating Blaschke products (like interpolating polynomials) can (but do not always) do wild things between the interpolation points. Hence to get a good value for m, need S to contain little more than $\sigma(A)$. But if $\kappa(V)$ is large, to get a good value for M, need S to contain significantly more than $\sigma(A)$.

- Perhaps the problem should be changed. Limit the class of polynomials. Or look for two different sets $S_m \subset \mathbf{C}$ for lower bounds on $\|p(A)\|$ and $S_M \subset \mathbf{C}$ for upper bounds.

Crouzeix's Conjecture: For any polynomial p,

$$\|p(A)\| \leq 2 \max_{z \in W(A)} |p(z)|.$$

- "If true it would be astounding." (Peter Lax)
- Need only consider $p = B \circ g$ where g is a conformal mapping from $W(A)$ to $\mathcal{D}$ and B is a finite Blaschke product. Show $\|B(g(A))\| \leq 2$.

- Crouzeix proved $\|p(A)\| \leq 11.08 \max_{z \in W(A)} |p(z)|$, but proof is *complicated* and does not appear to be extendable to yield smaller constant.

- Constant 2 can be attained:

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

  $W(A)$ is disk of radius 1/2 about 0.
  $\|A\| = 1 = 2 \max_{z \in \mathcal{D}_{1/2}} |z|$.

- Another open question: If constant 2 is attained, is $W(A)$ necessarily a disk? (Yes, for 2 by 2 matrices.)

- For more information and interesting open problems, see:

  `http://perso.univ-rennes1.fr/michel.crouzeix`