Pseudospectra and the Behavior of Dynamical Systems

Mark Embree
Computational and Applied Mathematics
Rice University
Houston, Texas

Trogir, Croatia
October 2011

Overview of Lectures

These lectures address modern tools
for the spectral analysis of dynamical systems.
We shall cover a mix of theory, computation, and applications.

Goal: By the end, you should be equipped to understand phenomena that many people find quite mysterious when encountered in the wild.

Overview of Lectures

These lectures address modern tools
for the spectral analysis of dynamical systems.
We shall cover a mix of theory, computation, and applications.

Goal: By the end, you should be equipped to understand phenomena that many people find quite mysterious when encountered in the wild.

Lecture 1:

- Normality and Nonnormality
- Pseudospectra
- Bounding Functions of Matrices

Overview of Lectures

These lectures address modern tools
for the spectral analysis of dynamical systems.
We shall cover a mix of theory, computation, and applications.

Goal: By the end, you should be equipped to understand phenomena that many people find quite mysterious when encountered in the wild.

Lecture 1:

- Normality and Nonnormality
- Pseudospectra
- Bounding Functions of Matrices

Lecture 2:

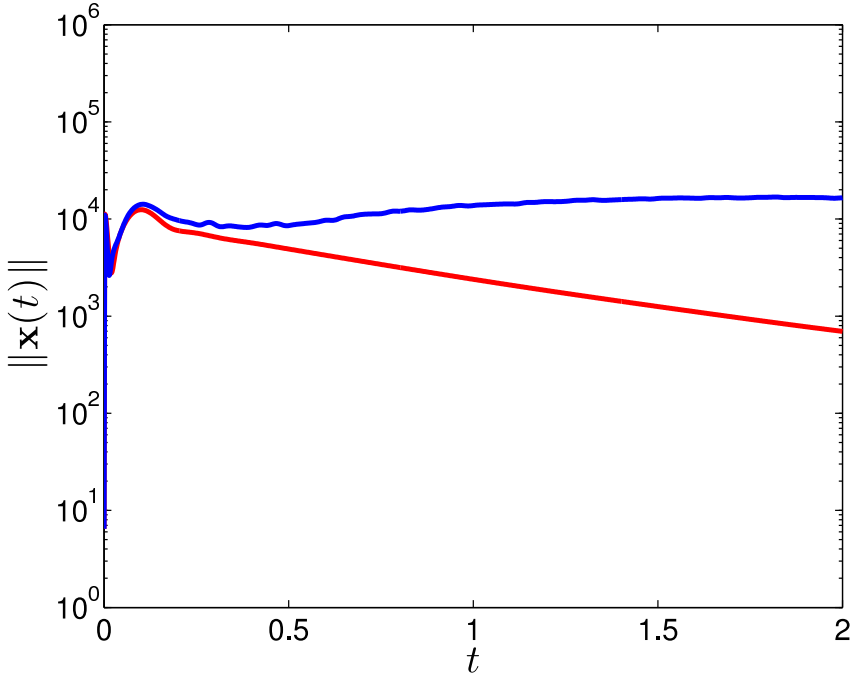
- Balanced Truncation and Lyapunov Equations
- Moment Matching Model Reduction
- Differential Algebraic Equations

Quiz

The plots below show the solution to two dynamical systems, $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$.

Which system is stable?

That is, for which system, does $\mathbf{x}(t) \to \mathbf{0}$ as $t \to \infty$?



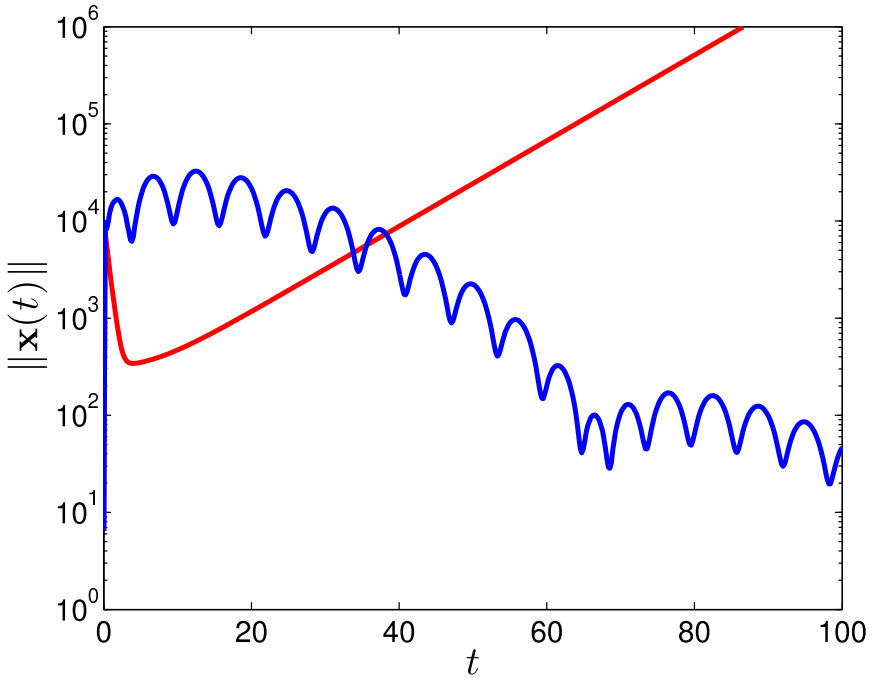
(a) neither system is stable
(b) only the one on the blue system is stable
(c) only the one on the red system is stable
(d) both are stable

Quiz

The plots below show the solution to two dynamical systems, $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$.

Which system is stable?

That is, for which system, does $\mathbf{x}(t) \to \mathbf{0}$ as $t \to \infty$?



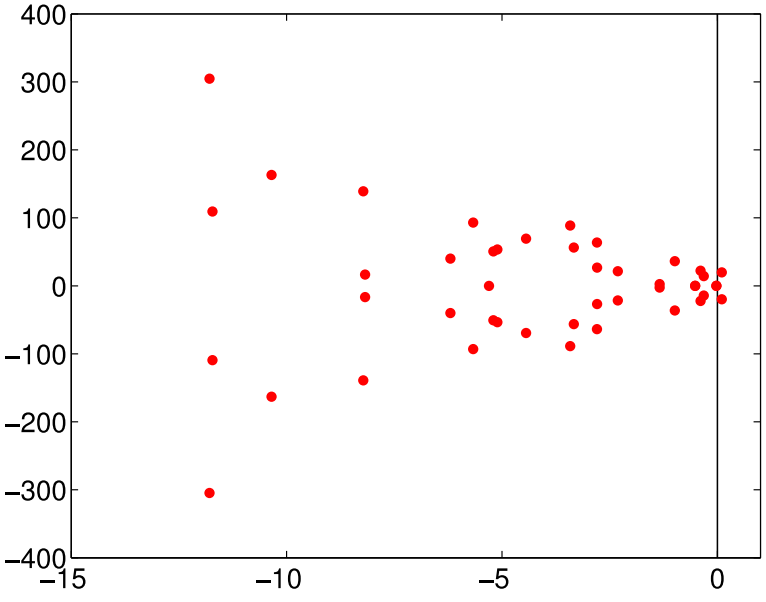
(a) neither system is stable
(b) only the one on the blue system is stable
(c) only the one on the red system is stable
(d) both are stable

Quiz

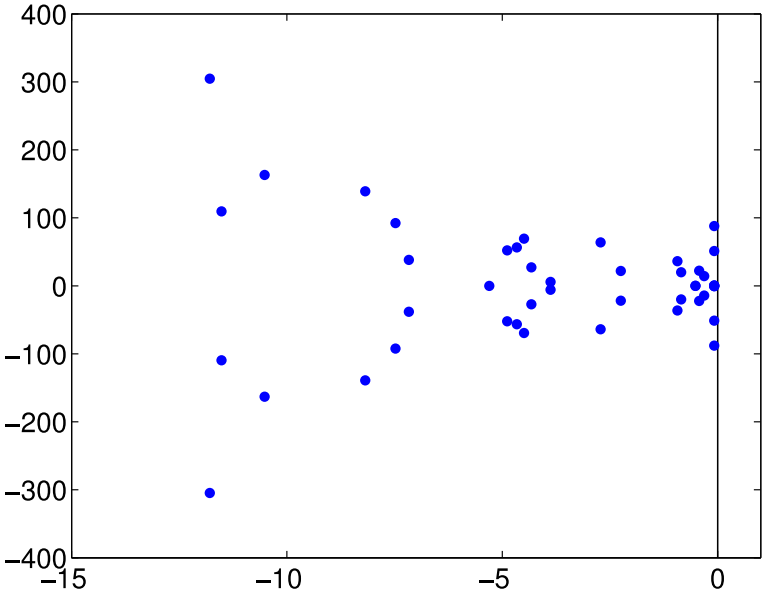
The plots below show the solution to two dynamical systems, $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$.

Which system is stable?

That is, for which system, does $\mathbf{x}(t) \to \mathbf{0}$ as $t \to \infty$?



original unstable system



stabilized system

Eigenvalues of 55×55 Boeing 767 flutter models [Burke, Lewis, Overton].

These eigenvalues do not reveal the exotic transient behavior.

Why is Transient Growth a Big Deal?

Many linear systems arise from the linearization of nonlinear equations, e.g., Navier–Stokes. We compute eigenvalues as part of *linear stability analysis*.

Transient growth in a stable linearized system has implications for the behavior of the associated nonlinear system.

Trefethen, Trefethen, Reddy, Driscoll
"Hydrodynamic stability without eigenvalues," *Science*, 1993.

Why is Transient Growth a Big Deal?

Many linear systems arise from the linearization of nonlinear equations, e.g., Navier–Stokes. We compute eigenvalues as part of *linear stability analysis*.

Transient growth in a stable linearized system has implications for the behavior of the associated nonlinear system.

Consider the autonomous nonlinear system $\mathbf{u}'(t) = \mathbf{f}(\mathbf{u})$.

- Find a steady state $\mathbf{u}_*$, i.e., $\mathbf{f}(\mathbf{u}_*) = \mathbf{0}$.
- Linearize $\mathbf{f}$ about this steady state and analyze small perturbations, $\mathbf{u} = \mathbf{u}_* + \mathbf{v}$:

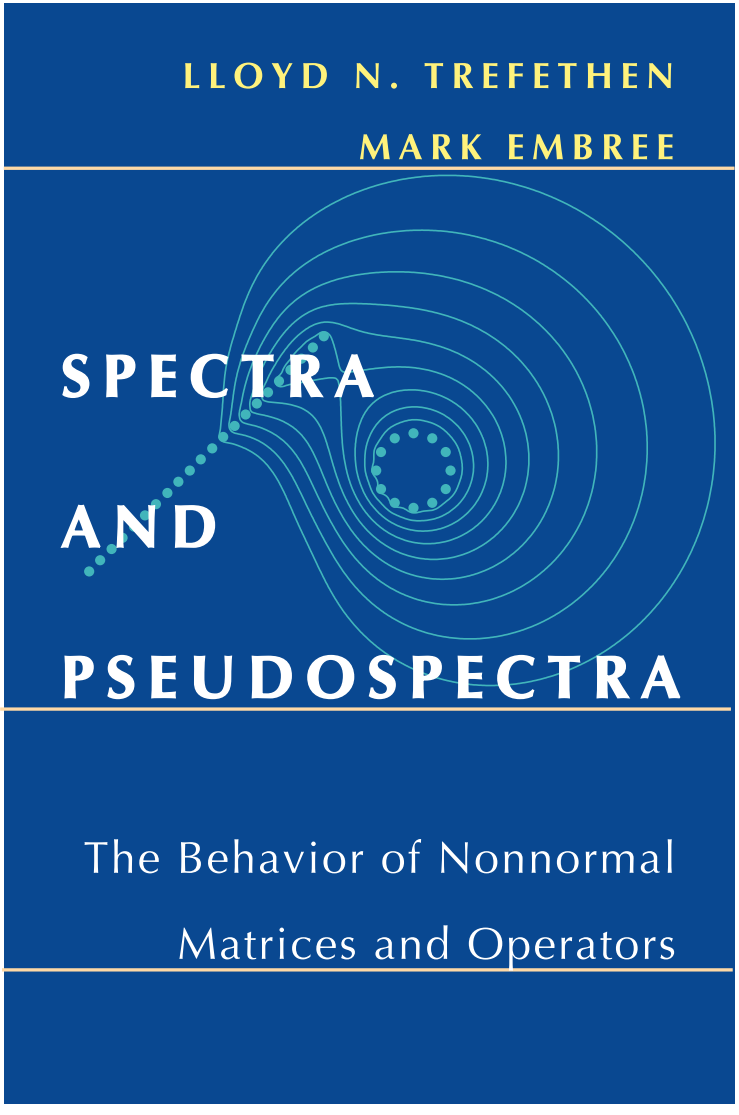
$$\begin{aligned} \mathbf{v}'(t) = \mathbf{u}'(t) &= \mathbf{f}(\mathbf{u}_* + \mathbf{v}) \\ &= \mathbf{f}(\mathbf{u}_*) + \mathbf{A}\mathbf{v} + \mathcal{O}(\|\mathbf{v}\|^2) \\ &= \mathbf{A}\mathbf{v} + \mathcal{O}(\|\mathbf{v}\|^2). \end{aligned}$$

- Ignore higher-order effects, and analyze the linear system $\mathbf{v}'(t) = \mathbf{A}\mathbf{v}(t)$. The steady state $\mathbf{u}_*$ is *stable* provided $\mathbf{A}$ is stable.

But what if the small perturbation $\mathbf{v}(t)$ grows by orders of magnitude before eventually decaying?

Spectra and Pseudospectra

Source for much of the content of these lectures:



Princeton University Press
2005

1. Normality and Nonnormality

Motivating Applications

Many applications lead to models involving linear, non-self-adjoint operators.

- convective fluid flows
- damped mechanical systems
- atmospheric science
- magnetohydrodynamics
- neutron transport
- population dynamics
- food webs
- directed social networks
- Markov chains
- lasers

$$\mathbf{u}_t(\mathbf{x}, t) = \nu \Delta \mathbf{u}(\mathbf{x}, t)$$

$$\mathbf{M}\mathbf{x}''(t) = -\mathbf{K}\mathbf{x}(t)$$

Motivating Applications

Many applications lead to models involving linear, non-self-adjoint operators.

- convective fluid flows
- damped mechanical systems
- atmospheric science
- magnetohydrodynamics
- neutron transport
- population dynamics
- food webs
- directed social networks
- Markov chains
- lasers

$$\begin{aligned} \mathbf{u}_t(\mathbf{x},t) &= \nu\Delta\mathbf{u}(\mathbf{x},t) - (\mathbf{a}\cdot\nabla)\mathbf{u}(\mathbf{x},t) \\ \mathbf{M}\mathbf{x}''(t) &= -\mathbf{K}\mathbf{x}(t) - \mathbf{D}\mathbf{x}'(t) \end{aligned}$$

Normality and Nonnormality

Unless otherwise noted, all matrices are of size $n \times n$, with complex entries.

The *adjoint* is denoted by $\mathbf{A}^* = \overline{\mathbf{A}^T}$.

Definition (normal)

The matrix $\mathbf{A}$ is *normal* if it commutes with its adjoint, $\mathbf{A}^*\mathbf{A} = \mathbf{A}\mathbf{A}^*$.

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} : \mathbf{A}^*\mathbf{A} = \mathbf{A}\mathbf{A}^* = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \implies \text{normal}$$

$$\mathbf{A} = \begin{bmatrix} -1 & 1 \\ 0 & 1 \end{bmatrix} : \mathbf{A}^*\mathbf{A} = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} \neq \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} = \mathbf{A}\mathbf{A}^* \implies \text{nonnormal}$$

Normality and Nonnormality

Unless otherwise noted, all matrices are of size $n \times n$, with complex entries.

The *adjoint* is denoted by $\mathbf{A}^* = \overline{\mathbf{A}^T}$.

Definition (normal)

The matrix $\mathbf{A}$ is *normal* if it commutes with its adjoint, $\mathbf{A}^*\mathbf{A} = \mathbf{A}\mathbf{A}^*$.

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} : \mathbf{A}^*\mathbf{A} = \mathbf{A}\mathbf{A}^* = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \implies \text{normal}$$

$$\mathbf{A} = \begin{bmatrix} -1 & 1 \\ 0 & 1 \end{bmatrix} : \mathbf{A}^*\mathbf{A} = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} \neq \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} = \mathbf{A}\mathbf{A}^* \implies \text{nonnormal}$$

Important note:
The adjoint is defined via the inner product: $\langle \mathbf{A}\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{A}^*\mathbf{y} \rangle$.
hence the definition of normality depends on the inner product.
Here we always use the standard Euclidean inner product, unless noted.
In applications, one *must* use the physically relevant inner product.

Conditions for Normality

Many (~ 100) equivalent definitions of normality are known;
see [Grone et al. 1987] . . . .

By far, the most important of these concerns the eigenvectors of $\mathbf{A}$.

Theorem

The matrix $\mathbf{A}$ *is normal if and only if it is unitarily diagonalizable,*

$$\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^*,$$

for $\mathbf{U}$ *unitary* ($\mathbf{U}^*\mathbf{U} = \mathbf{I}$) *and* $\mathbf{\Lambda}$ *diagonal.*

Equivalently, $\mathbf{A}$ *is normal if and only if is possesses an orthonormal basis of eigenvectors (i.e., the columns of* $\mathbf{U}$*).*

Conditions for Normality

Many (~ 100) equivalent definitions of normality are known;
see [Grone et al. 1987] . . . .

By far, the most important of these concerns the eigenvectors of $\mathbf{A}$.

Theorem

The matrix $\mathbf{A}$ *is normal if and only if it is unitarily diagonalizable,*

$$\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^*,$$

for $\mathbf{U}$ *unitary* ($\mathbf{U}^*\mathbf{U} = \mathbf{I}$) *and* $\mathbf{\Lambda}$ *diagonal.*

Equivalently, $\mathbf{A}$ *is normal if and only if is possesses an orthonormal basis of eigenvectors (i.e., the columns of* $\mathbf{U}$*).*

Hence, any nondiagonalizable (defective) matrix is nonnormal.
But there are many interesting diagonalizable nonnormal matrices.
Our fixation with diagonalizability has caused us to overlook these matrices.

Orthogonality of Eigenvectors

Theorem

The matrix $\mathbf{A}$ *is normal if and only if it is unitarily diagonalizable,*

$$\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^*,$$

for $\mathbf{U}$ *unitary* ($\mathbf{U}^*\mathbf{U} = \mathbf{I}$) *and* $\mathbf{\Lambda}$ *diagonal.*

Equivalently, $\mathbf{A}$ *is normal if and only if is possesses an orthonormal basis of eigenvectors (i.e., the columns of* $\mathbf{U}$*).*

An orthogonal basis of eigenvectors gives a perfect coordinate system for studying dynamical systems:

$$\begin{aligned} \mathbf{x}'(t) = \mathbf{A}\mathbf{x}(t) \quad &\Longrightarrow \quad \mathbf{U}^*\mathbf{x}'(t) = \mathbf{U}^*\mathbf{A}\mathbf{U}\mathbf{U}^*\mathbf{x}(t) \\ &\Longrightarrow \quad \mathbf{z}'(t) = \mathbf{\Lambda}\mathbf{z}(t) \\ &\Longrightarrow \quad z_j'(t) = \lambda_j z_j(t), \end{aligned}$$

with $\|\mathbf{x}(t)\| = \|\mathbf{z}(t)\|$ for all t.

The Perils of Oblique Eigenvectors

Now suppose we only have a diagonalization, $\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}$:

$$\begin{aligned} \mathbf{x}'(t) = \mathbf{A}\mathbf{x}(t) \quad &\Longrightarrow \quad \mathbf{V}^{-1}\mathbf{x}'(t) = \mathbf{V}^{-1}\mathbf{A}\mathbf{V}\mathbf{V}^{-1}\mathbf{x}(t) \\ &\Longrightarrow \quad \mathbf{z}'(t) = \mathbf{\Lambda}\mathbf{z}(t) \\ &\Longrightarrow \quad z_j'(t) = \lambda_j z_j(t), \end{aligned}$$

with $\|\mathbf{x}(t)\| \neq \|\mathbf{z}(t)\|$ in general.

The Perils of Oblique Eigenvectors

Now suppose we only have a diagonalization, $\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}$:

$$\begin{aligned} \mathbf{x}'(t) = \mathbf{A}\mathbf{x}(t) \quad &\Longrightarrow \quad \mathbf{V}^{-1}\mathbf{x}'(t) = \mathbf{V}^{-1}\mathbf{A}\mathbf{V}\mathbf{V}^{-1}\mathbf{x}(t) \\ &\Longrightarrow \quad \mathbf{z}'(t) = \mathbf{\Lambda}\mathbf{z}(t) \\ &\Longrightarrow \quad z_j'(t) = \lambda_j z_j(t), \end{aligned}$$

with $\|\mathbf{x}(t)\| \neq \|\mathbf{z}(t)\|$ in general.

The exact solution is easy:

$$\mathbf{x}(t) = \mathbf{V}\mathbf{z}(t) = \sum_{k=1}^{n} \mathrm{e}^{t\lambda_k} z_k(0)\mathbf{v}_k.$$

Suppose $\|\mathbf{x}(0)\| = 1$. The coefficients $z_k(0)$ might still be quite large:

$$\mathbf{x}(0) = \mathbf{V}\mathbf{z}(0) = \sum_{k=0}^{n} z_k(0)\mathbf{v}_k.$$

The "cancellation" that gives $\mathbf{x}(0)$ is washed out for $t > 0$ by the $\mathrm{e}^{t\lambda_k}$ terms.

Oblique Eigenvectors: Example

Example

$$\begin{bmatrix} x_1'(t) \\ x_2'(t) \end{bmatrix} = \begin{bmatrix} -1/2 & 500 \\ 0 & -5 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}, \qquad \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Eigenvalues and eigenvectors:

$$\lambda_1 = -1/2, \quad \mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad \lambda_2 = -5, \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ -.009 \end{bmatrix}$$

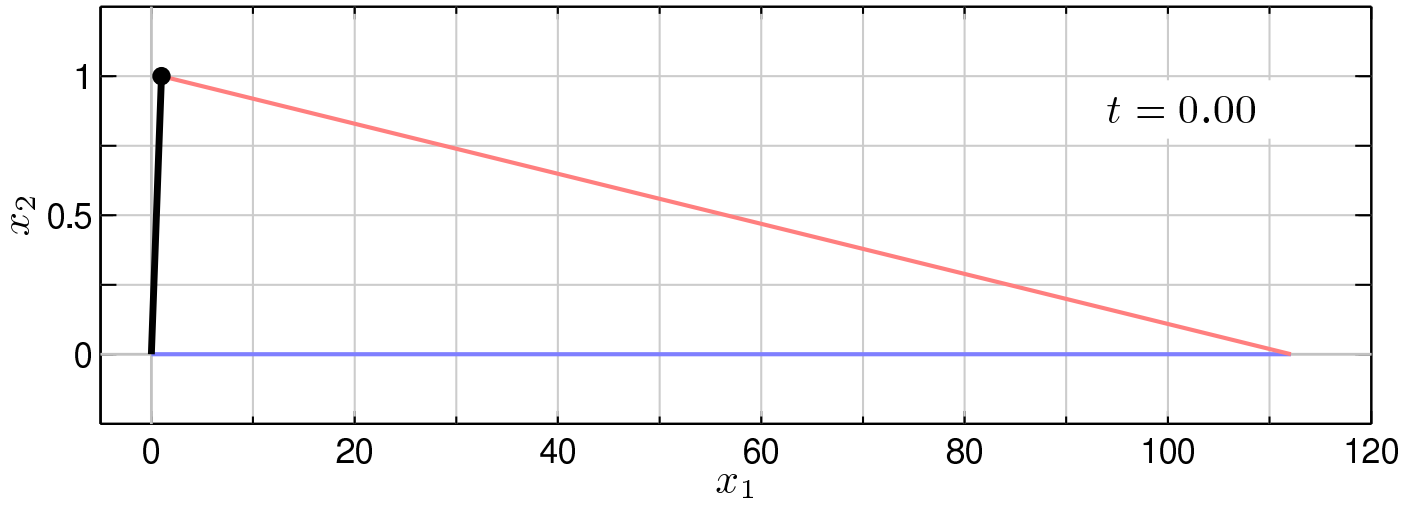
Initial condition:

$$\mathbf{x}(0) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1000}{9}\mathbf{v}_1 - \frac{1009}{9}\mathbf{v}_2.$$

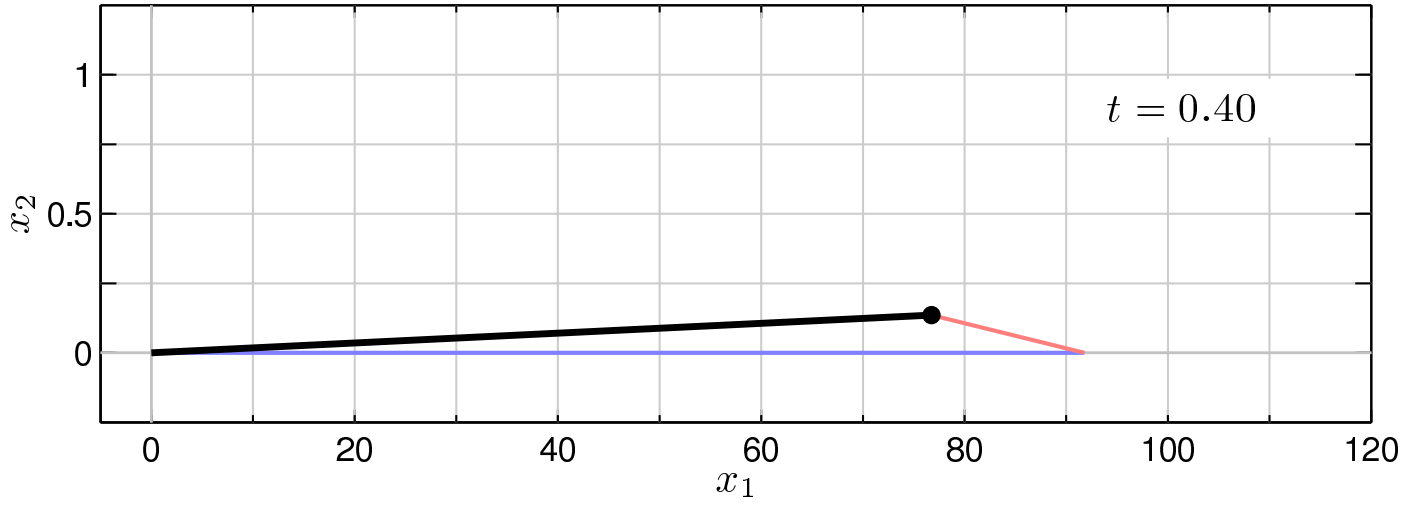
Exact solution:

$$\mathbf{x}(t) = \frac{1000}{9}\mathrm{e}^{-\lambda_1 t}\mathbf{v}_1 - \frac{1009}{9}\mathrm{e}^{-\lambda_2 t}\mathbf{v}_2.$$

Oblique Eigenvectors: Example



$\|\mathbf{x}(0)\| = 1$

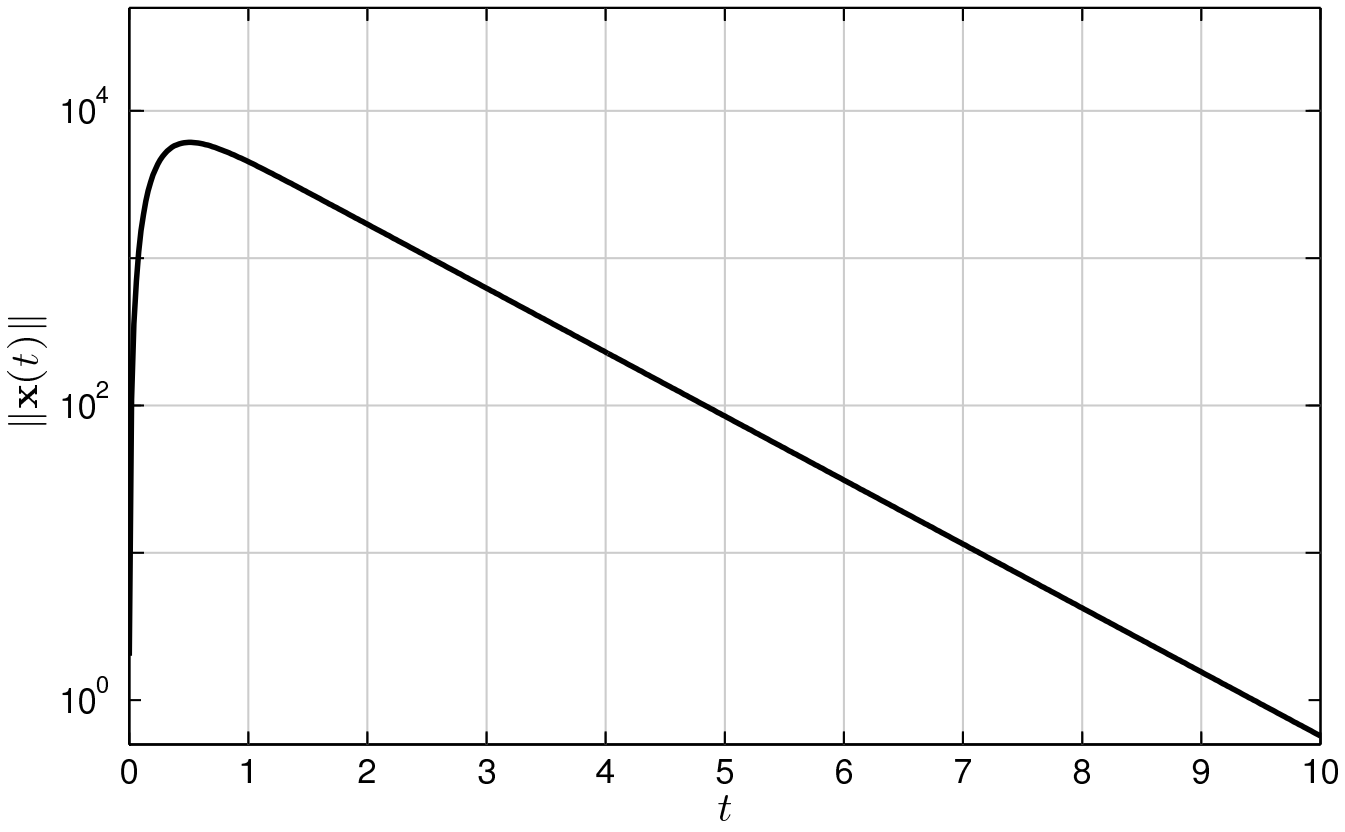


$\|\mathbf{x}(.4)\| = 76.75$

Note the different scales of the horizontal and vertical axes.

Oblique Eigenvectors Can Lead to Transient Growth

Transient growth in the solution: a consequence of nonnormality.



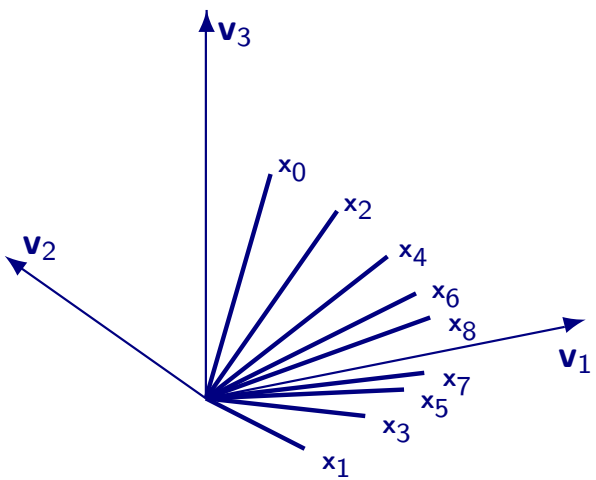
See Figure 1 in "Nineteen Dubious Ways to Compute the Exponential of a Matrix" by Moler and Van Loan, *SIAM Review*, 1978.

Transient Behavior of the Power Method for Computing Eigenvalues

Large coefficients in the expansion of $\mathbf{x}_0$ in the eigenvector basis can lead to cancellation effects in $\mathbf{x}_k = \mathbf{A}^k \mathbf{x}_0$.

Example: here different choices of α and β affect eigenvalue conditioning,

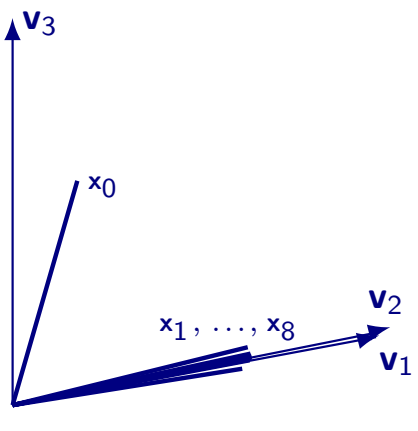
$$\mathbf{A} = \begin{bmatrix} 1 & \alpha & 0 \\ 0 & 3/4 & \beta \\ 0 & 0 & -3/4 \end{bmatrix}, \quad \mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -4\alpha \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 8\alpha\beta/21 \\ -2\beta/3 \\ 1 \end{bmatrix}.$$



$\alpha = \beta = 0$

normal

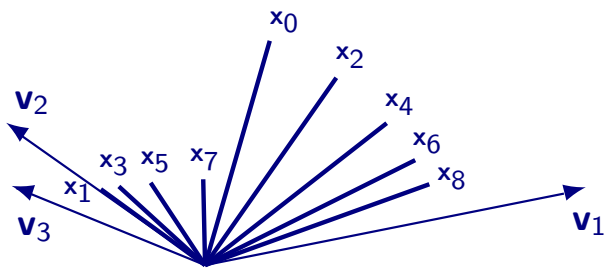
$\mathbf{v}_1 \perp \mathbf{v}_2 \perp \mathbf{v}_3$



$\alpha = 10, \ \beta = 0$

nonnormal

$\mathbf{v}_1 \not\perp \mathbf{v}_2 \perp \mathbf{v}_3$



$\alpha = 0, \ \beta = 10$

nonnormal

$\mathbf{v}_1 \perp \mathbf{v}_2 \not\perp \mathbf{v}_3$

Cf. [Beattie, E., Rossi 2004; Beattie, E., Sorensen 2005]

Tools for Measuring Nonnormality

Given a matrix, we would like some effective way to measure whether we should be concerned about the effects of nonnormality.

Tools for Measuring Nonnormality

Given a matrix, we would like some effective way to measure whether we should be concerned about the effects of nonnormality.

First, we might seek a scalar measure of normality.
Any definition of normality leads to one such gauge.

Tools for Measuring Nonnormality

Given a matrix, we would like some effective way to measure whether we should be concerned about the effects of nonnormality.

First, we might seek a scalar measure of normality.
Any definition of normality leads to one such gauge.

- $\|\mathbf{A}^*\mathbf{A} - \mathbf{A}\mathbf{A}^*\|$

Tools for Measuring Nonnormality

Given a matrix, we would like some effective way to measure whether we should be concerned about the effects of nonnormality.

First, we might seek a scalar measure of normality.
Any definition of normality leads to one such gauge.

- $\|\mathbf{A}^*\mathbf{A} - \mathbf{A}\mathbf{A}^*\|$
- $\min_{\mathbf{Z}\ \text{normal}} \|\mathbf{A} - \mathbf{Z}\|$
  (See work on computing the nearest normal matrix by Gabriel, Ruhe 1987.)

Tools for Measuring Nonnormality

Given a matrix, we would like some effective way to measure whether we should be concerned about the effects of nonnormality.

First, we might seek a scalar measure of normality.
Any definition of normality leads to one such gauge.

- $\|\mathbf{A}^*\mathbf{A} - \mathbf{A}\mathbf{A}^*\|$
- $\min_{\mathbf{Z}\ \text{normal}} \|\mathbf{A} - \mathbf{Z}\|$
  (See work on computing the nearest normal matrix by Gabriel, Ruhe 1987.)
- Henrici's departure from normality:

$$\mathrm{dep}_2(\mathbf{A}) = \min_{\substack{\mathbf{A}=\mathbf{U}(\mathbf{D}+\mathbf{N})\mathbf{U}^* \\ \text{Schur factorization}}} \|\mathbf{N}\|.$$

  No minimization is needed in the Frobenius norm:

$$\mathrm{dep}_F(\mathbf{A}) = \min_{\substack{\mathbf{A}=\mathbf{U}(\mathbf{D}+\mathbf{N})\mathbf{U}^* \\ \text{Schur factorization}}} \|\mathbf{N}\|_F = \sqrt{\|\mathbf{A}\|_F^2 - \sum_{j=1}^{n} |\lambda_j|^2}.$$

Tools for Measuring Nonnormality

Given a matrix, we would like some effective way to measure whether we should be concerned about the effects of nonnormality.

First, we might seek a scalar measure of normality.
Any definition of normality leads to one such gauge.

- $\|\mathbf{A}^*\mathbf{A} - \mathbf{A}\mathbf{A}^*\|$
- $\min_{\mathbf{Z}\ \text{normal}} \|\mathbf{A} - \mathbf{Z}\|$
  (See work on computing the nearest normal matrix by Gabriel, Ruhe 1987.)
- Henrici's departure from normality:

$$\mathrm{dep}_2(\mathbf{A}) = \min_{\substack{\mathbf{A}=\mathbf{U}(\mathbf{D}+\mathbf{N})\mathbf{U}^* \\ \text{Schur factorization}}} \|\mathbf{N}\|.$$

  No minimization is needed in the Frobenius norm:

$$\mathrm{dep}_F(\mathbf{A}) = \min_{\substack{\mathbf{A}=\mathbf{U}(\mathbf{D}+\mathbf{N})\mathbf{U}^* \\ \text{Schur factorization}}} \|\mathbf{N}\|_F = \sqrt{\|\mathbf{A}\|_F^2 - \sum_{j=1}^{n} |\lambda_j|^2}.$$

These are related by equivalence constants [Elsner, Paardekooper, 1987].

Tools for Measuring Nonnormality

Given a matrix, we would like some effective way to measure whether we should be concerned about the effects of nonnormality.

First, we might seek a scalar measure of normality.
Any definition of normality leads to one such gauge.

- $\|\mathbf{A}^*\mathbf{A} - \mathbf{A}\mathbf{A}^*\|$
- $\min_{\mathbf{Z}\ \text{normal}} \|\mathbf{A} - \mathbf{Z}\|$
  (See work on computing the nearest normal matrix by Gabriel, Ruhe 1987.)
- Henrici's departure from normality:

$$\mathrm{dep}_2(\mathbf{A}) = \min_{\substack{\mathbf{A}=\mathbf{U}(\mathbf{D}+\mathbf{N})\mathbf{U}^* \\ \text{Schur factorization}}} \|\mathbf{N}\|.$$

  No minimization is needed in the Frobenius norm:

$$\mathrm{dep}_F(\mathbf{A}) = \min_{\substack{\mathbf{A}=\mathbf{U}(\mathbf{D}+\mathbf{N})\mathbf{U}^* \\ \text{Schur factorization}}} \|\mathbf{N}\|_F = \sqrt{\|\mathbf{A}\|_F^2 - \sum_{j=1}^{n} |\lambda_j|^2}.$$

These are related by equivalence constants [Elsner, Paardekooper, 1987].
None of these measures is of much use in practice.

Tools for Measuring Nonnormality

If $\mathbf{A}$ is diagonalizable,

$$\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1},$$

it is often helpful to characterize nonnormality by $\kappa(\mathbf{V}) := \|\mathbf{V}\|\|\mathbf{V}^{-1}\| \geq 1$.

- This quantity depends on the choice of eigenvectors;
  scaling each column of $\mathbf{V}$ to be a unit vector gets within $\sqrt{n}$
  of the optimal value, if the eigenvalues are distinct [van der Sluis 1969].
- For normal matrices, one can take $\mathbf{V}$ unitary, so $\kappa(\mathbf{V}) = 1$.
- If $\kappa(\mathbf{V})$ is not much more than 1 for some diagonalization, then the effects of nonnormality will be minimal.

Tools for Measuring Nonnormality

If $\mathbf{A}$ is diagonalizable,

$$\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1},$$

it is often helpful to characterize nonnormality by $\kappa(\mathbf{V}) := \|\mathbf{V}\|\|\mathbf{V}^{-1}\| \geq 1$.

- This quantity depends on the choice of eigenvectors; scaling each column of $\mathbf{V}$ to be a unit vector gets within $\sqrt{n}$ of the optimal value, if the eigenvalues are distinct [van der Sluis 1969].
- For normal matrices, one can take $\mathbf{V}$ unitary, so $\kappa(\mathbf{V}) = 1$.
- If $\kappa(\mathbf{V})$ is not much more than 1 for some diagonalization, then the effects of nonnormality will be minimal.

Example (Bound for Continuous Systems)

$$\begin{aligned} \|\mathbf{x}(t)\| = \|e^{t\mathbf{A}}\mathbf{x}(0)\| &\leq \|e^{t\mathbf{A}}\|\|\mathbf{x}(0)\| \\ &\leq \|\mathbf{V}e^{t\mathbf{\Lambda}}\mathbf{V}^{-1}\|\|\mathbf{x}(0)\| \\ &\leq \kappa(\mathbf{V}) \max_{\lambda\in\sigma(\mathbf{A})} |e^{t\lambda}|\|\mathbf{x}(0)\|. \end{aligned}$$

Numerical Range (Field of Values)

Another approach: identify a set in the complex plane to replace the spectrum. This dates to the early 20th century literature in functional analysis, e.g., the *numerical range*, Von Neumann's *spectral sets*, and *sectorial operators*.

Definition (Numerical Range, a.k.a. Field of Values)

The *numerical range* of a matrix is the set

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

Numerical Range (Field of Values)

Another approach: identify a set in the complex plane to replace the spectrum. This dates to the early 20th century literature in functional analysis, e.g., the *numerical range*, Von Neumann's *spectral sets*, and *sectorial operators*.

Definition (Numerical Range, a.k.a. Field of Values)

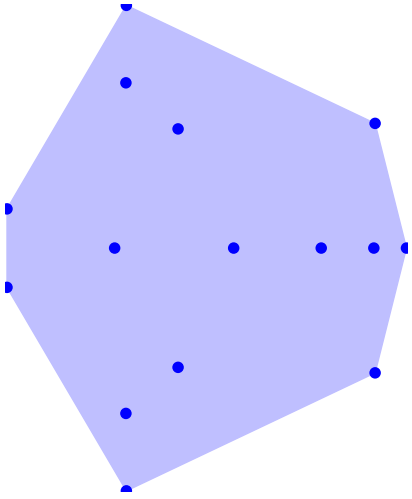
The *numerical range* of a matrix is the set

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

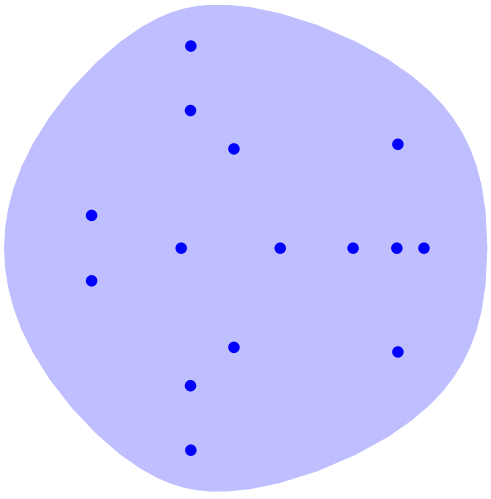
- $W(\mathbf{A})$ is the set of all *Rayleigh quotient* eigenvalue estimates.
- $\sigma(\mathbf{A}) \subset W(\mathbf{A})$ [Proof: take $\mathbf{x}$ to be an eigenvector.]
- $W(\mathbf{A})$ is a closed, convex subset of $\mathbf{C}$.
- If $\mathbf{A}$ is normal, then $W(\mathbf{A})$ is the convex hull of $\sigma(\mathbf{A})$.
- Unlike $\sigma(\mathbf{A})$, the numerical range is robust to perturbations:

$$W(\mathbf{A}+\mathbf{E}) \subseteq W(\mathbf{A}) + \{z \in \mathbf{C} : |z| \leq \|\mathbf{E}\|\}.$$

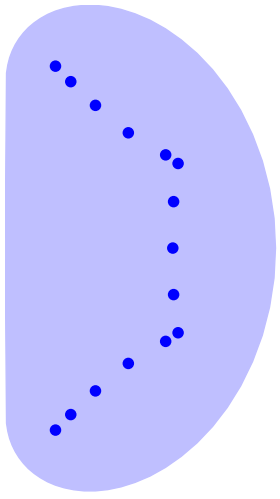
A Gallery of Numerical Ranges



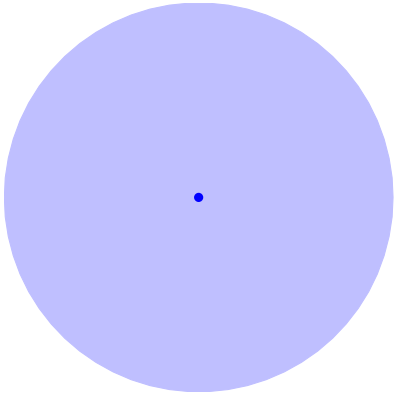
normal



random

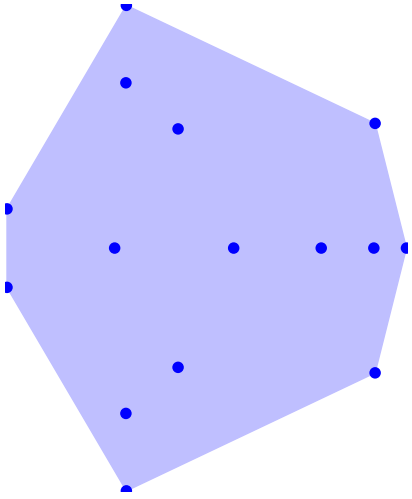


Grcar

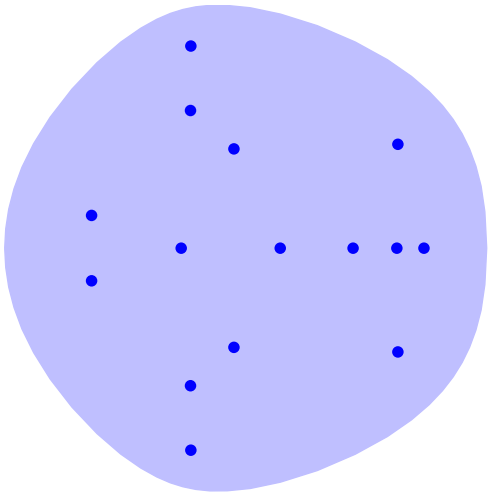


Jordan

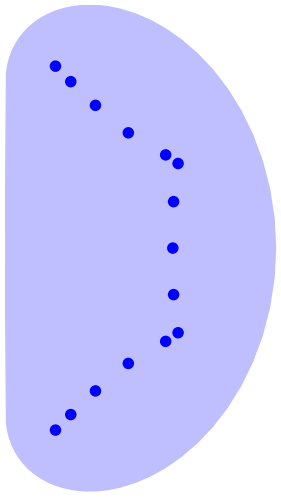
A Gallery of Numerical Ranges



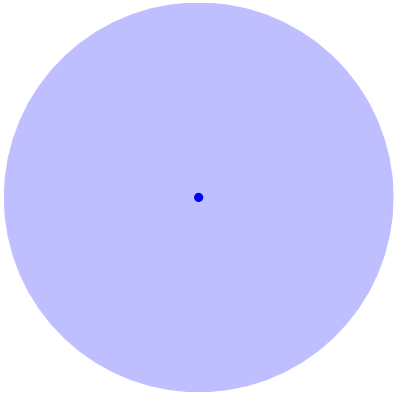
normal



random



Grcar



Jordan

If $z \in W(\mathbf{A})$, then

$$\operatorname{Re} z = \frac{z+\overline{z}}{2} = \frac{1}{2}(\mathbf{x}^*\mathbf{A}\mathbf{x} + \mathbf{x}^*\mathbf{A}^*\mathbf{x}) = \mathbf{x}^*\left(\frac{\mathbf{A}+\mathbf{A}^*}{2}\right)\mathbf{x}.$$

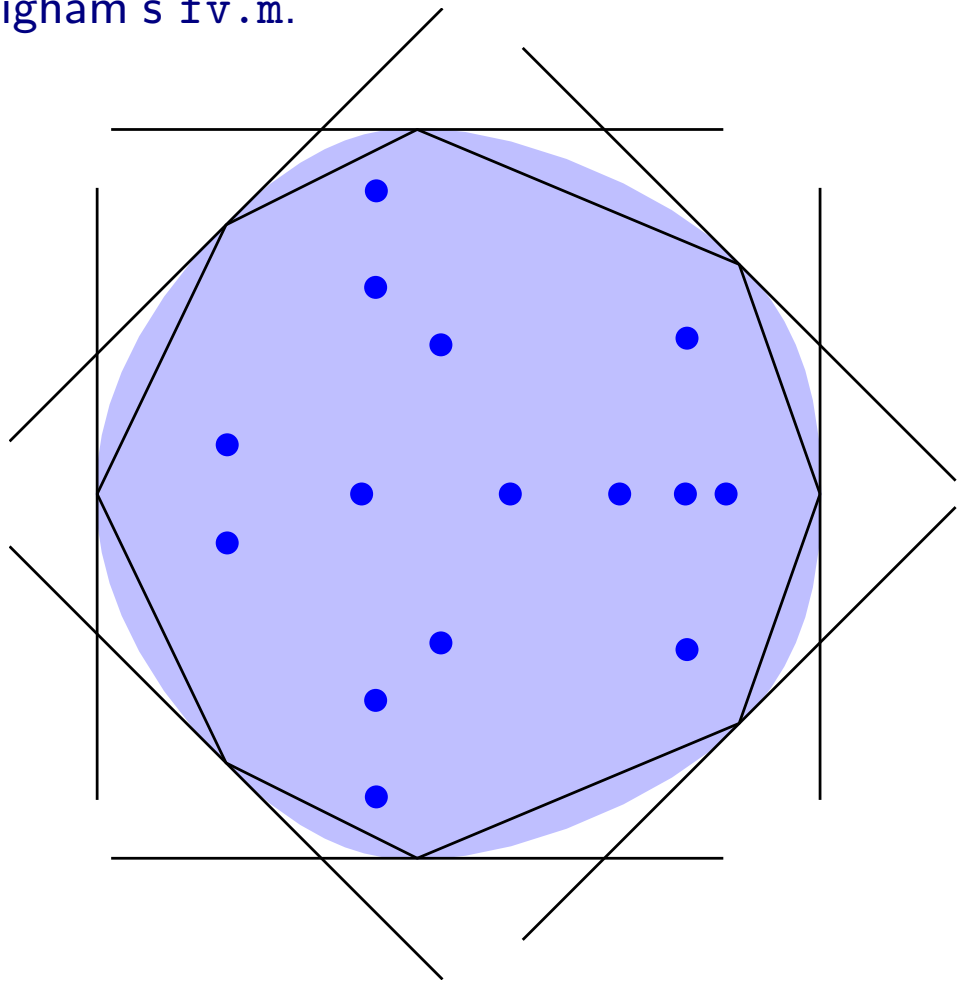
Using properties of Hermitian matrices, we conclude that

$$\operatorname{Re}(W(\mathbf{A})) = \left[\lambda_{\min}\left(\frac{\mathbf{A}+\mathbf{A}^*}{2}\right), \lambda_{\max}\left(\frac{\mathbf{A}+\mathbf{A}^*}{2}\right)\right].$$

Similarly, one can determine the intersection of $W(\mathbf{A})$ with any line in $\mathbf{C}$.

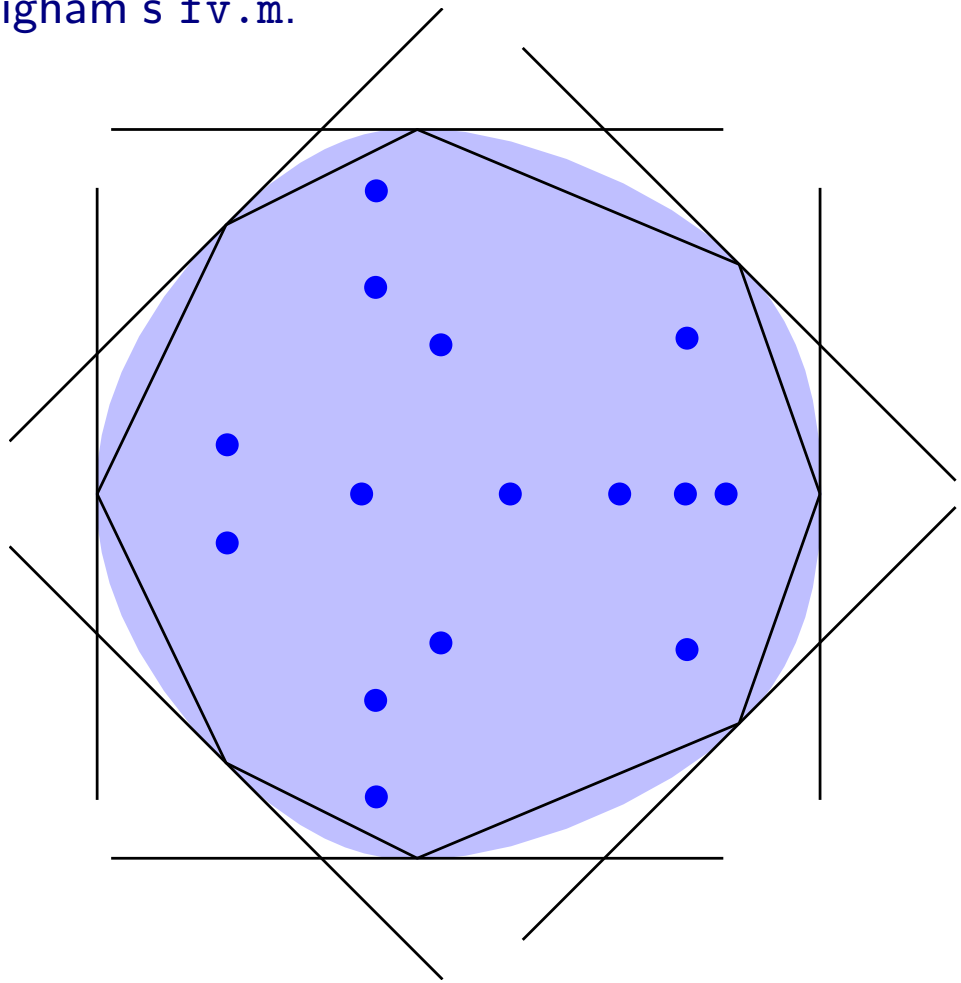
Computation of the Numerical Range

This calculation yields points on the boundary of the numerical range.
Use convexity to obtain polygonal outer and inner approximations
[Johnson 1980]; Higham's `fv.m`.



Computation of the Numerical Range

This calculation yields points on the boundary of the numerical range.
Use convexity to obtain polygonal outer and inner approximations
[Johnson 1980]; Higham's `fv.m`.



Neat problem: Given $z \in W(\mathbf{A})$, find unit vector $\mathbf{x}$ such that $z = \mathbf{x}^* \mathbf{A} \mathbf{x}$
[Uhlig 2008; Carden 2009].

Numerical Range and the Matrix Exponential

What does the numerical range reveal about matrix behavior?

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\|\mathrm{e}^{t\mathbf{A}}\mathbf{x}_0\|\Big|_{t=0} &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\mathbf{x}_0^*\mathrm{e}^{t\mathbf{A}^*}\mathrm{e}^{t\mathbf{A}}\mathbf{x}_0\right)^{1/2} \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\left(\mathbf{x}_0^*(\mathbf{I}+t\mathbf{A}^*)(\mathbf{I}+t\mathbf{A})\mathbf{x}_0\right)^{1/2} \\
&= \frac{1}{\|\mathbf{x}_0\|}\mathbf{x}_0^*\left(\frac{\mathbf{A}+\mathbf{A}^*}{2}\right)\mathbf{x}_0
\end{aligned}$$

So, the rightmost point in $W(\mathbf{A})$ reveals the maximal slope of $\|\mathrm{e}^{t\mathbf{A}}\|$ at $t = 0$.

Numerical Range and the Matrix Exponential

What does the numerical range reveal about matrix behavior?

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\|\mathrm{e}^{t\mathbf{A}}\mathbf{x}_0\|\Big|_{t=0} &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\mathbf{x}_0^*\mathrm{e}^{t\mathbf{A}^*}\mathrm{e}^{t\mathbf{A}}\mathbf{x}_0\right)^{1/2} \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\left(\mathbf{x}_0^*(\mathbf{I}+t\mathbf{A}^*)(\mathbf{I}+t\mathbf{A})\mathbf{x}_0\right)^{1/2} \\
&= \frac{1}{\|\mathbf{x}_0\|}\mathbf{x}_0^*\left(\frac{\mathbf{A}+\mathbf{A}^*}{2}\right)\mathbf{x}_0
\end{aligned}$$

So, the rightmost point in $W(\mathbf{A})$ reveals the maximal slope of $\|\mathrm{e}^{t\mathbf{A}}\|$ at $t = 0$.

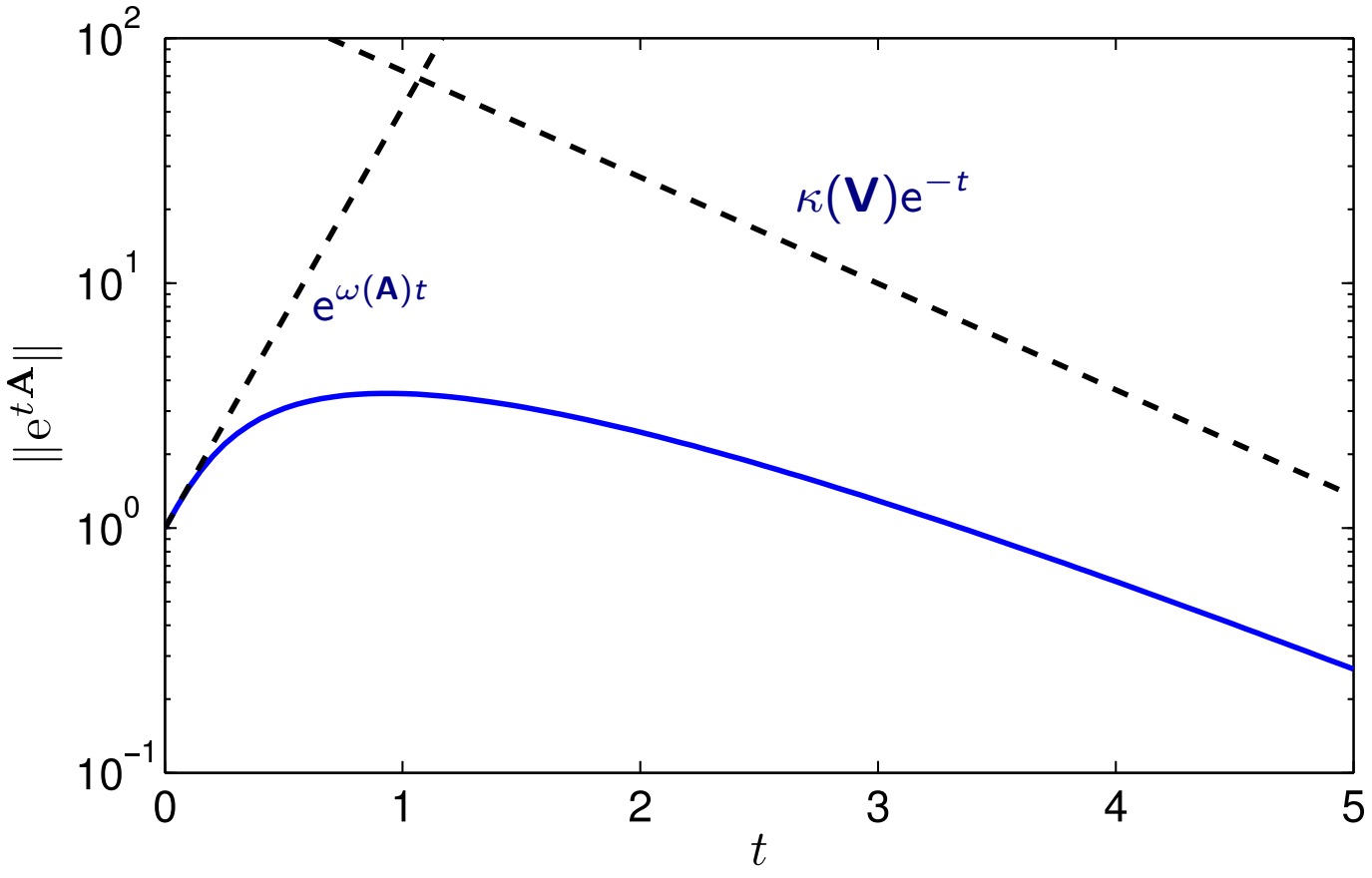
Definition (numerical abscissa)

The *numerical abscissa* is the rightmost in $W(\mathbf{A})$:

$$\omega(\mathbf{A}) := \max_{z\in W(\mathbf{A})} \mathrm{Re}\, z.$$

Intitial Transient Growth via Numerical Abscissa

$$\mathbf{A} = \begin{bmatrix} -1.1 & 10 \\ 0 & -1 \end{bmatrix}$$



Crouzeix's Conjecture

The numerical range can be used to help gauge the size of matrix functions.

Theorem (Crouzeix, 2007)

Let f be a function analytic on $W(\mathbf{A})$. Then

$$\|f(\mathbf{A})\| \leq 11.08 \max_{z \in W(\mathbf{A})} |f(z)|.$$

Crouzeix's Conjecture

The numerical range can be used to help gauge the size of matrix functions.

Theorem (Crouzeix, 2007)

Let f be a function analytic on $W(\mathbf{A})$. Then

$$\|f(\mathbf{A})\| \leq 11.08 \max_{z \in W(\mathbf{A})} |f(z)|.$$

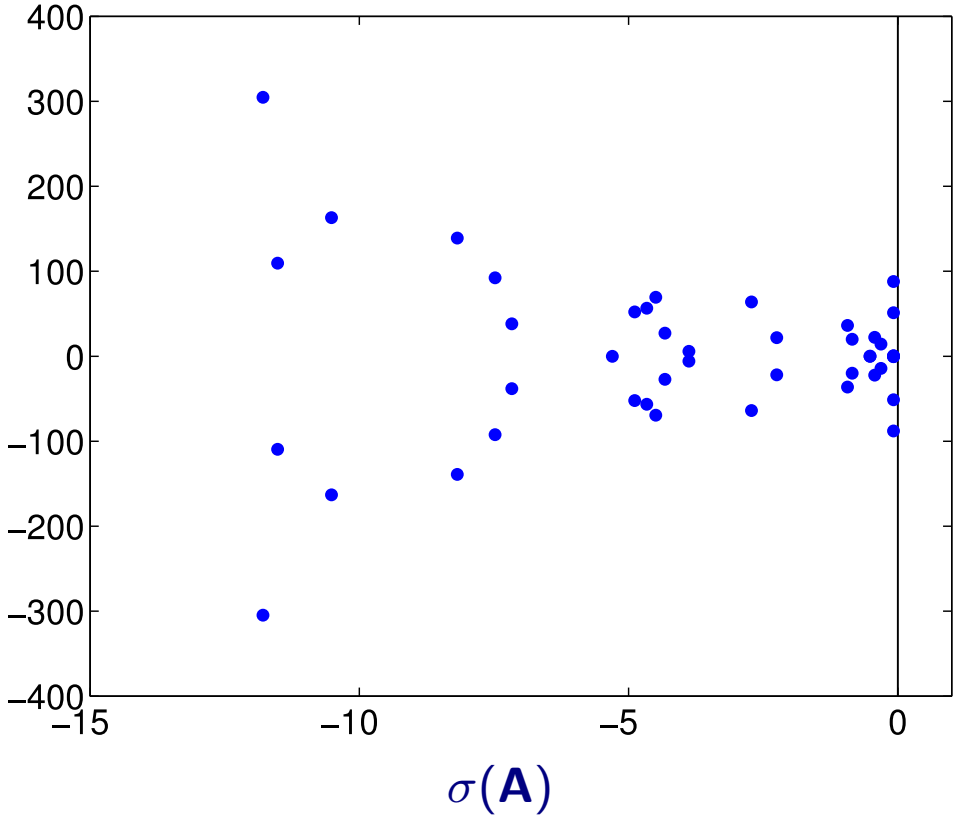
Crouzeix's Conjecture:

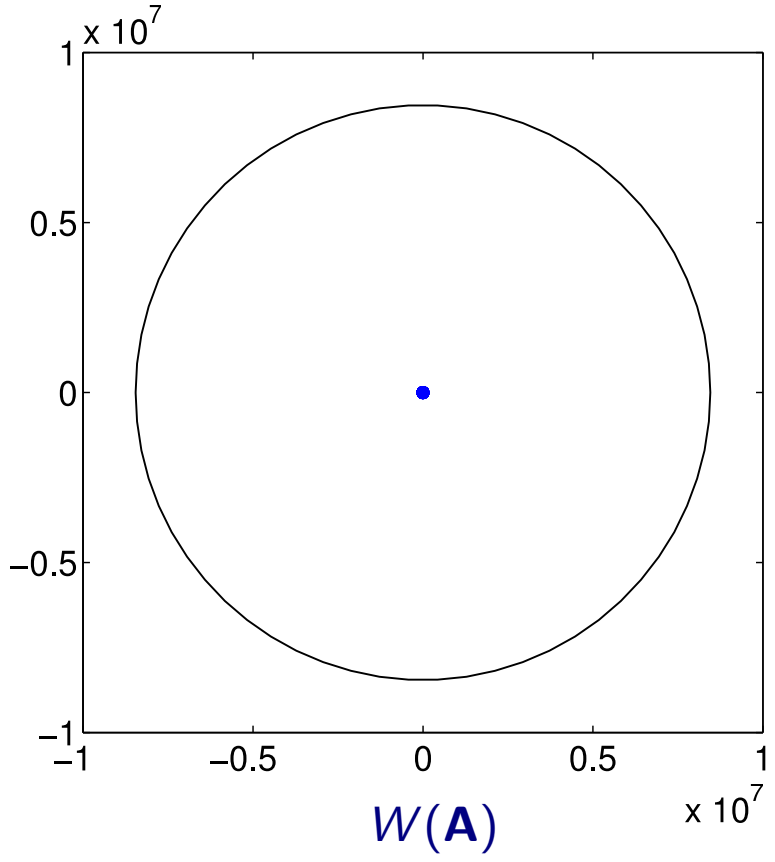
$$\|f(\mathbf{A})\| \leq 2 \max_{z \in W(\mathbf{A})} |f(z)|.$$

Obstacle: in applications, $W(\mathbf{A})$ is often too large, e.g., it contains the origin, or points in the right-half plane, or points of magnitude larger than one.

Boeing 737 Example, Revisited

Numerical range of 55×55 stable matrix.





2. Pseudospectra

Pseudospectra

The numerical range and the eigenvalues both have limitations.
For nonnormal matrices, $W(\mathbf{A})$ can be "too big," while $\sigma(\mathbf{A})$ is "too small."

Here we shall explore another option that, loosely speaking,
interpolates between $\sigma(\mathbf{A})$ and $W(\mathbf{A})$.

Pseudospectra

The numerical range and the eigenvalues both have limitations.
For nonnormal matrices, $W(\mathbf{A})$ can be "too big," while $\sigma(\mathbf{A})$ is "too small."

Here we shall explore another option that, loosely speaking,
interpolates between $\sigma(\mathbf{A})$ and $W(\mathbf{A})$.

Example

Compute eigenvalues of three *similar* 100×100 matrices using MATLAB's `eig`.

$$\begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 1 & & & 0 \end{bmatrix} \qquad \begin{bmatrix} 0 & 1/2 & & \\ & \ddots & \ddots & \\ & & \ddots & 1/2 \\ 2 & & & 0 \end{bmatrix} \qquad \begin{bmatrix} 0 & 1/3 & & \\ & \ddots & \ddots & \\ & & \ddots & 1/3 \\ 3 & & & 0 \end{bmatrix}$$

Pseudospectra

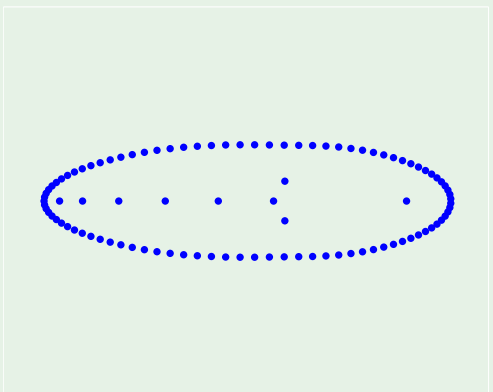
The numerical range and the eigenvalues both have limitations.
For nonnormal matrices, $W(\mathbf{A})$ can be "too big," while $\sigma(\mathbf{A})$ is "too small."

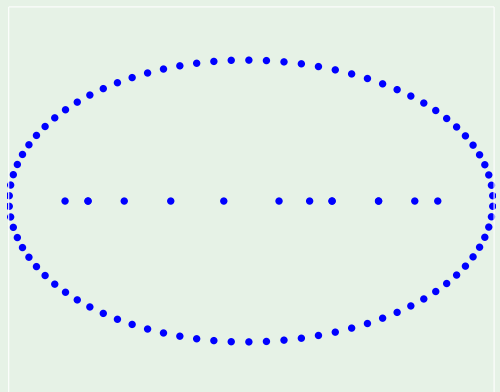
Here we shall explore another option that, loosely speaking,
interpolates between $\sigma(\mathbf{A})$ and $W(\mathbf{A})$.

Example

Compute eigenvalues of three *similar* 100×100 matrices using MATLAB's `eig`.

$$\begin{bmatrix} 0 & 1 & & \\ 1 & 0 & \ddots & \\ & \ddots & \ddots & 1 \\ & & 1 & 0 \end{bmatrix} \qquad \begin{bmatrix} 0 & 1/2 & & \\ 2 & 0 & \ddots & \\ & \ddots & \ddots & 1/2 \\ & & 2 & 0 \end{bmatrix} \qquad \begin{bmatrix} 0 & 1/3 & & \\ 3 & 0 & \ddots & \\ & \ddots & \ddots & 1/3 \\ & & 3 & 0 \end{bmatrix}$$





Pseudospectra

Definition (ε-pseudospectrum)

For any $\varepsilon > 0$, the *ε-pseudospectrum of* $\mathbf{A}$, denoted $\sigma_\varepsilon(\mathbf{A})$, is the set

$$\sigma_\varepsilon(\mathbf{A}) = \{z \in \mathbf{C} : z \in \sigma(\mathbf{A} + \mathbf{E}) \text{ for some } \mathbf{E} \in \mathbf{C}^{n\times n} \text{ with } \|\mathbf{E}\| < \varepsilon\}.$$

Pseudospectra

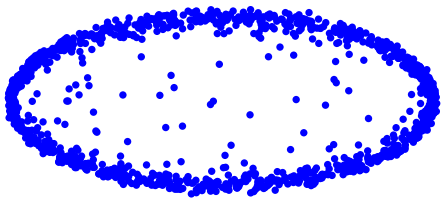
Definition (ε-pseudospectrum)

For any $\varepsilon > 0$, the *ε-pseudospectrum of* $\mathbf{A}$, denoted $\sigma_\varepsilon(\mathbf{A})$, is the set

$$\sigma_\varepsilon(\mathbf{A}) = \{z \in \mathbf{C} : z \in \sigma(\mathbf{A} + \mathbf{E}) \text{ for some } \mathbf{E} \in \mathbf{C}^{n\times n} \text{ with } \|\mathbf{E}\| < \varepsilon\}.$$

We can estimate $\sigma_\varepsilon(\mathbf{A})$ by conducting experiments with random perturbations.

For the 20×20 version of $\mathbf{A} = \text{tridiag}(2, 0, 1/2)$, 50 trials each:



$\varepsilon = 10^{-2}$

$\varepsilon = 10^{-4}$

$\varepsilon = 10^{-6}$

Equivalent Definitions of the Pseudospectrum

Theorem

The following three definitions of the ε-pseudospectrum are equivalent:

1. $\sigma_\varepsilon(\mathbf{A}) = \{z \in \mathbb{C} : z \in \sigma(\mathbf{A} + \mathbf{E})$ *for some* $\mathbf{E} \in \mathbb{C}^{n\times n}$ *with* $\|\mathbf{E}\| < \varepsilon\}$;
2. $\sigma_\varepsilon(\mathbf{A}) = \{z \in \mathbb{C} : \|(z - \mathbf{A})^{-1}\| > 1/\varepsilon\}$;
3. $\sigma_\varepsilon(\mathbf{A}) = \{z \in \mathbb{C} : \|\mathbf{A}\mathbf{v} - z\mathbf{v}\| < \varepsilon$ *for some unit vector* $\mathbf{v} \in \mathbb{C}^n\}$.

Equivalent Definitions of the Pseudospectrum

Theorem

The following three definitions of the ε-pseudospectrum are equivalent:

1. $\sigma_\varepsilon(\mathbf{A}) = \{z \in C : z \in \sigma(\mathbf{A}+\mathbf{E})$ *for some* $\mathbf{E} \in C^{n\times n}$ *with* $\|\mathbf{E}\| < \varepsilon\}$;
2. $\sigma_\varepsilon(\mathbf{A}) = \{z \in C : \|(z-\mathbf{A})^{-1}\| > 1/\varepsilon\}$;
3. $\sigma_\varepsilon(\mathbf{A}) = \{z \in C : \|\mathbf{A}\mathbf{v} - z\mathbf{v}\| < \varepsilon$ *for some unit vector* $\mathbf{v} \in C^n\}$.

Proof. (1) $\implies$ (2)

If $z \in \sigma(\mathbf{A}+\mathbf{E})$ for some $\mathbf{E}$ with $\|\mathbf{E}\| < \varepsilon$, there exists a unit vector $\mathbf{v}$ such that $(\mathbf{A}+\mathbf{E})\mathbf{v} = z\mathbf{v}$. Rearrange to obtain

$$\mathbf{v} = (z-\mathbf{A})^{-1}\mathbf{E}\mathbf{v}.$$

Take norms:

$$\|\mathbf{v}\| = \|(z-\mathbf{A})^{-1}\mathbf{E}\mathbf{v}\| \le \|(z-\mathbf{A})^{-1}\|\|\mathbf{E}\|\|\mathbf{v}\| < \varepsilon\|(z-\mathbf{A})^{-1}\|\|\mathbf{v}\|.$$

Hence $\|(z-\mathbf{A})^{-1}\| > 1/\varepsilon$. □

Equivalent Definitions of the Pseudospectrum

Theorem

The following three definitions of the ε-pseudospectrum are equivalent:

1. $\sigma_\varepsilon(\mathbf{A}) = \{z \in C : z \in \sigma(\mathbf{A}+\mathbf{E})$ *for some* $\mathbf{E} \in C^{n\times n}$ *with* $\|\mathbf{E}\| < \varepsilon\}$;
2. $\sigma_\varepsilon(\mathbf{A}) = \{z \in C : \|(z-\mathbf{A})^{-1}\| > 1/\varepsilon\}$;
3. $\sigma_\varepsilon(\mathbf{A}) = \{z \in C : \|\mathbf{A}\mathbf{v} - z\mathbf{v}\| < \varepsilon$ *for some unit vector* $\mathbf{v} \in C^n\}$.

Proof. (2) $\implies$ (3)

If $\|(z-\mathbf{A})^{-1}\| > 1/\varepsilon$, there exists a unit vector $\mathbf{w}$ with $\|(z-\mathbf{A})^{-1}\mathbf{w}\| > 1/\varepsilon$. Define $\widehat{\mathbf{v}} := (z-\mathbf{A})^{-1}\mathbf{w}$, so that $1/\|\widehat{\mathbf{v}}\| < \varepsilon$, and

$$\frac{\|(z-\mathbf{A})\widehat{\mathbf{v}}\|}{\|\widehat{\mathbf{v}}\|} = \frac{\|\mathbf{w}\|}{\|\widehat{\mathbf{v}}\|} < \varepsilon.$$

Hence we have found a unit vector $\mathbf{v} := \widehat{\mathbf{v}}/\|\widehat{\mathbf{v}}\|$ for which $\|\mathbf{A}\mathbf{v} - z\mathbf{v}\| < \varepsilon$. □

Equivalent Definitions of the Pseudospectrum

Theorem

The following three definitions of the ε-pseudospectrum are equivalent:

1. $\sigma_\varepsilon(\mathbf{A}) = \{z \in C : z \in \sigma(\mathbf{A}+\mathbf{E})$ *for some* $\mathbf{E} \in C^{n\times n}$ *with* $\|\mathbf{E}\| < \varepsilon\}$;
2. $\sigma_\varepsilon(\mathbf{A}) = \{z \in C : \|(z-\mathbf{A})^{-1}\| > 1/\varepsilon\}$;
3. $\sigma_\varepsilon(\mathbf{A}) = \{z \in C : \|\mathbf{A}\mathbf{v} - z\mathbf{v}\| < \varepsilon$ *for some unit vector* $\mathbf{v} \in C^n\}$.

Proof. (3) $\implies$ (1)

Given a unit vector $\mathbf{v}$ such that $\|\mathbf{A}\mathbf{v} - z\mathbf{v}\| < \varepsilon$, define $\mathbf{r} := \mathbf{A}\mathbf{v} - z\mathbf{v}$.
Now set $\mathbf{E} := -\mathbf{r}\mathbf{v}^*$, so that

$$(\mathbf{A}+\mathbf{E})\mathbf{v} = (\mathbf{A} - \mathbf{r}\mathbf{v}^*)\mathbf{v} = \mathbf{A}\mathbf{v} - \mathbf{r} = z\mathbf{v}.$$

Hence $z \in \sigma(\mathbf{A}+\mathbf{E})$. □

Equivalent Definitions of the Pseudospectrum

Theorem

The following three definitions of the ε-pseudospectrum are equivalent:

1. $\sigma_\varepsilon(\mathbf{A}) = \{z \in \mathbb{C} : z \in \sigma(\mathbf{A} + \mathbf{E})$ *for some* $\mathbf{E} \in \mathbb{C}^{n\times n}$ *with* $\|\mathbf{E}\| < \varepsilon\}$;
2. $\sigma_\varepsilon(\mathbf{A}) = \{z \in \mathbb{C} : \|(z - \mathbf{A})^{-1}\| > 1/\varepsilon\}$;
3. $\sigma_\varepsilon(\mathbf{A}) = \{z \in \mathbb{C} : \|\mathbf{A}\mathbf{v} - z\mathbf{v}\| < \varepsilon$ *for some unit vector* $\mathbf{v} \in \mathbb{C}^n\}$.

These different definitions are useful in different contexts:

1. interpreting numerically computed eigenvalues;
2. analyzing matrix behavior/functions of matrices;
   computing pseudospectra on a grid in $\mathbb{C}$;
3. proving bounds on a particular $\sigma_\varepsilon(\mathbf{A})$.

History of Pseudospectra

Invented at least four times, independently:

- Jim Varah in his 1967 Stanford PhD thesis.
- Henry Landau in a 1975 paper, motivated by laser theory.
- S. K. Godunov and colleagues in Novosibirsk in the 1980s.
- L. N. Trefethen, motivated by stability of spectral methods (1990).

Early adopters include Wilkinson (1986), Demmel (1987), Chatelin (1990s).

Properties of Pseudospectra

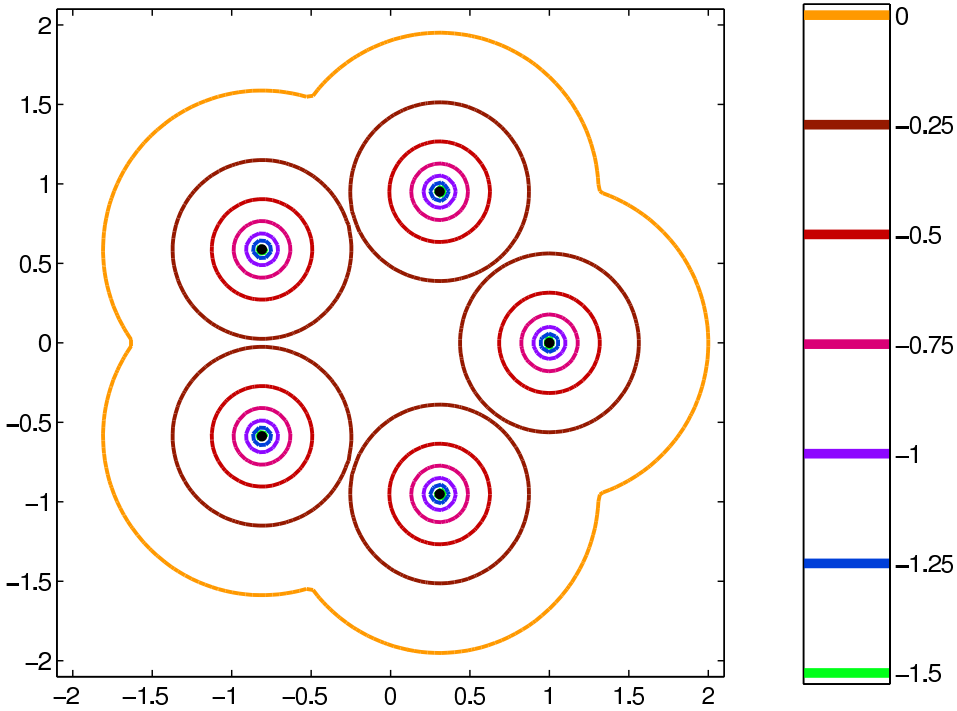
A is normal $\iff$ $\sigma_\varepsilon(\mathbf{A})$ is the union of open ε-balls about each eigenvalue:

$$\mathbf{A} \text{ normal} \quad \Longrightarrow \quad \sigma_\varepsilon(\mathbf{A}) = \bigcup_j \lambda_j + \Delta_\varepsilon$$

$$\mathbf{A} \text{ nonnormal} \quad \Longrightarrow \quad \sigma_\varepsilon(\mathbf{A}) \supset \bigcup_j \lambda_j + \Delta_\varepsilon$$

A circulant (hence normal) matrix:

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$



Properties of Pseudospectra

A is normal $\iff$ $\sigma_\varepsilon(\mathbf{A})$ is the union of open ε-balls about each eigenvalue:

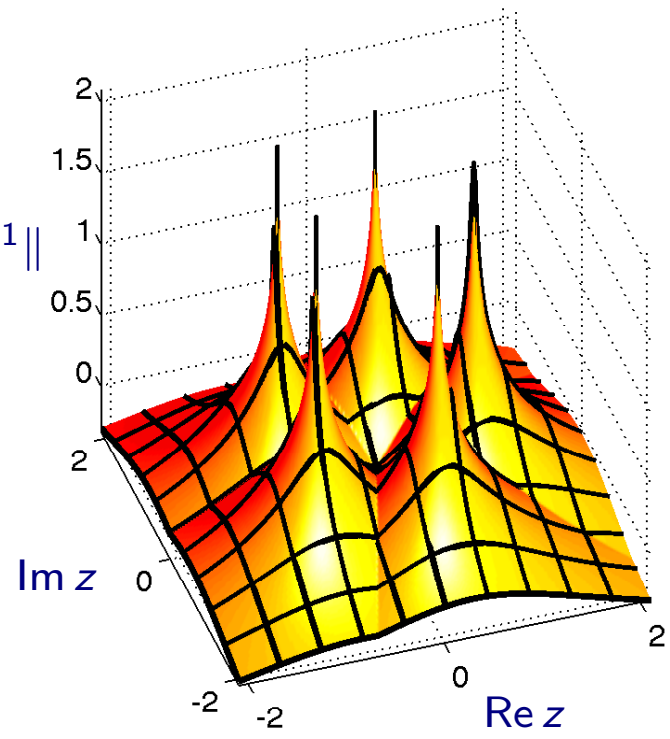
$$\mathbf{A} \text{ normal} \quad \Longrightarrow \quad \sigma_\varepsilon(\mathbf{A}) = \bigcup_j \lambda_j + \Delta_\varepsilon$$

$$\mathbf{A} \text{ nonnormal} \quad \Longrightarrow \quad \sigma_\varepsilon(\mathbf{A}) \supset \bigcup_j \lambda_j + \Delta_\varepsilon$$

A circulant (hence normal) matrix:

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$\|(z - \mathbf{A})^{-1}\|$



Normal Matrices: matching distance

An easy misinterpretation: "A size ε perturbation to a normal matrix can move the eigenvalues by no more than ε."

This holds for Hermitian matrices, but not all normal matrices.

An example constructed by Gerd Krause (see Bhatia, *Matrix Analysis*, 1997):

$$\mathbf{A} = \mathrm{diag}(1, (4 + 5\sqrt{-3})/13, (-1 + 2\sqrt{-3})/13).$$

Now construct the (unitary) Householder reflector

$$\mathbf{U} = \mathbf{I} - 2\mathbf{v}\mathbf{v}^*$$

for $\mathbf{v} = [\sqrt{5/8}, 1/2, \sqrt{1/8}]^*$, and define $\mathbf{E}$ via

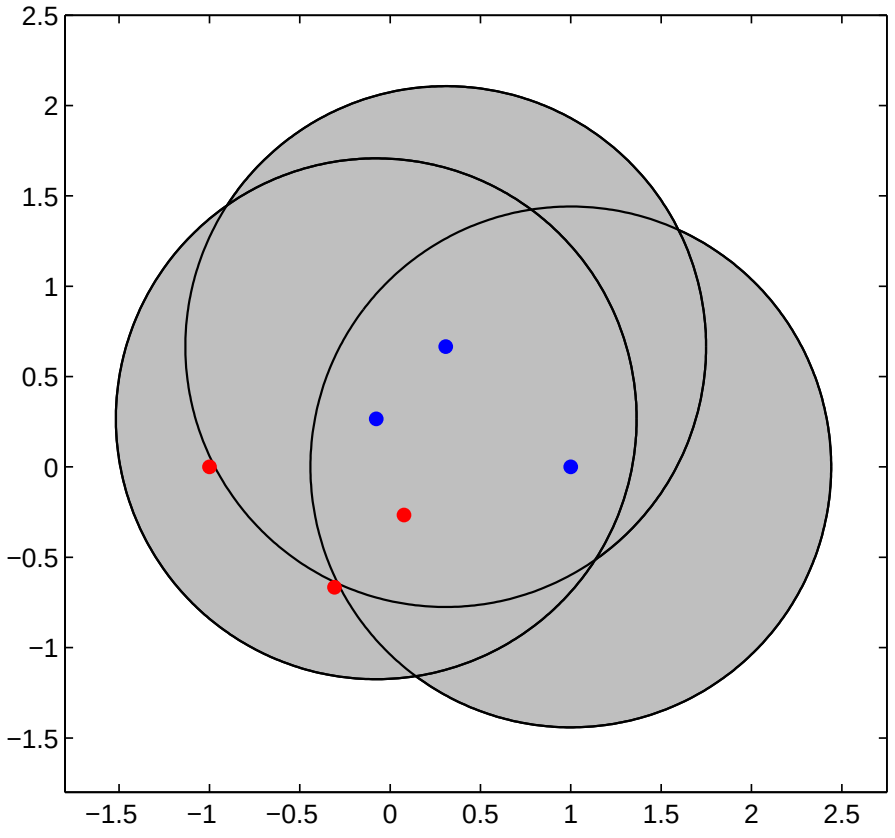
$$\mathbf{A} + \mathbf{E} = -\mathbf{U}^*\mathbf{A}\mathbf{U}.$$

By construction $\mathbf{A} + \mathbf{E}$ is normal and $\sigma(\mathbf{A} + \mathbf{E}) = -\sigma(\mathbf{A})$.

Normal Matrices: matching distance

$$\mathbf{A} = \mathrm{diag}(1, (4 + 5\sqrt{-3})/13, (-1 + 2\sqrt{-3})/13).$$

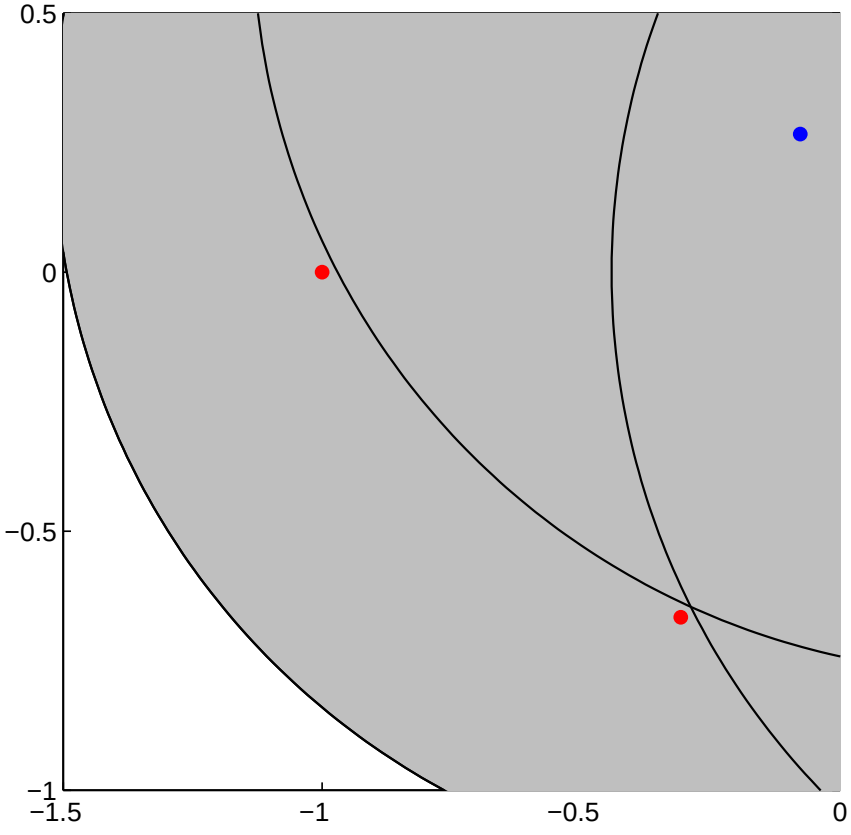
$$\mathbf{A} + \mathbf{E} = -\mathbf{U}^* \mathbf{A} \mathbf{U}.$$



Normal Matrices: matching distance

$$\mathbf{A} = \mathrm{diag}(1, (4 + 5\sqrt{-3})/13, (-1 + 2\sqrt{-3})/13).$$

$$\mathbf{A} + \mathbf{E} = -\mathbf{U}^* \mathbf{A} \mathbf{U}.$$



Properties of Pseudospectra

Theorem (Facts about Pseudospectra)

For all $\varepsilon > 0$,

- $\sigma_\varepsilon(\mathbf{A})$ *is an open, finite set that contains the spectrum.*
- $\sigma_\varepsilon(\mathbf{A})$ *is stable to perturbations:* $\sigma_\varepsilon(\mathbf{A}+\mathbf{E}) \subseteq \sigma_{\varepsilon+\|\mathbf{E}\|}(\mathbf{A})$.
- *If* $\mathbf{U}$ *is unitary,* $\sigma_\varepsilon(\mathbf{UAU}^*) = \sigma_\varepsilon(\mathbf{A})$.
- *For* $\mathbf{V}$ *invertible,* $\sigma_{\varepsilon/\kappa(\mathbf{V})}(\mathbf{VAV}^{-1}) \subseteq \sigma_\varepsilon(\mathbf{A}) \subseteq \sigma_{\varepsilon\kappa(\mathbf{V})}(\mathbf{VAV}^{-1})$.
- $\sigma_\varepsilon(\mathbf{A}+\alpha) = \alpha + \sigma_\varepsilon(\mathbf{A})$.
- $\sigma_{|\gamma|\varepsilon}(\gamma\mathbf{A}) = \gamma\sigma_\varepsilon(\mathbf{A})$.
- $\sigma_\varepsilon\left(\begin{bmatrix}\mathbf{A} & \mathbf{0}\\ \mathbf{0} & \mathbf{B}\end{bmatrix}\right) = \sigma_\varepsilon(\mathbf{A}) \cup \sigma_\varepsilon(\mathbf{B})$.

Relationship to $\kappa(\mathbf{V})$ and $\mathbf{W}(\mathbf{A})$

Let $\Delta_r := \{z \in \mathbf{C} : |z| < r\}$ denote the open disk of radius $r > 0$.

Theorem (Bauer–Fike, 1963)

Let $\mathbf{A}$ *be diagonalizable,* $\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}$. *Then for all* $\varepsilon > 0$,

$$\sigma_\varepsilon(\mathbf{A}) \subseteq \sigma(\mathbf{A}) + \Delta_{\varepsilon\kappa(\mathbf{V})}.$$

If $\kappa(\mathbf{V})$ is small, then $\sigma_\varepsilon(\mathbf{A})$ cannot contain points far from $\sigma(\mathbf{A})$.

Relationship to $\kappa(\mathbf{V})$ and $W(\mathbf{A})$

Let $\Delta_r := \{z \in \mathbf{C} : |z| < r\}$ denote the open disk of radius $r > 0$.

Theorem (Bauer–Fike, 1963)

Let $\mathbf{A}$ *be diagonalizable,* $\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}$. *Then for all* $\varepsilon > 0$,

$$\sigma_\varepsilon(\mathbf{A}) \subseteq \sigma(\mathbf{A}) + \Delta_{\varepsilon\kappa(\mathbf{V})}.$$

If $\kappa(\mathbf{V})$ is small, then $\sigma_\varepsilon(\mathbf{A})$ cannot contain points far from $\sigma(\mathbf{A})$.

Theorem (Stone, 1932)

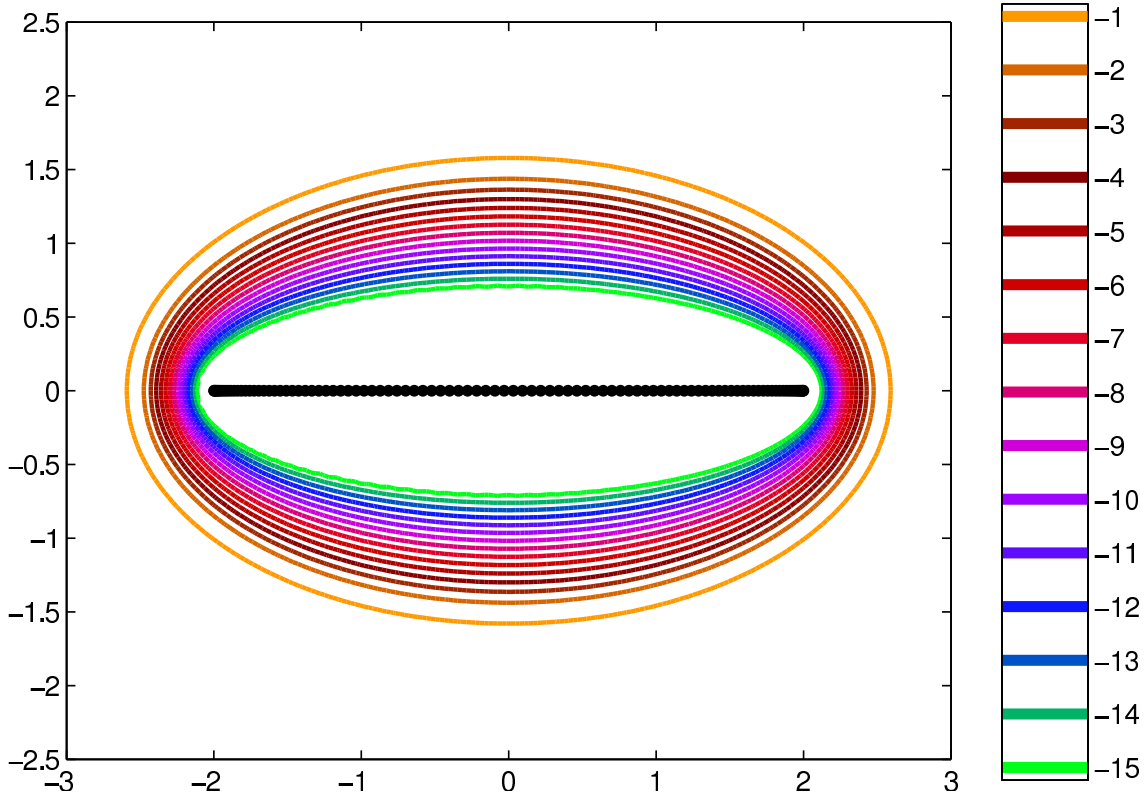
For any $\mathbf{A}$,

$$\sigma_\varepsilon(\mathbf{A}) \subseteq W(\mathbf{A}) + \Delta_\varepsilon.$$

The pseudospectrum $\sigma_\varepsilon(\mathbf{A})$ cannot be bigger than $W(\mathbf{A})$ in an interesting way.

Toeplitz Matrices

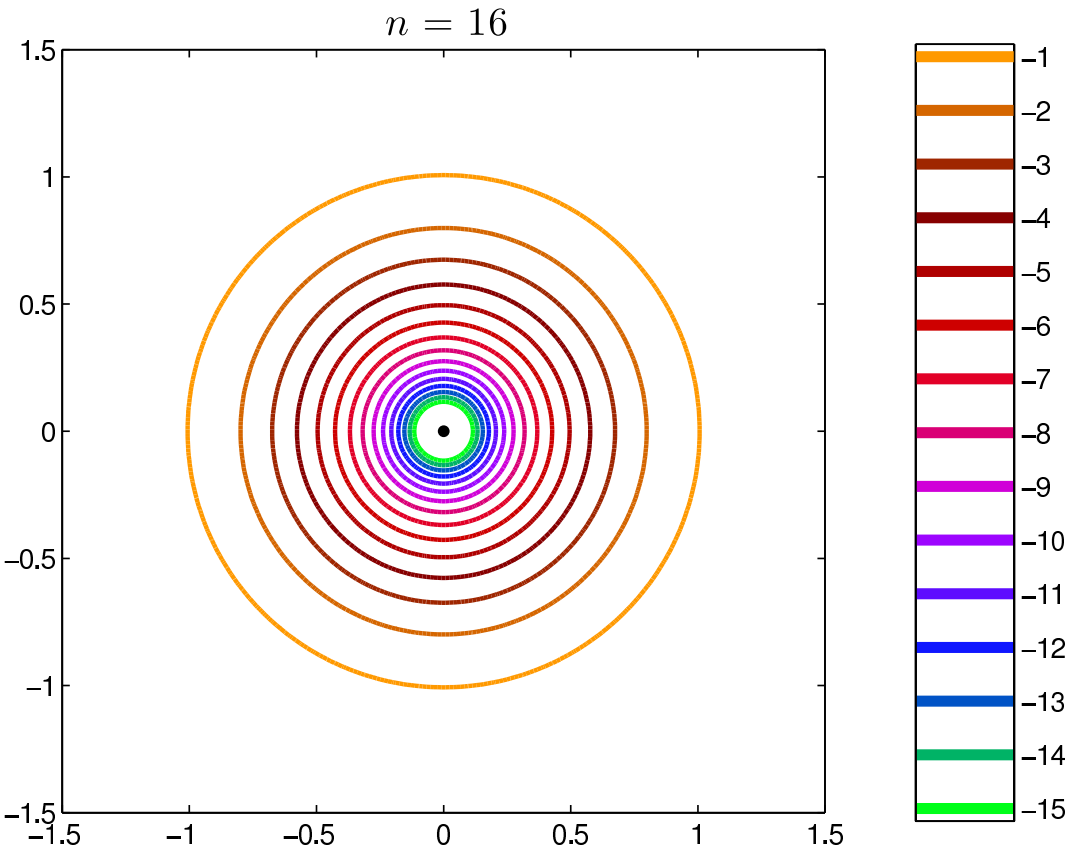
Pseudospectra of the 100×100 matrix $\mathbf{A} = \text{tridiag}(2, 0, 1/2)$ that began our investigation of eigenvalue perturbations.



A is diagonalizable (it has distinct eigenvalues), but Bauer–Fike is useless here: $\kappa(\mathbf{V}) = 2^{99} \approx 6 \times 10^{29}$.

Jordan Blocks

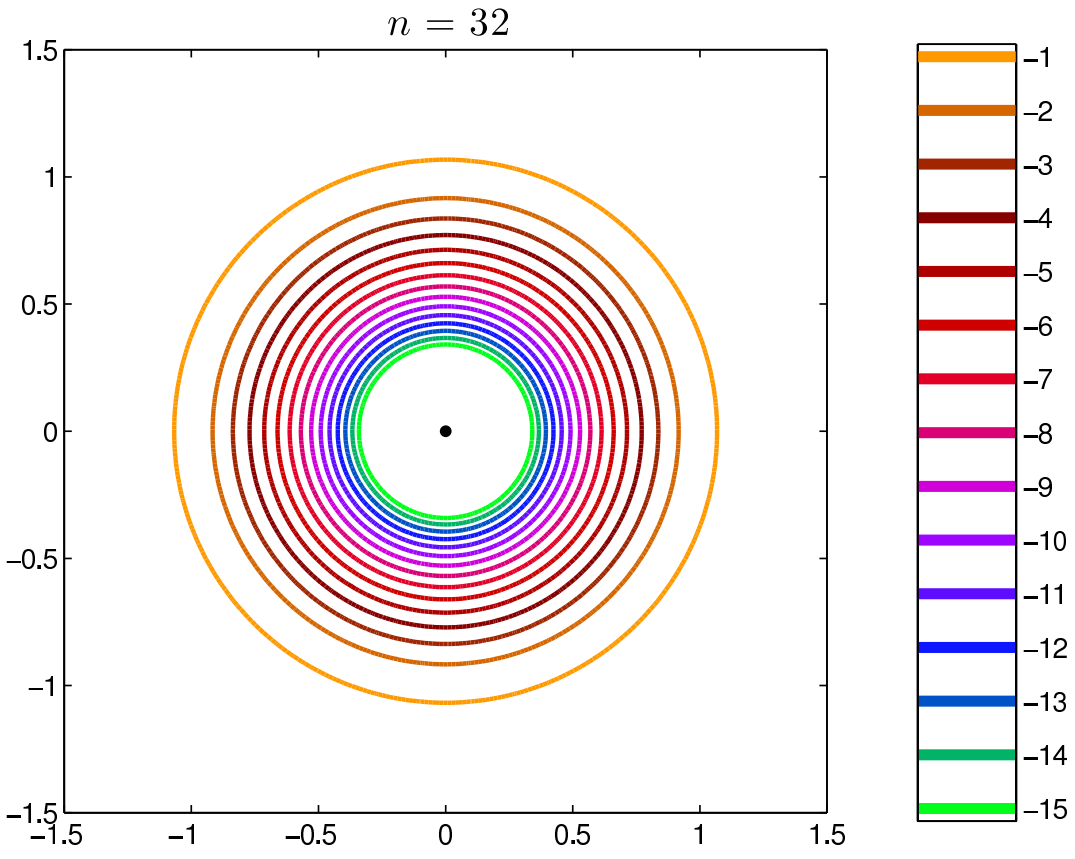
$$\mathbf{S}_n = \begin{bmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix}$$



Near the eigenvalue, the resolvent norm grows with dimension n; outside the unit disk, the resolvent norm does not seem to get big. We would like to prove this.

Jordan Blocks

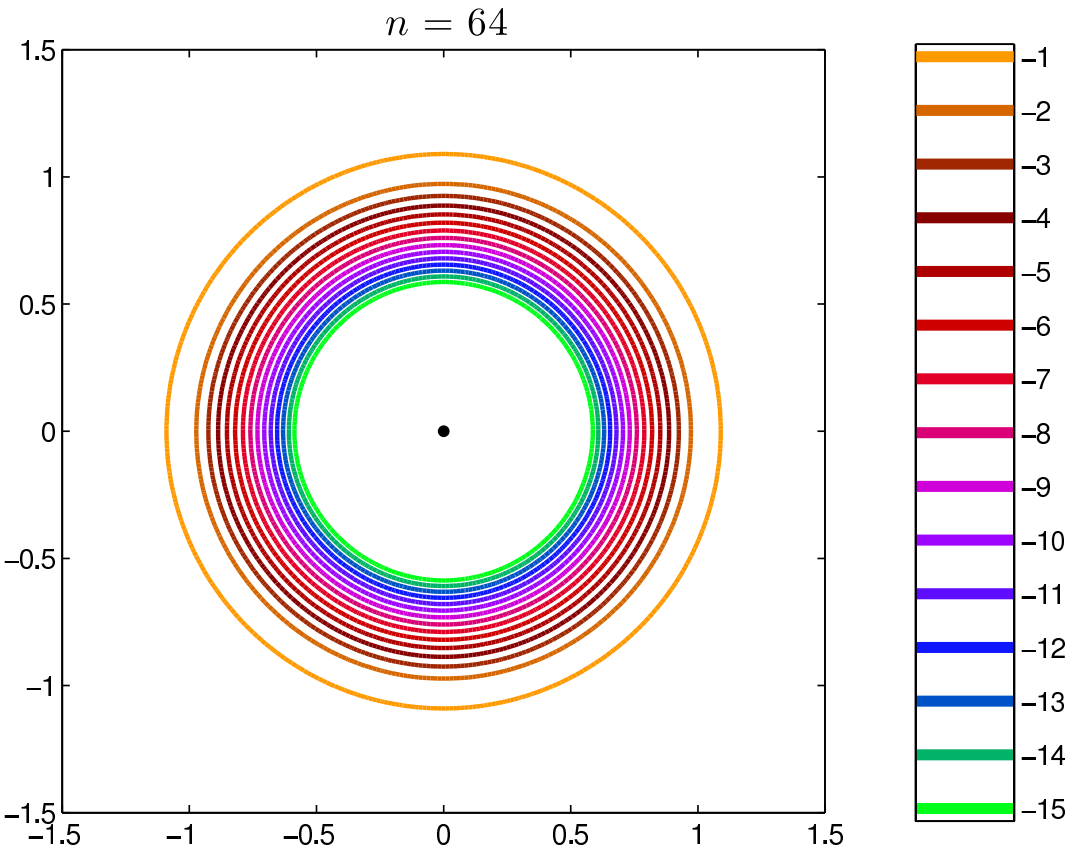
$$\mathbf{S}_n = \begin{bmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix}$$



Near the eigenvalue, the resolvent norm grows with dimension n; outside the unit disk, the resolvent norm does not seem to get big. We would like to prove this.

Jordan Blocks

$$\mathbf{S}_n = \begin{bmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix}$$



Near the eigenvalue, the resolvent norm grows with dimension n; outside the unit disk, the resolvent norm does not seem to get big. We would like to prove this.

Jordan Blocks/Shift Operator

Consider the generalization of the Jordan block to the domain

$$\ell^2(\mathbf{N}) = \{(x_1, x_2, \ldots) : \sum_{j=1}^{\infty} |x_j|^2 < \infty\}.$$

The shift operator $\mathbf{S}$ on $\ell^2(\mathbf{N})$ is defined as

$$\mathbf{S}(x_1, x_2, \ldots) = (x_2, x_3, \ldots).$$

Jordan Blocks/Shift Operator

Consider the generalization of the Jordan block to the domain

$$\ell^2(\mathbf{N}) = \{(x_1, x_2, \ldots) : \sum_{j=1}^{\infty} |x_j|^2 < \infty\}.$$

The shift operator $\mathbf{S}$ on $\ell^2(\mathbf{N})$ is defined as

$$\mathbf{S}(x_1, x_2, \ldots) = (x_2, x_3, \ldots).$$

In particular,

$$\begin{aligned} \mathbf{S}(1, z, z^2, \ldots) &= (z, z^2, z^3, \ldots) \\ &= z(1, z, z^2, \ldots). \end{aligned}$$

So if $(1, z, z^2, \ldots) \in \ell^2(\mathbf{N})$, then $z \in \sigma(\mathbf{S})$.

Jordan Blocks/Shift Operator

Consider the generalization of the Jordan block to the domain

$$\ell^2(\mathbf{N}) = \{(x_1, x_2, \ldots) : \sum_{j=1}^{\infty} |x_j|^2 < \infty\}.$$

The shift operator $\mathbf{S}$ on $\ell^2(\mathbf{N})$ is defined as

$$\mathbf{S}(x_1, x_2, \ldots) = (x_2, x_3, \ldots).$$

In particular,

$$\begin{aligned} \mathbf{S}(1, z, z^2, \ldots) &= (z, z^2, z^3, \ldots) \\ &= z(1, z, z^2, \ldots). \end{aligned}$$

So if $(1, z, z^2, \ldots) \in \ell^2(\mathbf{N})$, then $z \in \sigma(\mathbf{S})$.

If $|z| < 1$, then

$$\sum_{j=1}^{\infty} |z^{j-1}|^2 = \frac{1}{1 - |z|^2} < \infty.$$

So,

$$\{z \in \mathbf{C} : |z| < 1\} \subseteq \sigma(\mathbf{S}).$$

Jordan Blocks/Shift Operator

$$\mathbf{S}(x_1, x_2, \ldots) = (x_2, x_3, \ldots).$$

We have seen that

$$\{z \in \mathbf{C} : |z| < 1\} \subseteq \sigma(\mathbf{S}).$$

Observe that

$$\|\mathbf{S}\| = \sup_{\|\mathbf{x}\|=1} \|\mathbf{S}\mathbf{x}\| = 1,$$

and so

$$\sigma(\mathbf{S}) \subseteq \{z \in \mathbf{C} : |z| \leq 1\}.$$

The spectrum is closed, so

$$\sigma(\mathbf{A}) = \{z \in \mathbf{C} : |z| \leq 1\}.$$

For any finite dimensional $n \times n$ Jordan block $\mathbf{S}_n$,

$$\sigma(\mathbf{S}_n) = \{0\}.$$

So the $\mathbf{S}_n \to \mathbf{S}$ strongly, but there is a discontinuity in the spectrum:

$$\sigma(\mathbf{S}_n) \not\to \sigma(\mathbf{S}).$$

Pseudospectra of Jordan Blocks

Pseudospectra resolve this unpleasant discontinuity.

Recall the eigenvectors $(1, z, z^2, \ldots)$ for $\mathbf{S}$.

Truncate this vector to length n, and apply it to $\mathbf{S}_n$:

$$\begin{bmatrix} 0 & 1 & & & \\ & 0 & \ddots & & \\ & & \ddots & 1 & \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix} \begin{bmatrix} 1 \\ z \\ \vdots \\ z^{n-2} \\ z^{n-1} \end{bmatrix} = \begin{bmatrix} z \\ z^2 \\ \vdots \\ z^{n-1} \\ 0 \end{bmatrix}$$

Pseudospectra of Jordan Blocks

Pseudospectra resolve this unpleasant discontinuity.

Recall the eigenvectors $(1, z, z^2, \ldots)$ for $\mathbf{S}$.

Truncate this vector to length n, and apply it to $\mathbf{S}_n$:

$$\begin{bmatrix} 0 & 1 & & & \\ & 0 & \ddots & & \\ & & \ddots & 1 & \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix} \begin{bmatrix} 1 \\ z \\ \vdots \\ z^{n-2} \\ z^{n-1} \end{bmatrix} = \begin{bmatrix} z \\ z^2 \\ \vdots \\ z^{n-1} \\ 0 \end{bmatrix} = z \begin{bmatrix} 1 \\ z \\ \vdots \\ z^{n-2} \\ z^{n-1} \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ z^n \end{bmatrix}.$$

Pseudospectra of Jordan Blocks

Pseudospectra resolve this unpleasant discontinuity.

Recall the eigenvectors $(1, z, z^2, \ldots)$ for $\mathbf{S}$.

Truncate this vector to length n, and apply it to $\mathbf{S}_n$:

$$\begin{bmatrix} 0 & 1 & & & \\ & 0 & \ddots & & \\ & & \ddots & 1 & \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix} \begin{bmatrix} 1 \\ z \\ \vdots \\ z^{n-2} \\ z^{n-1} \end{bmatrix} = \begin{bmatrix} z \\ z^2 \\ \vdots \\ z^{n-1} \\ 0 \end{bmatrix} = z \begin{bmatrix} 1 \\ z \\ \vdots \\ z^{n-2} \\ z^{n-1} \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ z^n \end{bmatrix}.$$

Hence, $\|\mathbf{S}_n\mathbf{x} - z\mathbf{x}\| = |z|^n$, so for all $\varepsilon > \dfrac{|z|^n}{\|\mathbf{x}\|} = |z|^n \dfrac{\sqrt{1-|z|^{2n}}}{\sqrt{1-|z|^2}}$,

$$z \in \sigma_\varepsilon(\mathbf{S}_n).$$

We conclude that for fixed $|z| < 1$, the resolvent norm $\|(z - \mathbf{S}_n)^{-1}\|$ grows *exponentially* with n.

Pseudospectra of Toeplitz Matrices

Toeplitz matrices are described by their *symbol* a with Laurent expansion

$$a(z) = \sum_{k=-\infty}^{\infty} a_k z^k,$$

giving the matrix with constant diagonals containing the Laurent coefficients:

$$\mathbf{A}_n = \begin{bmatrix} a_0 & a_{-1} & a_{-2} & \cdots & \\ a_1 & a_0 & \ddots & \ddots & \vdots \\ a_2 & \ddots & \ddots & a_{-1} & a_{-2} \\ \vdots & \ddots & a_1 & a_0 & a_{-1} \\ & \cdots & a_2 & a_1 & a_0 \end{bmatrix} \in \mathbf{C}^{n\times n}.$$

Call the image of the unit circle $\mathbf{T}$ under a the *symbol curve*, $a(\mathbf{T})$.

Pseudospectra of Toeplitz Matrices

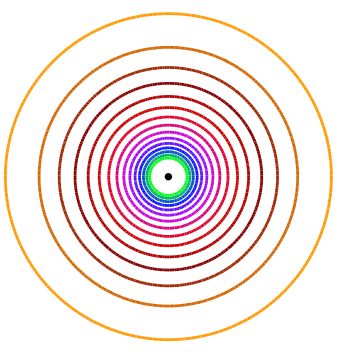
Theorem (Landau; Reichel and Trefethen; Böttcher)

Let $a(z) = \sum_{k=-M}^{M} a_k z^k$ be the symbol of a banded Toepiltz operator.

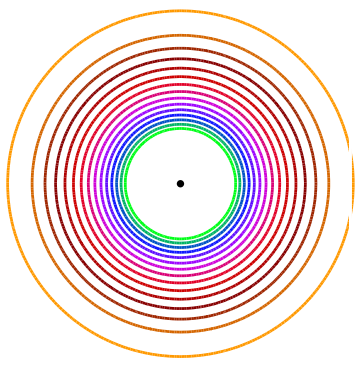
- *The pseudospectra of $\mathbf{A}_n$ converge to the pseudospectra of the Toeplitz operator on $\ell^2(\mathbf{N})$ as $n \to \infty$.*

Let $z \in C$ have nonzero winding number w.r.t. the symbol curve $a(T)$.

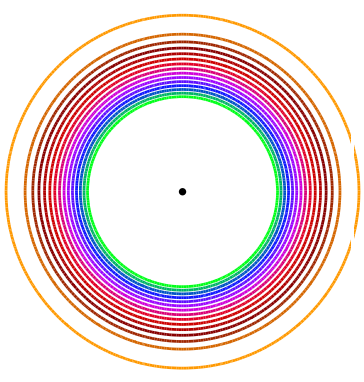
- $\|(z - \mathbf{A}_n)^{-1}\|$ *grows exponentially in n.*
- *For all $\varepsilon > 0$, $z \in \sigma_\varepsilon(\mathbf{A}_n)$ for all n sufficiently large.*



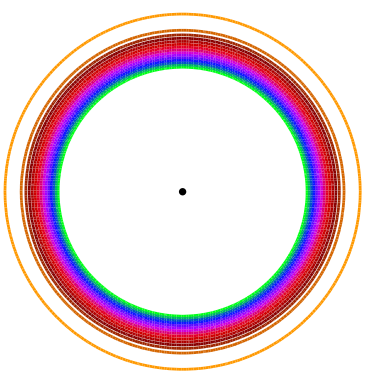
$n = 16$



$n = 32$

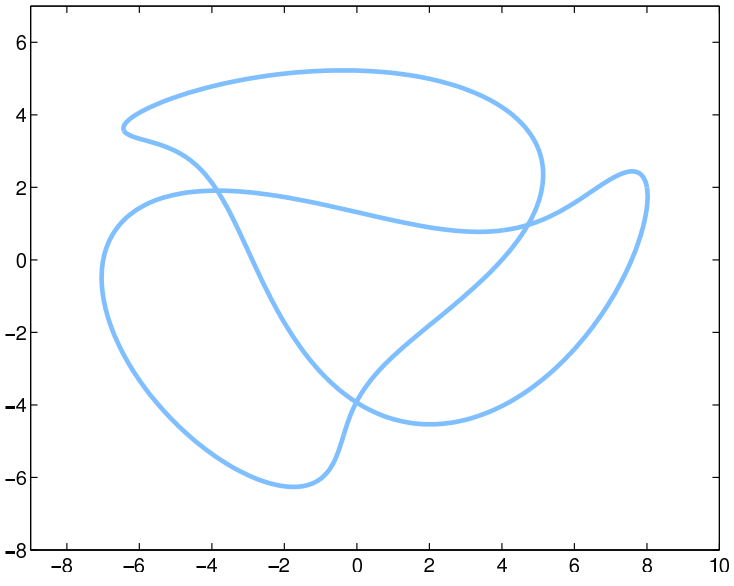


$n = 64$

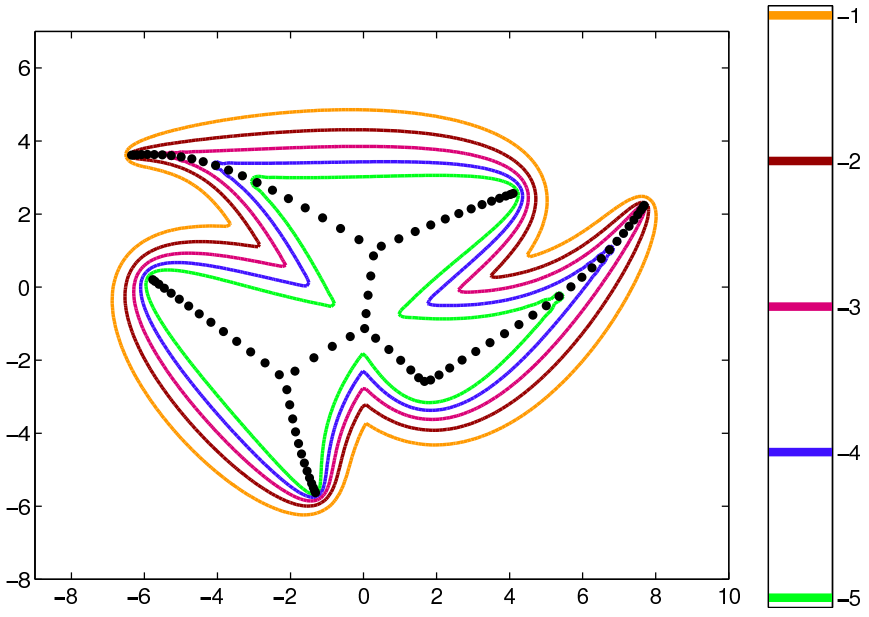


$n = 128$

Pseudospectra of Toeplitz Matrices

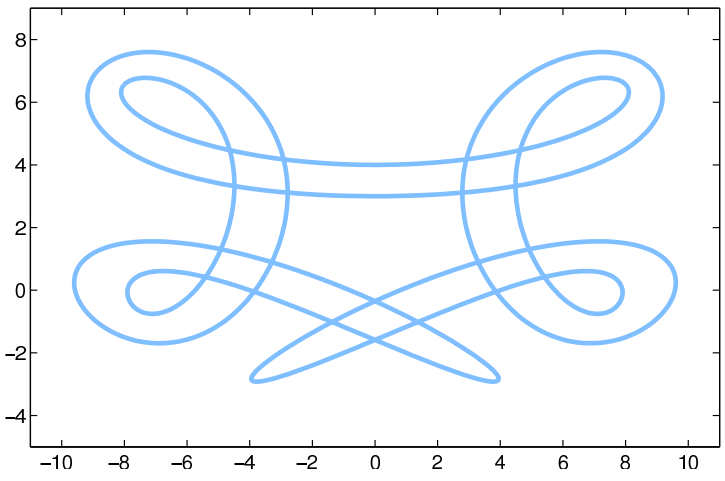


symbol curve

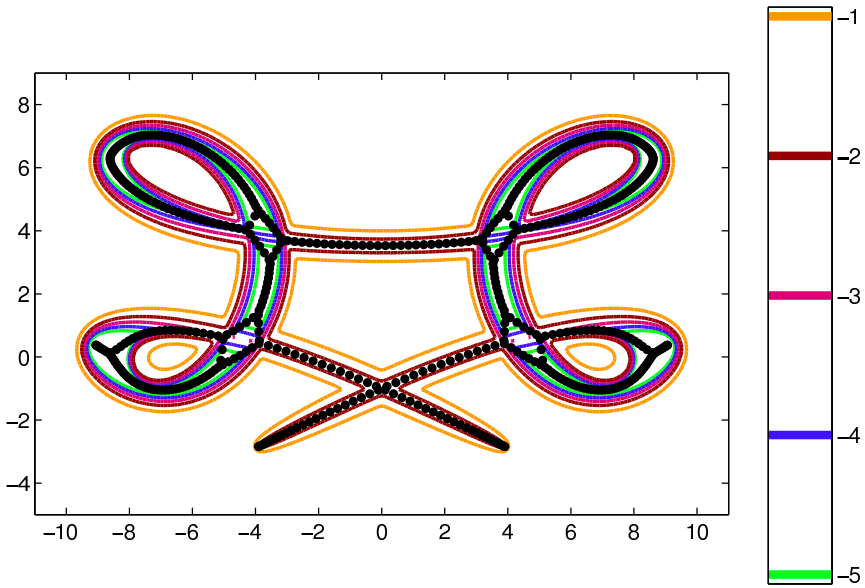


$\sigma_\varepsilon(\mathbf{A}_{100})$

Pseudospectra of Toeplitz Matrices



symbol curve



$\sigma_\varepsilon(\mathbf{A}_{500})$

Computation of Pseudospectra

Naive algorithm: $\mathcal{O}(n^3)$ per grid point

- Compute $\|(z - \mathbf{A})^{-1}\|$ using the SVD on a grid of points in $\mathbf{C}$.
- Send data to a contour plotting routine.

Modern algorithm: $\mathcal{O}(n^3) + \mathcal{O}(n^2)$ per grid point [Lui 1997; Trefethen 1999]

- Compute a Schur factorization $\mathbf{A} = \mathbf{UTU}^*$ and compute $\sigma_\varepsilon(\mathbf{T})$.
- Compute $\|(z - \mathbf{T})^{-1}\|$ on grid via inverse Lanczos on $(z - \mathbf{T})^{-*}(z - \mathbf{T})^{-1}$.

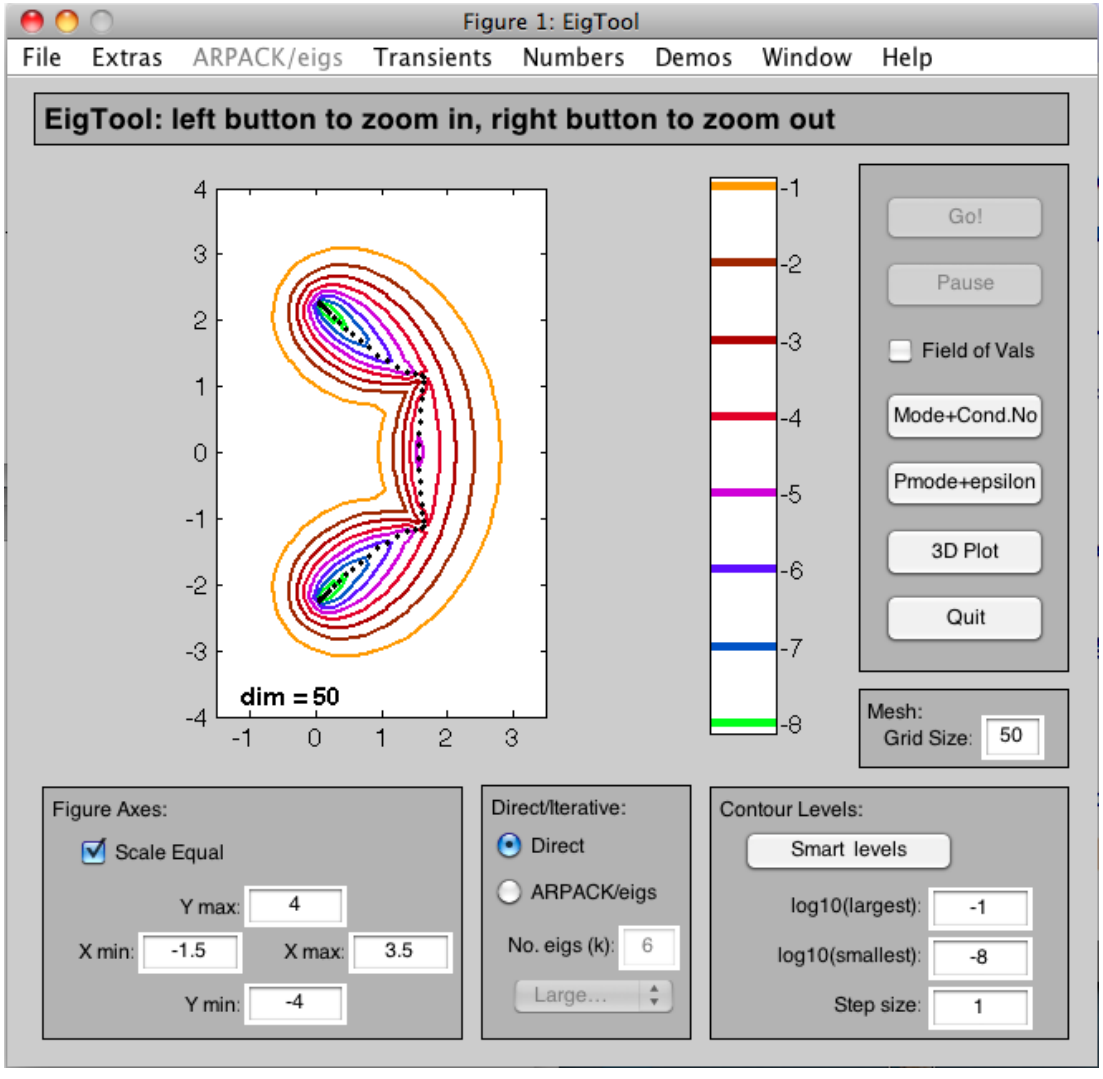
Large-scale problems: [Toh and Trefethen 1996; Wright and Trefethen 2001]

- Pick an invariant subspace Ran($\mathbf{V}$) for $\mathbf{V}^*\mathbf{V} = \mathbf{I}$ corresponding to eigenvalues of physical interest (e.g., using ARPACK).
- Compute $\sigma_\varepsilon(\mathbf{V}^*\mathbf{AV}) \subseteq \sigma_\varepsilon(\mathbf{A})$.

Alternative: [Brühl 1996; Bekas and Gallopoulos, . . .]

- Curve tracing: follow level sets of $\|(z - \mathbf{A})^{-1}\|$.

EigTool



EigTool: Thomas Wright, 2002

http://www.cs.ox.ac.uk/pseudospectra/eigtool

Pole Placement Example

Problem (Pole Placement Example of Mehrmann and Xu, 1996)

Given $\mathbf{A} = \mathrm{diag}(1, 2, \ldots, N)$ *and* $\mathbf{b} = [1, 1, \ldots, 1]^T$, *find* $\mathbf{f} \in C^n$ *such that*

$$\sigma(\mathbf{A} - \mathbf{b}\mathbf{f}^*) = \{-1, -2, \ldots, -N\}.$$

One can show that $\mathbf{f} = \mathbf{G}^{-*}\mathbf{e}$, where $\mathbf{e} = [1, 1, \ldots, 1]^T$, where

$$\mathbf{G}_{:,k} = (\mathbf{A} - \lambda_k)^{-1}\mathbf{b} = (\mathbf{A} + k)^{-1}\mathbf{b}.$$

This gives

$$\mathbf{A} - \mathbf{b}\mathbf{f}^* = \mathbf{G} \begin{bmatrix} -1 & & \\ & \ddots & \\ & & -n \end{bmatrix} \mathbf{G}^{-1},$$

with

$$\mathbf{G}_{j,k} = \frac{1}{j + k}.$$

Pole Placement Example: Numerics

In exact arithmetic, $\sigma(\mathbf{A} - \mathbf{b}\mathbf{f}^*) = \{-1, -2, \ldots, -N\}$.

All entries in $\mathbf{A} - \mathbf{b}\mathbf{f}^*$ are integers.
(To ensure this, we compute $\mathbf{f}$ *symbolically*.)

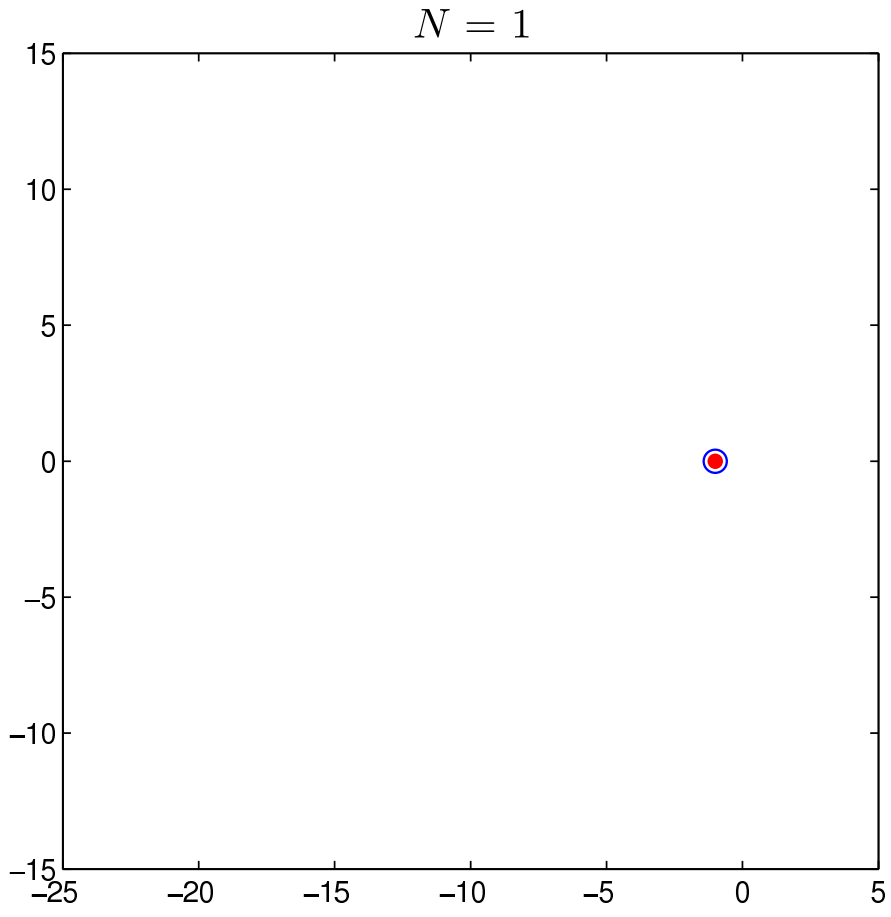
For example, when $N = 8$,

$$\mathbf{A}-\mathbf{b}\mathbf{f}^* = \begin{bmatrix} 73 & -2520 & 27720 & -138600 & 360360 & -504504 & 360360 & -102960 \\ 72 & -2518 & 27720 & -138600 & 360360 & -504504 & 360360 & -102960 \\ 72 & -2520 & 27723 & -138600 & 360360 & -504504 & 360360 & -102960 \\ 72 & -2520 & 27720 & -138596 & 360360 & -504504 & 360360 & -102960 \\ 72 & -2520 & 27720 & -138600 & 360365 & -504504 & 360360 & -102960 \\ 72 & -2520 & 27720 & -138600 & 360360 & -504498 & 360360 & -102960 \\ 72 & -2520 & 27720 & -138600 & 360360 & -504504 & 360367 & -102960 \\ 72 & -2520 & 27720 & -138600 & 360360 & -504504 & 360360 & -102952 \end{bmatrix}.$$

Pole Placement Example: Numerics

In exact arithmetic, $\sigma(\mathbf{A} - \mathbf{b}\mathbf{f}^*) = \{-1, -2, \ldots, -N\}$.

Computed eigenvalues of $\mathbf{A} - \mathbf{b}\mathbf{f}^*$ (matrix is *exact*: all integer entries).



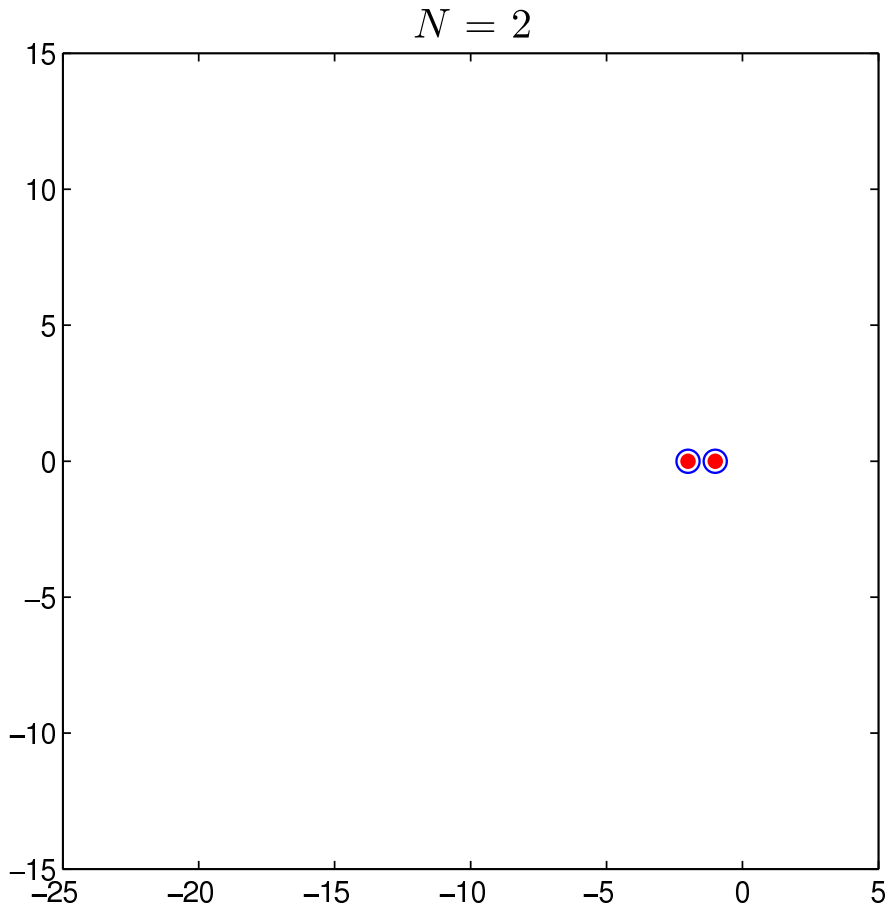
○ exact eigenvalues

● computed eigenvalues

Pole Placement Example: Numerics

In exact arithmetic, $\sigma(\mathbf{A} - \mathbf{b}\mathbf{f}^*) = \{-1, -2, \ldots, -N\}$.

Computed eigenvalues of $\mathbf{A} - \mathbf{b}\mathbf{f}^*$ (matrix is *exact*: all integer entries).



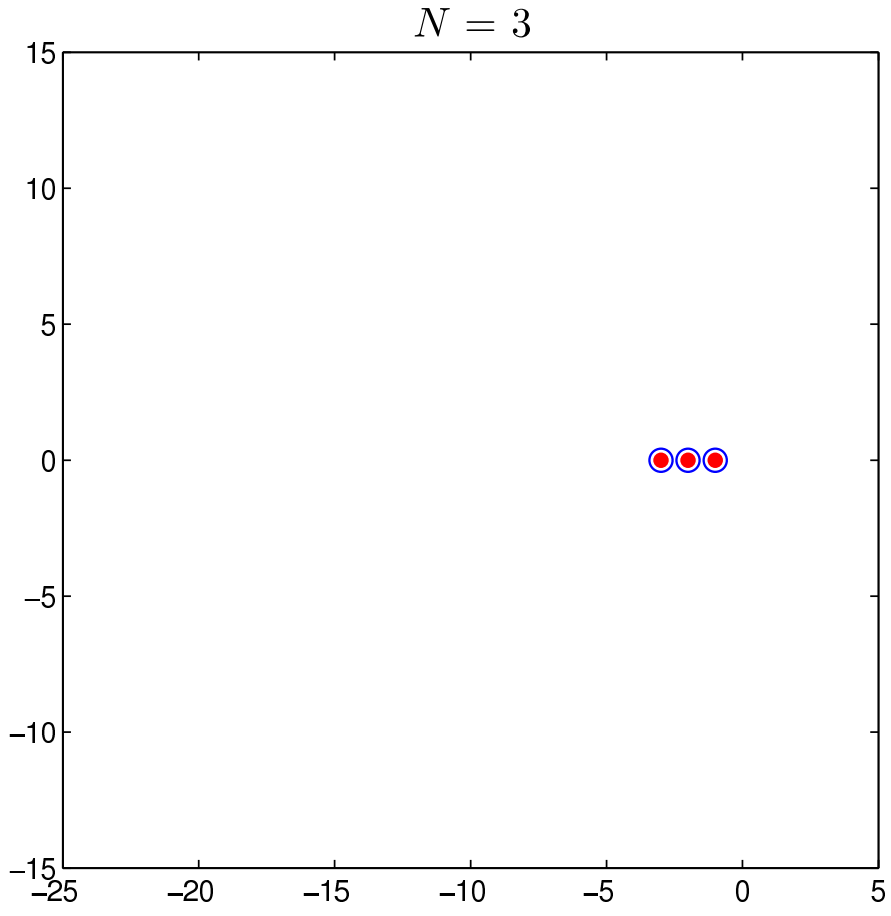
○ exact eigenvalues

● computed eigenvalues

Pole Placement Example: Numerics

In exact arithmetic, $\sigma(\mathbf{A} - \mathbf{b}\mathbf{f}^*) = \{-1, -2, \ldots, -N\}$.

Computed eigenvalues of $\mathbf{A} - \mathbf{b}\mathbf{f}^*$ (matrix is *exact*: all integer entries).



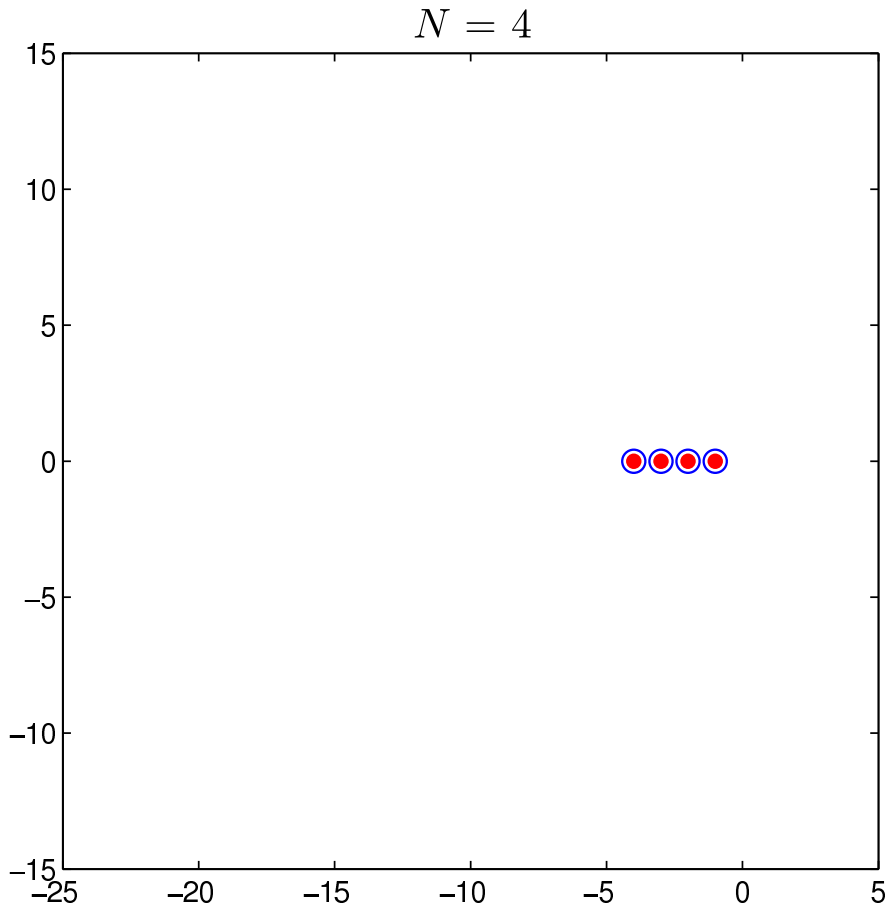
○ exact eigenvalues

● computed eigenvalues

Pole Placement Example: Numerics

In exact arithmetic, $\sigma(\mathbf{A} - \mathbf{b}\mathbf{f}^*) = \{-1, -2, \ldots, -N\}$.

Computed eigenvalues of $\mathbf{A} - \mathbf{b}\mathbf{f}^*$ (matrix is *exact*: all integer entries).



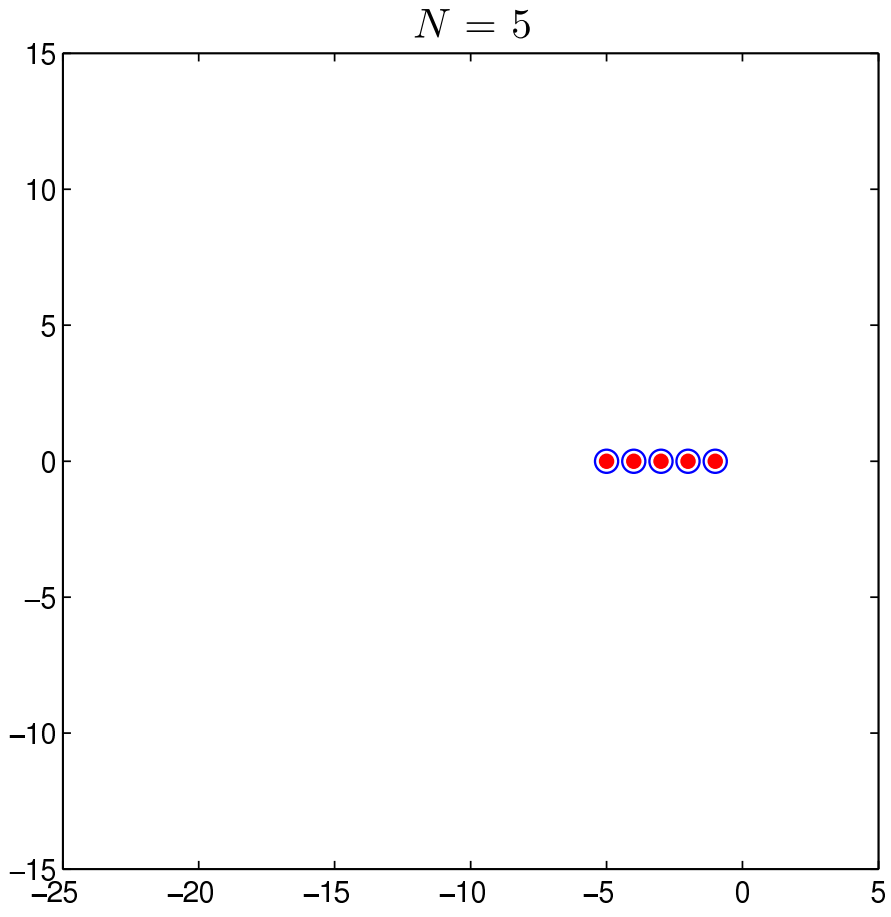
○ exact eigenvalues

● computed eigenvalues

Pole Placement Example: Numerics

In exact arithmetic, $\sigma(\mathbf{A} - \mathbf{b}\mathbf{f}^*) = \{-1, -2, \ldots, -N\}$.

Computed eigenvalues of $\mathbf{A} - \mathbf{b}\mathbf{f}^*$ (matrix is *exact*: all integer entries).



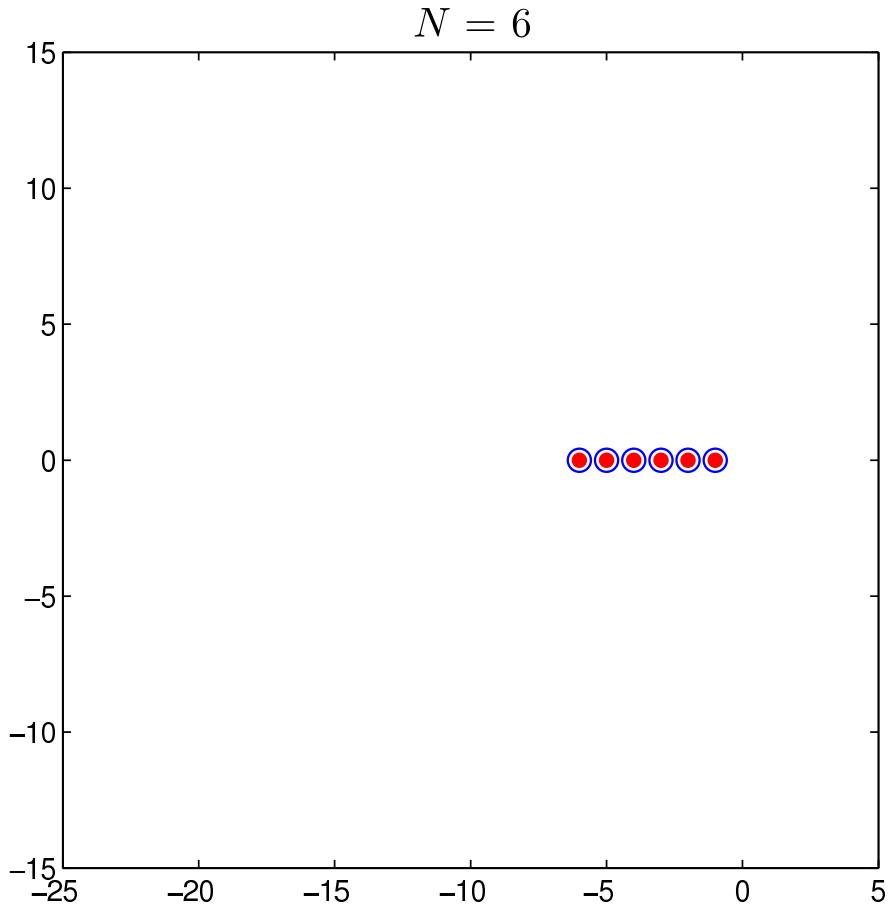
○ exact eigenvalues

● computed eigenvalues

Pole Placement Example: Numerics

In exact arithmetic, $\sigma(\mathbf{A} - \mathbf{b}\mathbf{f}^*) = \{-1, -2, \ldots, -N\}$.

Computed eigenvalues of $\mathbf{A} - \mathbf{b}\mathbf{f}^*$ (matrix is *exact*: all integer entries).



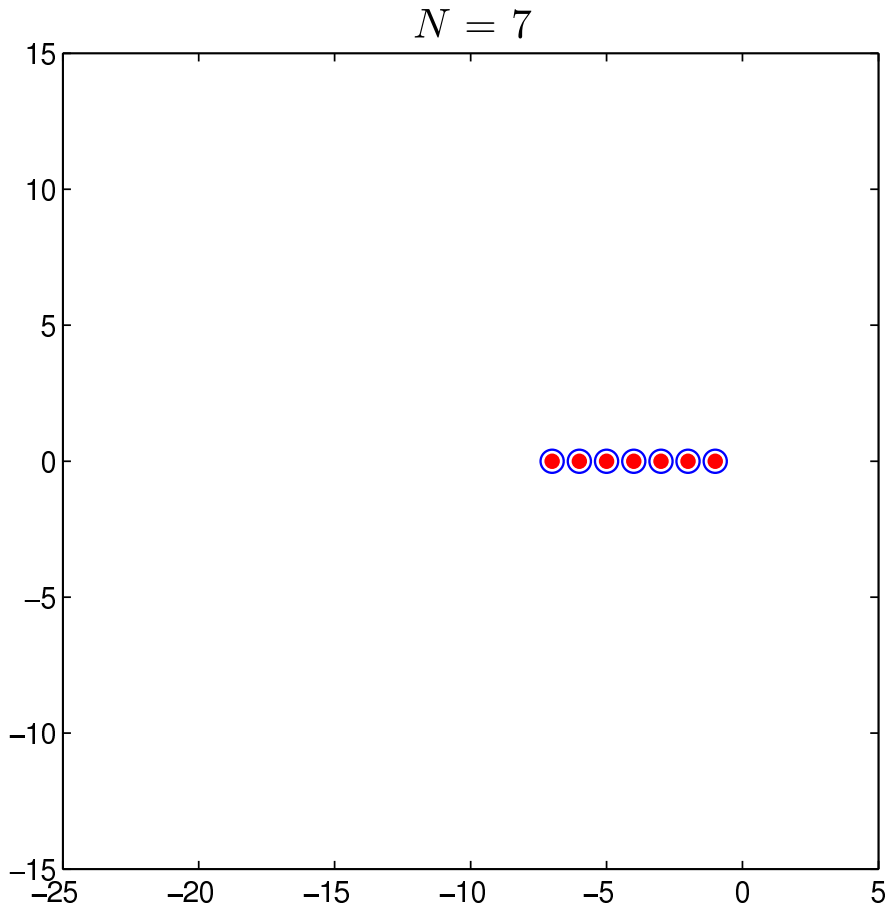
○ exact eigenvalues

● computed eigenvalues

Pole Placement Example: Numerics

In exact arithmetic, $\sigma(\mathbf{A} - \mathbf{b}\mathbf{f}^*) = \{-1, -2, \ldots, -N\}$.

Computed eigenvalues of $\mathbf{A} - \mathbf{b}\mathbf{f}^*$ (matrix is *exact*: all integer entries).



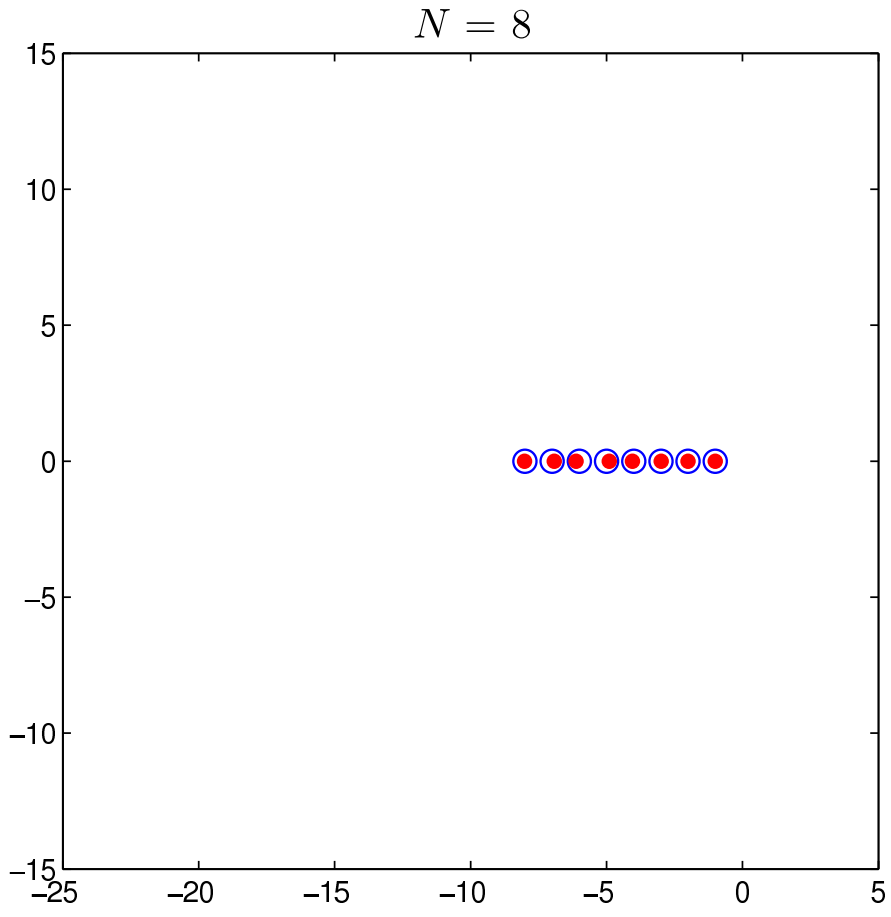
○ exact eigenvalues

● computed eigenvalues

Pole Placement Example: Numerics

In exact arithmetic, $\sigma(\mathbf{A} - \mathbf{b}\mathbf{f}^*) = \{-1, -2, \ldots, -N\}$.

Computed eigenvalues of $\mathbf{A} - \mathbf{b}\mathbf{f}^*$ (matrix is *exact*: all integer entries).



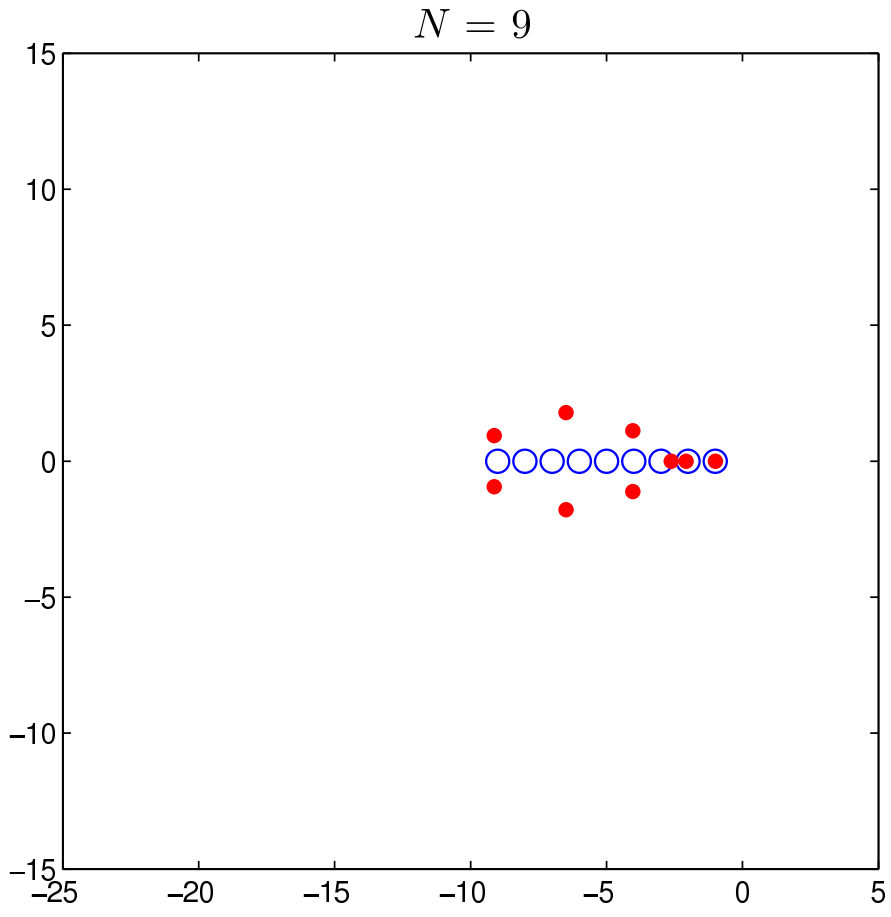
○ exact eigenvalues

● computed eigenvalues

Pole Placement Example: Numerics

In exact arithmetic, $\sigma(\mathbf{A} - \mathbf{b}\mathbf{f}^*) = \{-1, -2, \ldots, -N\}$.

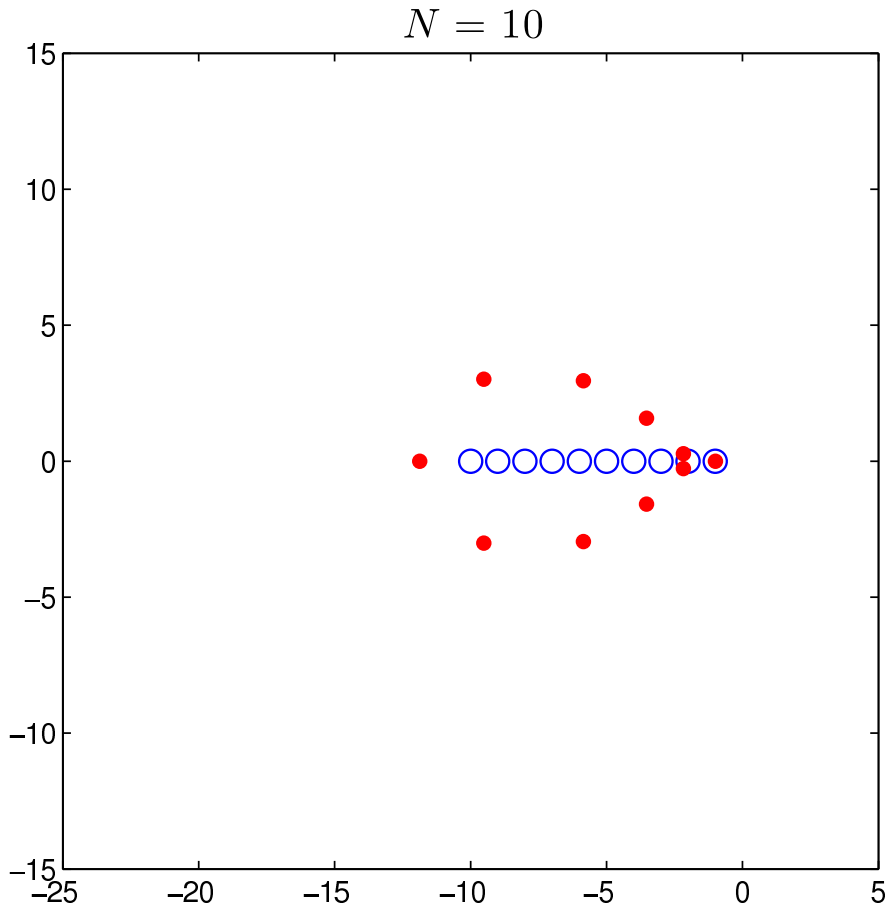
Computed eigenvalues of $\mathbf{A} - \mathbf{b}\mathbf{f}^*$ (matrix is *exact*: all integer entries).



Pole Placement Example: Numerics

In exact arithmetic, $\sigma(\mathbf{A} - \mathbf{b}\mathbf{f}^*) = \{-1, -2, \ldots, -N\}$.

Computed eigenvalues of $\mathbf{A} - \mathbf{b}\mathbf{f}^*$ (matrix is *exact*: all integer entries).



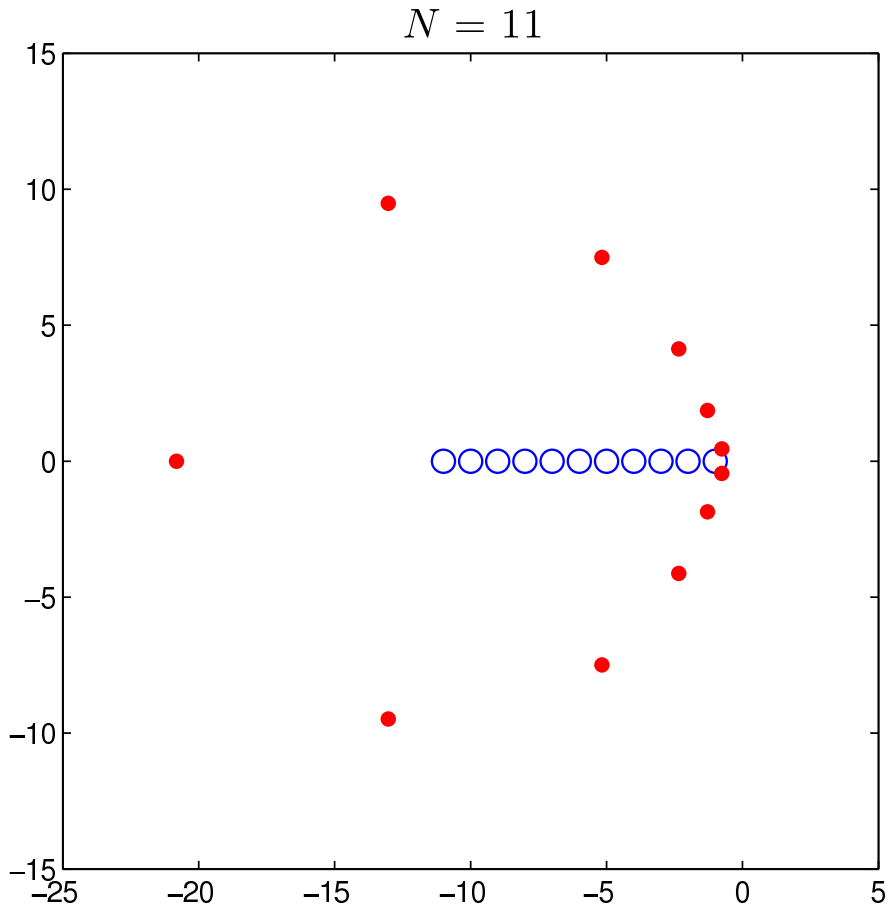
○ exact eigenvalues

● computed eigenvalues

Pole Placement Example: Numerics

In exact arithmetic, $\sigma(\mathbf{A} - \mathbf{b}\mathbf{f}^*) = \{-1, -2, \ldots, -N\}$.

Computed eigenvalues of $\mathbf{A} - \mathbf{b}\mathbf{f}^*$ (matrix is *exact*: all integer entries).

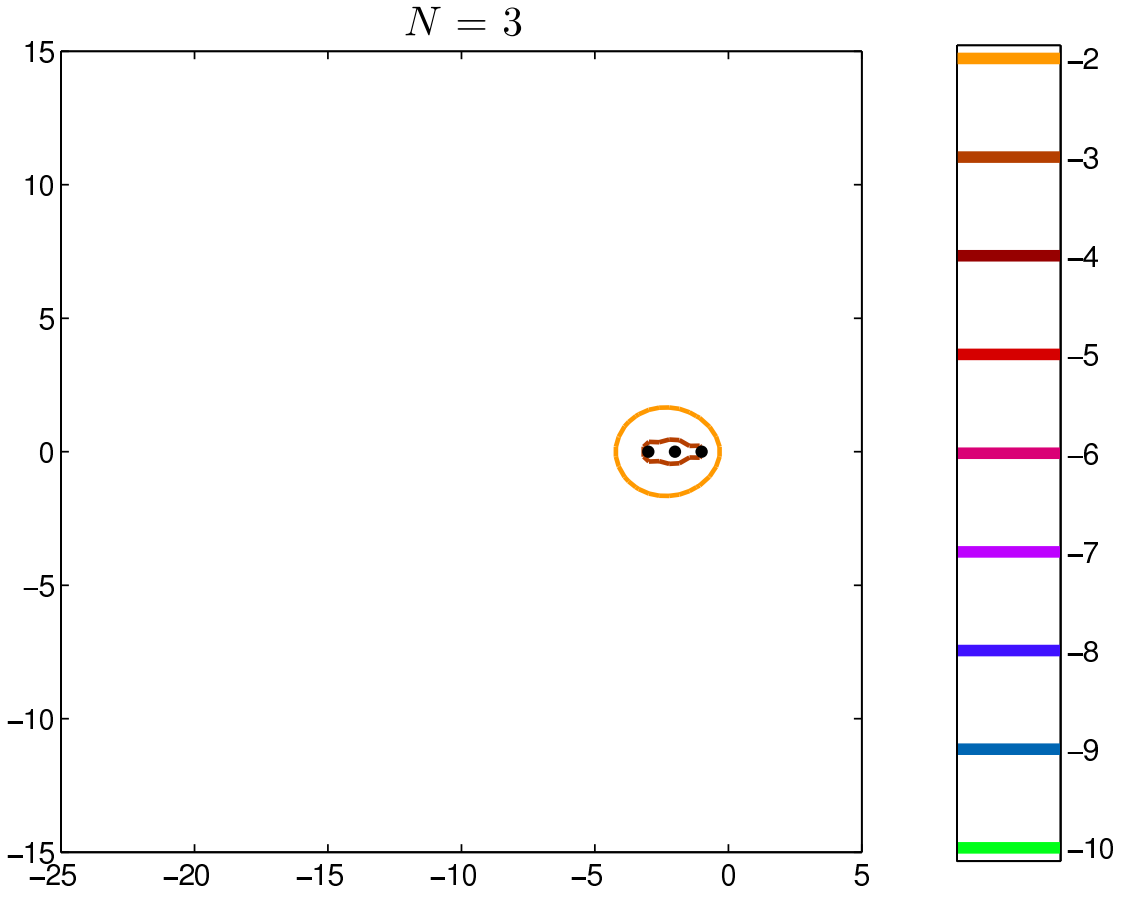


○ exact eigenvalues

● computed eigenvalues

Pole Placement Example: Pseudospectra

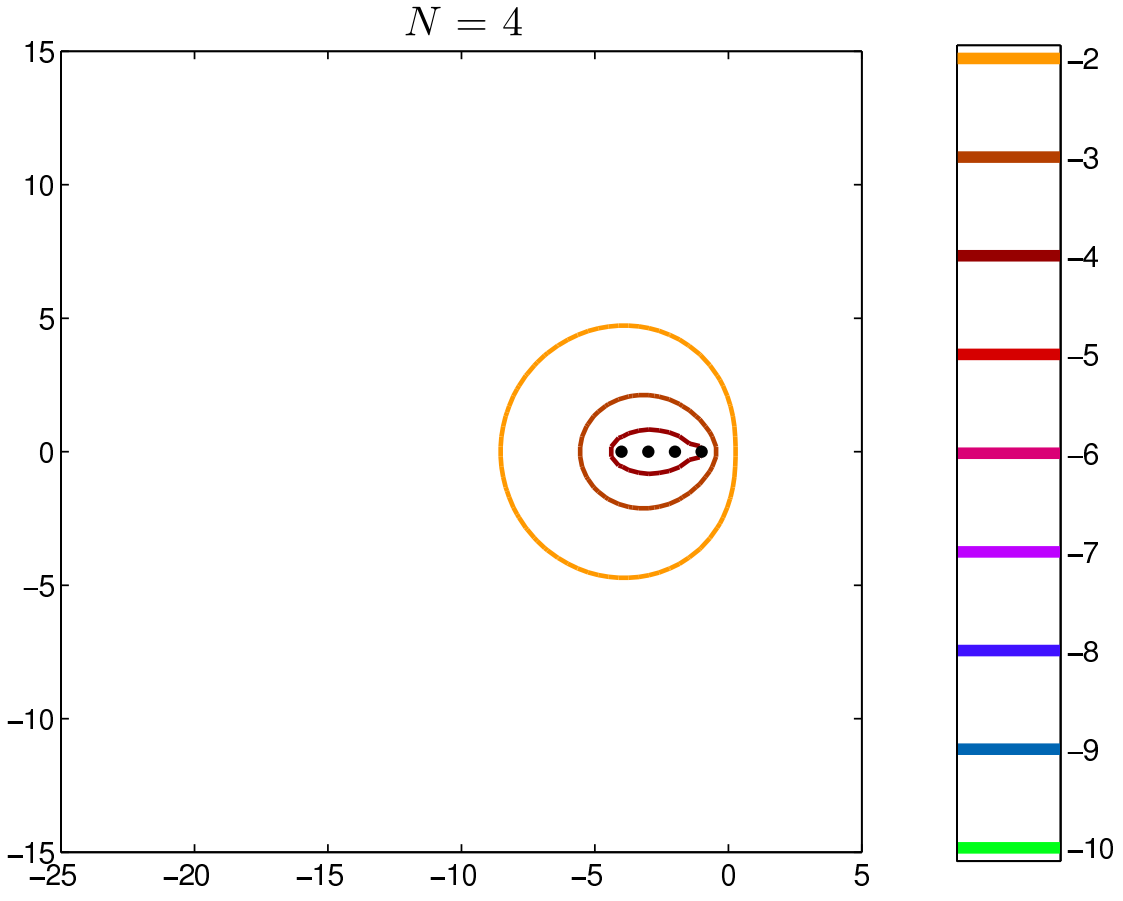
Computed pseudospectra of $\mathbf{A} - \mathbf{b}\mathbf{f}^*$.



Computed eigenvalues should be accurate to roughly $\varepsilon_{\text{mach}} \|\mathbf{A} - \mathbf{b}\mathbf{f}^*\|$.
For example, when $N = 11$, $\|\mathbf{A} - \mathbf{b}\mathbf{f}^*\| = 5.26 \times 10^8$.

Pole Placement Example: Pseudospectra

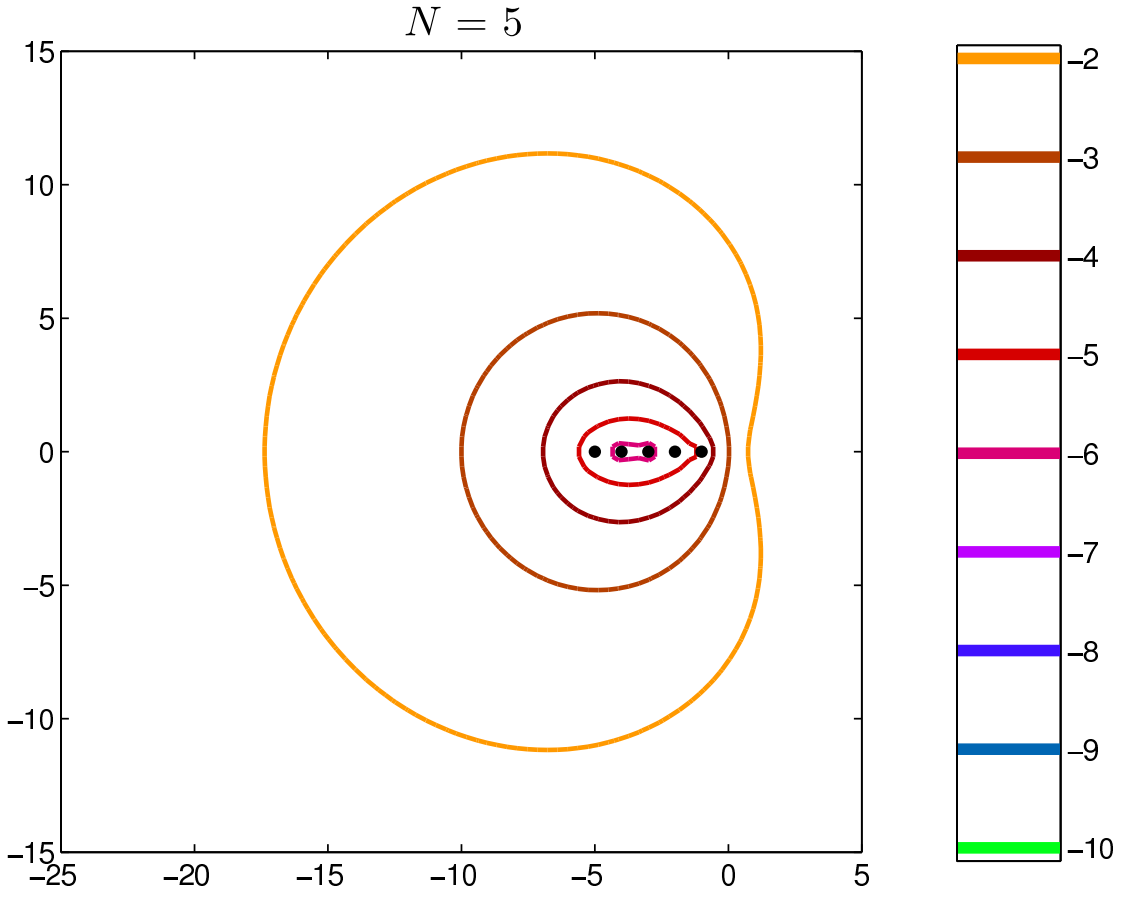
Computed pseudospectra of $\mathbf{A} - \mathbf{b}\mathbf{f}^*$.



Computed eigenvalues should be accurate to roughly $\varepsilon_{\text{mach}} \|\mathbf{A} - \mathbf{b}\mathbf{f}^*\|$.
For example, when $N = 11$, $\|\mathbf{A} - \mathbf{b}\mathbf{f}^*\| = 5.26 \times 10^8$.

Pole Placement Example: Pseudospectra

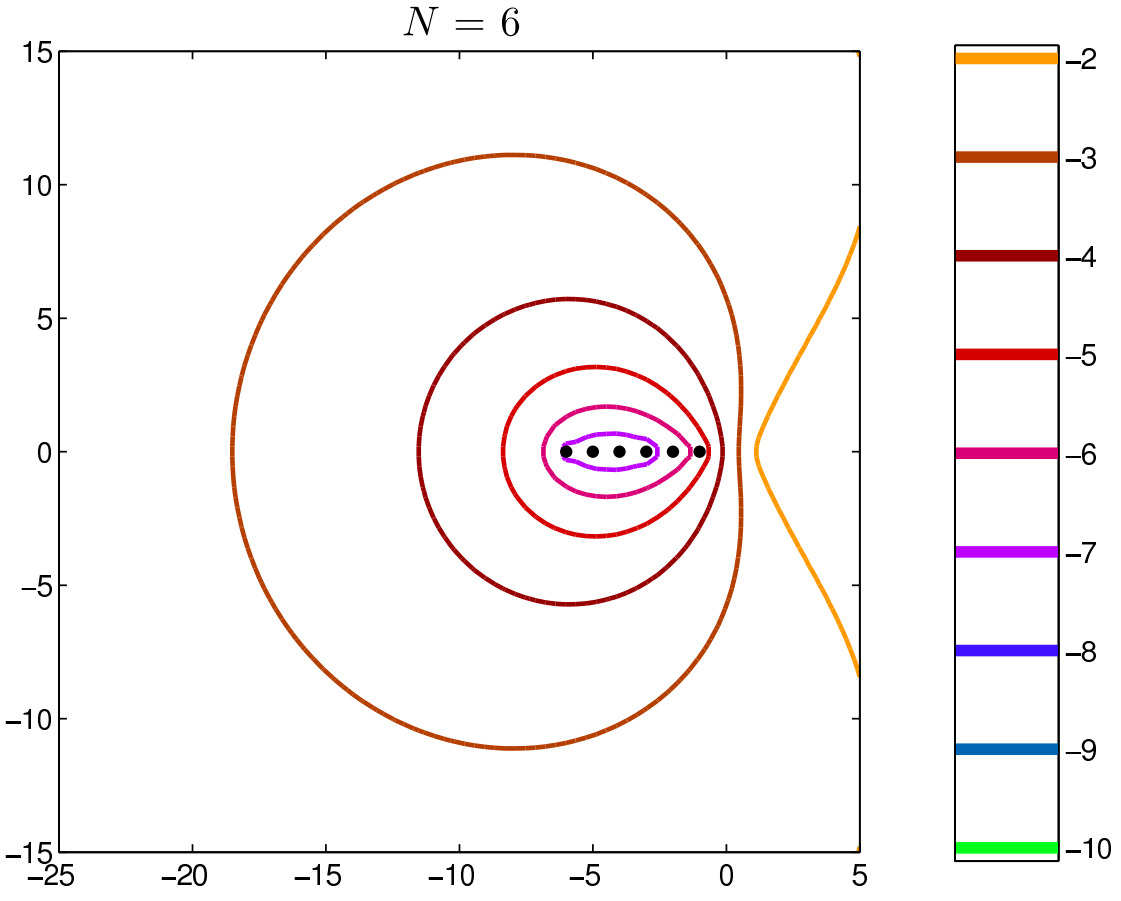
Computed pseudospectra of $\mathbf{A} - \mathbf{b}\mathbf{f}^*$.



Computed eigenvalues should be accurate to roughly $\varepsilon_{\text{mach}} \|\mathbf{A} - \mathbf{b}\mathbf{f}^*\|$.
For example, when $N = 11$, $\|\mathbf{A} - \mathbf{b}\mathbf{f}^*\| = 5.26 \times 10^8$.

Pole Placement Example: Pseudospectra

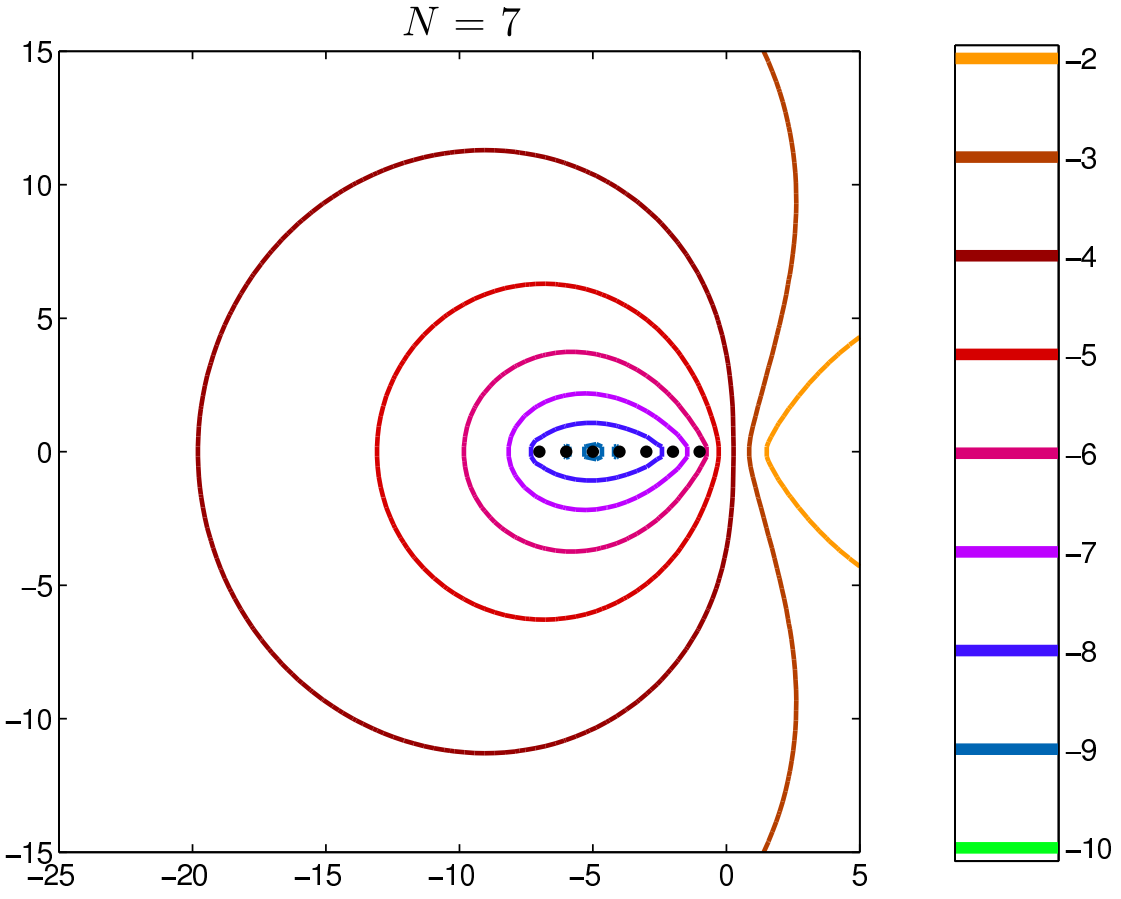
Computed pseudospectra of $\mathbf{A} - \mathbf{b}\mathbf{f}^*$.



Computed eigenvalues should be accurate to roughly $\varepsilon_{\text{mach}} \|\mathbf{A} - \mathbf{b}\mathbf{f}^*\|$.
For example, when $N = 11$, $\|\mathbf{A} - \mathbf{b}\mathbf{f}^*\| = 5.26 \times 10^8$.

Pole Placement Example: Pseudospectra

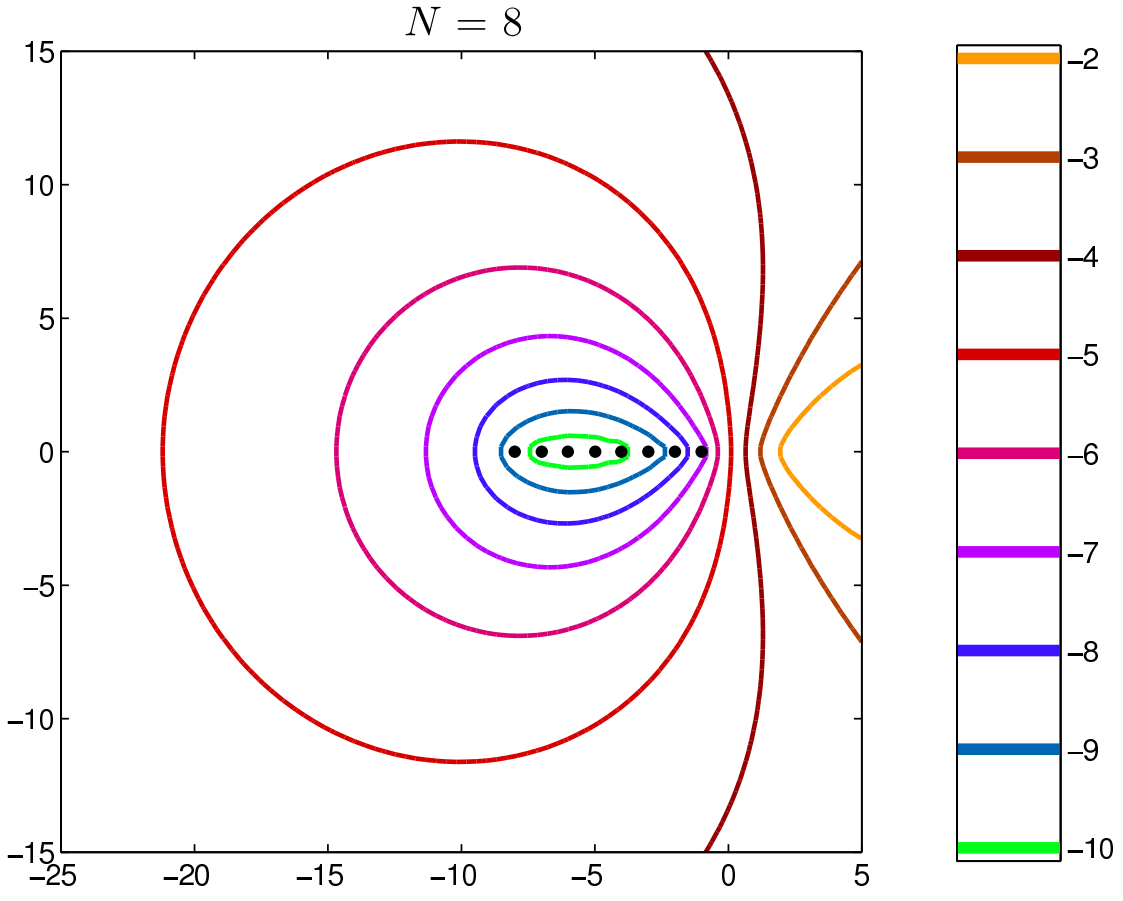
Computed pseudospectra of $\mathbf{A} - \mathbf{b}\mathbf{f}^*$.



Computed eigenvalues should be accurate to roughly $\varepsilon_{\text{mach}} \|\mathbf{A} - \mathbf{b}\mathbf{f}^*\|$.
For example, when $N = 11$, $\|\mathbf{A} - \mathbf{b}\mathbf{f}^*\| = 5.26 \times 10^8$.

Pole Placement Example: Pseudospectra

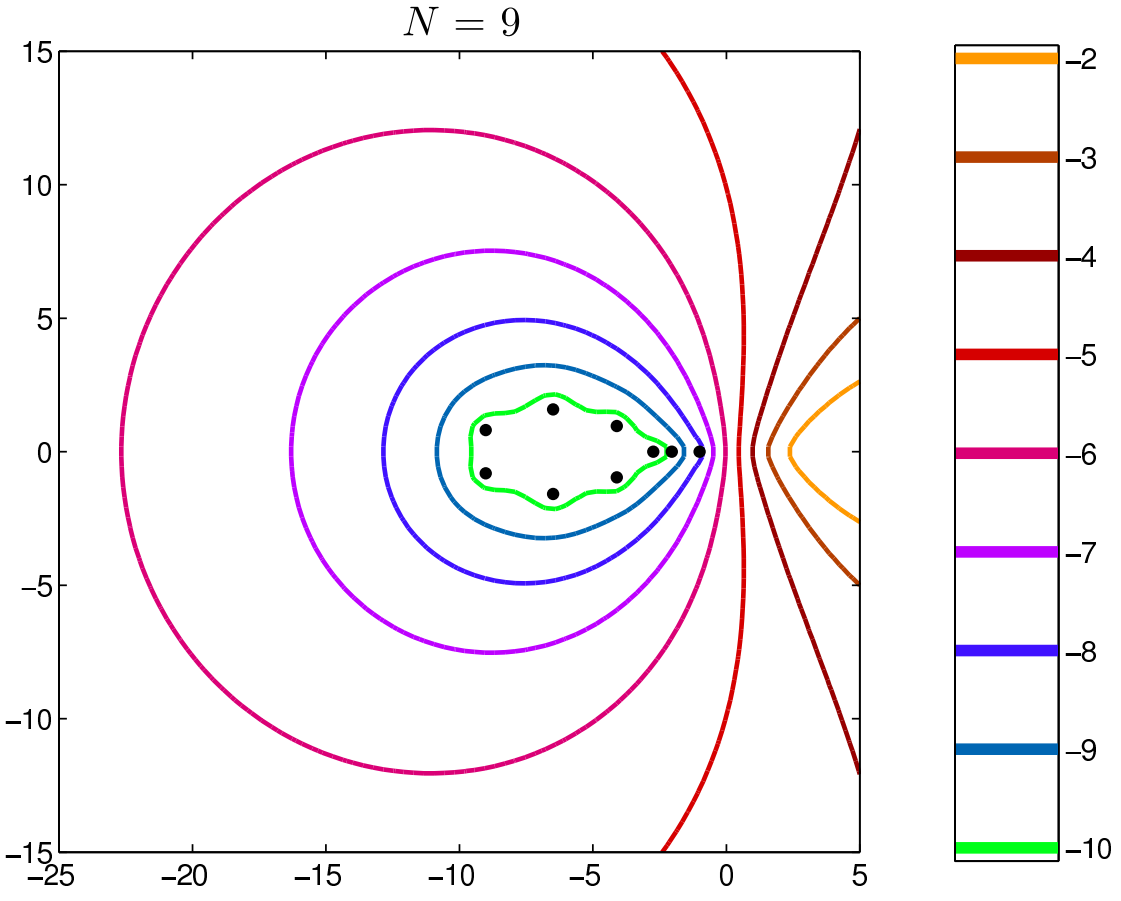
Computed pseudospectra of $\mathbf{A} - \mathbf{b}\mathbf{f}^*$.



Computed eigenvalues should be accurate to roughly $\varepsilon_{\text{mach}} \|\mathbf{A} - \mathbf{b}\mathbf{f}^*\|$.
For example, when $N = 11$, $\|\mathbf{A} - \mathbf{b}\mathbf{f}^*\| = 5.26 \times 10^8$.

Pole Placement Example: Pseudospectra

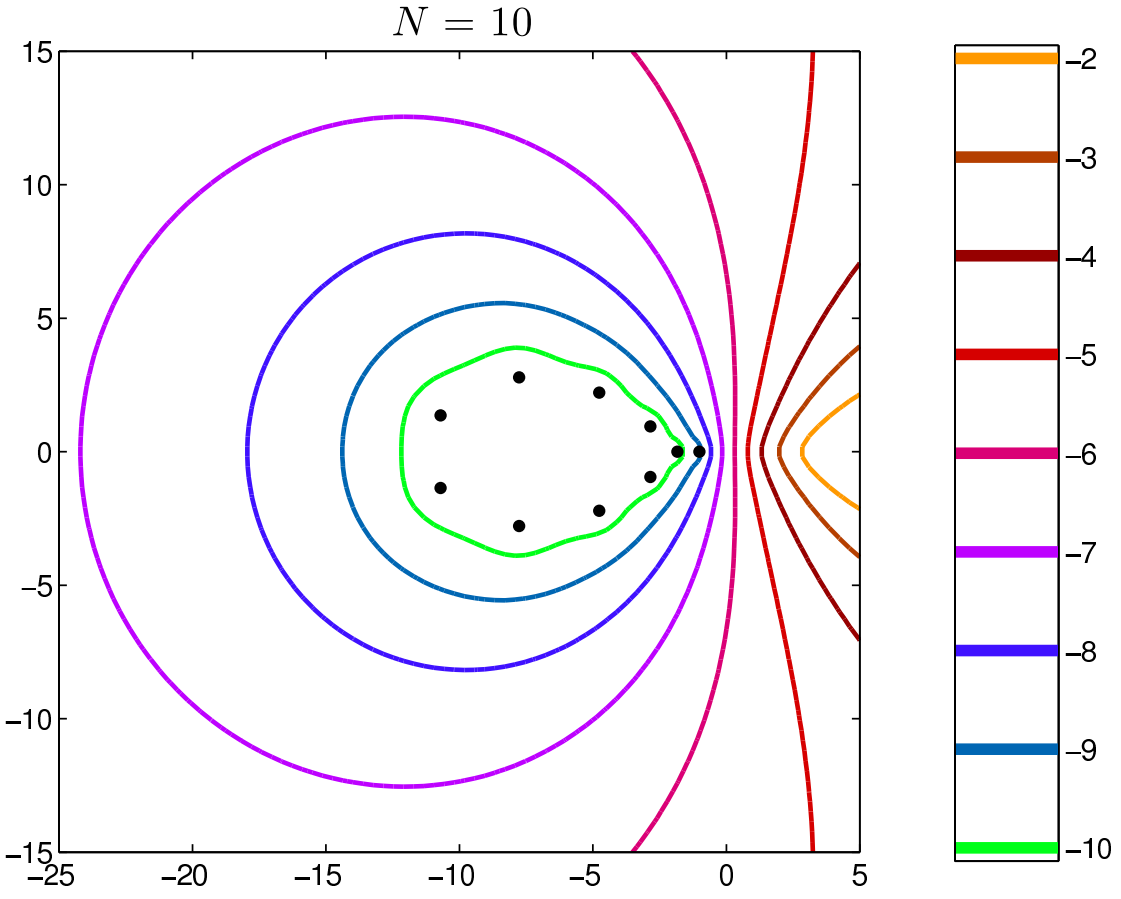
Computed pseudospectra of $\mathbf{A} - \mathbf{b}\mathbf{f}^*$.



Computed eigenvalues should be accurate to roughly $\varepsilon_{\text{mach}} \|\mathbf{A} - \mathbf{b}\mathbf{f}^*\|$.
For example, when $N = 11$, $\|\mathbf{A} - \mathbf{b}\mathbf{f}^*\| = 5.26 \times 10^8$.

Pole Placement Example: Pseudospectra

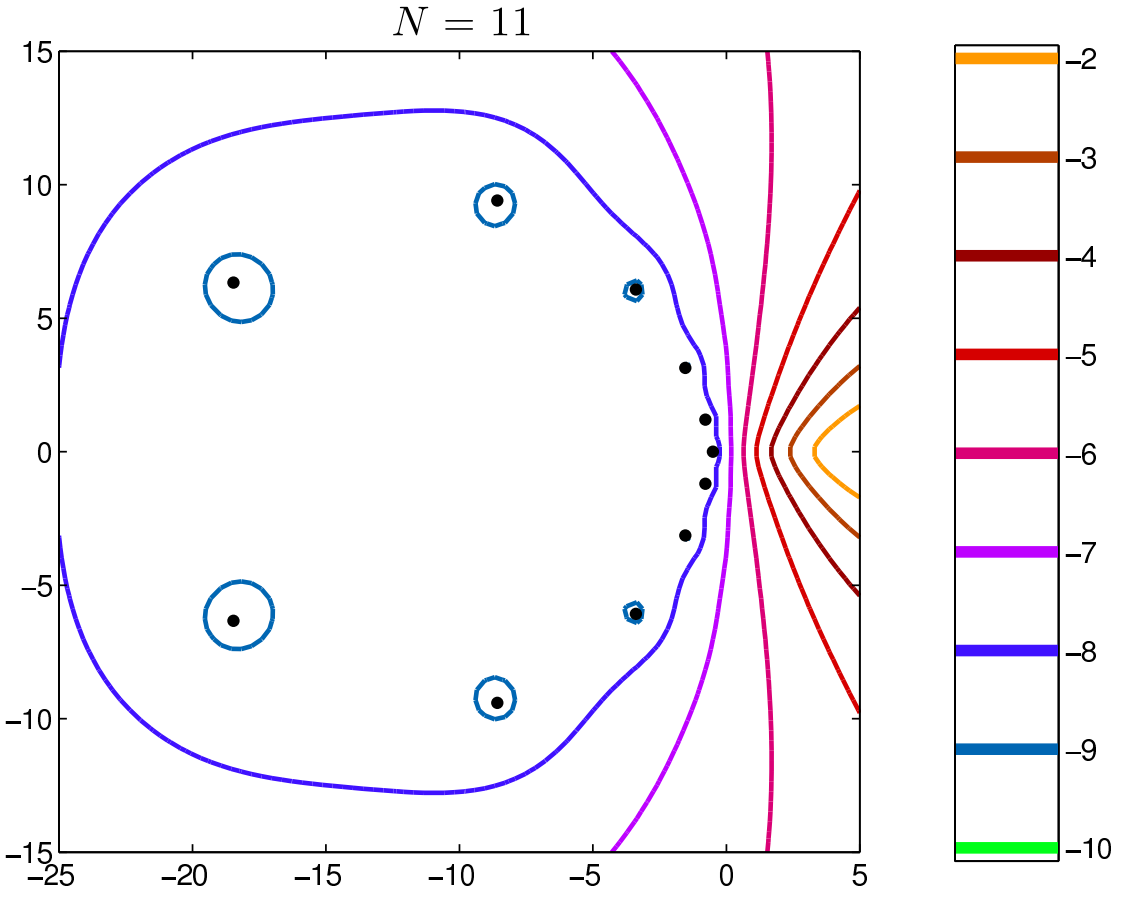
Computed pseudospectra of $\mathbf{A} - \mathbf{b}\mathbf{f}^*$.



Computed eigenvalues should be accurate to roughly $\varepsilon_{\text{mach}} \|\mathbf{A} - \mathbf{b}\mathbf{f}^*\|$.
For example, when $N = 11$, $\|\mathbf{A} - \mathbf{b}\mathbf{f}^*\| = 5.26 \times 10^8$.

Pole Placement Example: Pseudospectra

Computed pseudospectra of $\mathbf{A} - \mathbf{b}\mathbf{f}^*$.



Computed eigenvalues should be accurate to roughly $\varepsilon_{\text{mach}} \|\mathbf{A} - \mathbf{b}\mathbf{f}^*\|$.
For example, when $N = 11$, $\|\mathbf{A} - \mathbf{b}\mathbf{f}^*\| = 5.26 \times 10^8$.

Polynomial Zeros and Companion Matrices

MATLAB's `roots` command computes polynomial zeros by computing the eigenvalues of a companion matrix.

For example, given $p(z) = c_0 + c_1 z + c_2 z^2 + c_3 z^3 + c_4 z^4$, MATLAB builds

$$\mathbf{A} = \begin{bmatrix} -c_4/c_0 & -c_3/c_0 & -c_2/c_0 & -c_1/c_0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

whose characteristic polynomial is p.

Polynomial Zeros and Companion Matrices

MATLAB's `roots` command computes polynomial zeros by computing the eigenvalues of a companion matrix.

For example, given $p(z) = c_0 + c_1 z + c_2 z^2 + c_3 z^3 + c_4 z^4$, MATLAB builds

$$\mathbf{A} = \begin{bmatrix} -c_4/c_0 & -c_3/c_0 & -c_2/c_0 & -c_1/c_0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

whose characteristic polynomial is p.

Problem (Wilkinson's "Perfidious Polynomial")

Find the zeros of the polynomial

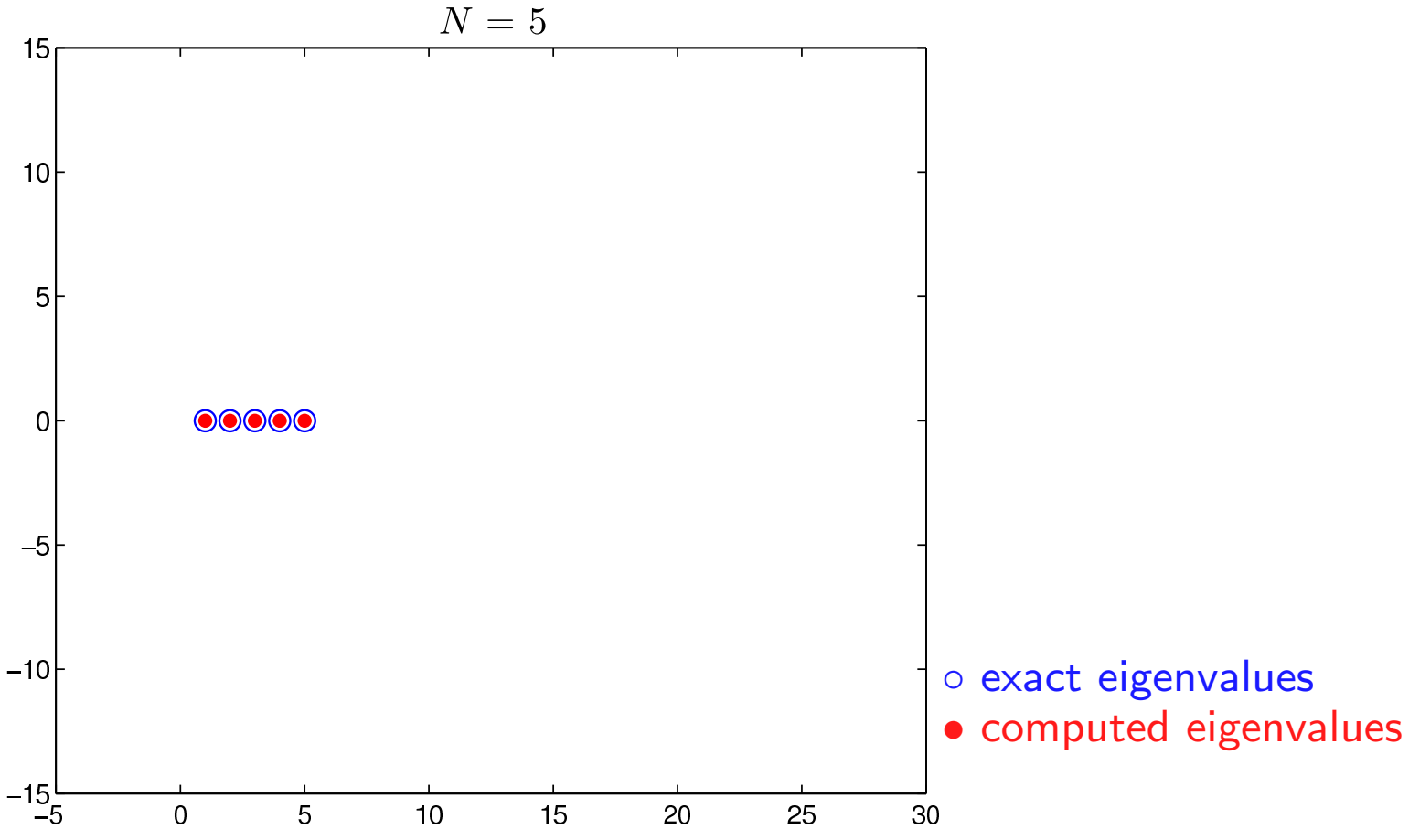
$$p(z) = (z-1)(z-2)\cdots(z-N)$$

from coefficients in the monomial basis.

MATLAB gives this as an example: `roots(poly(1:20))`.

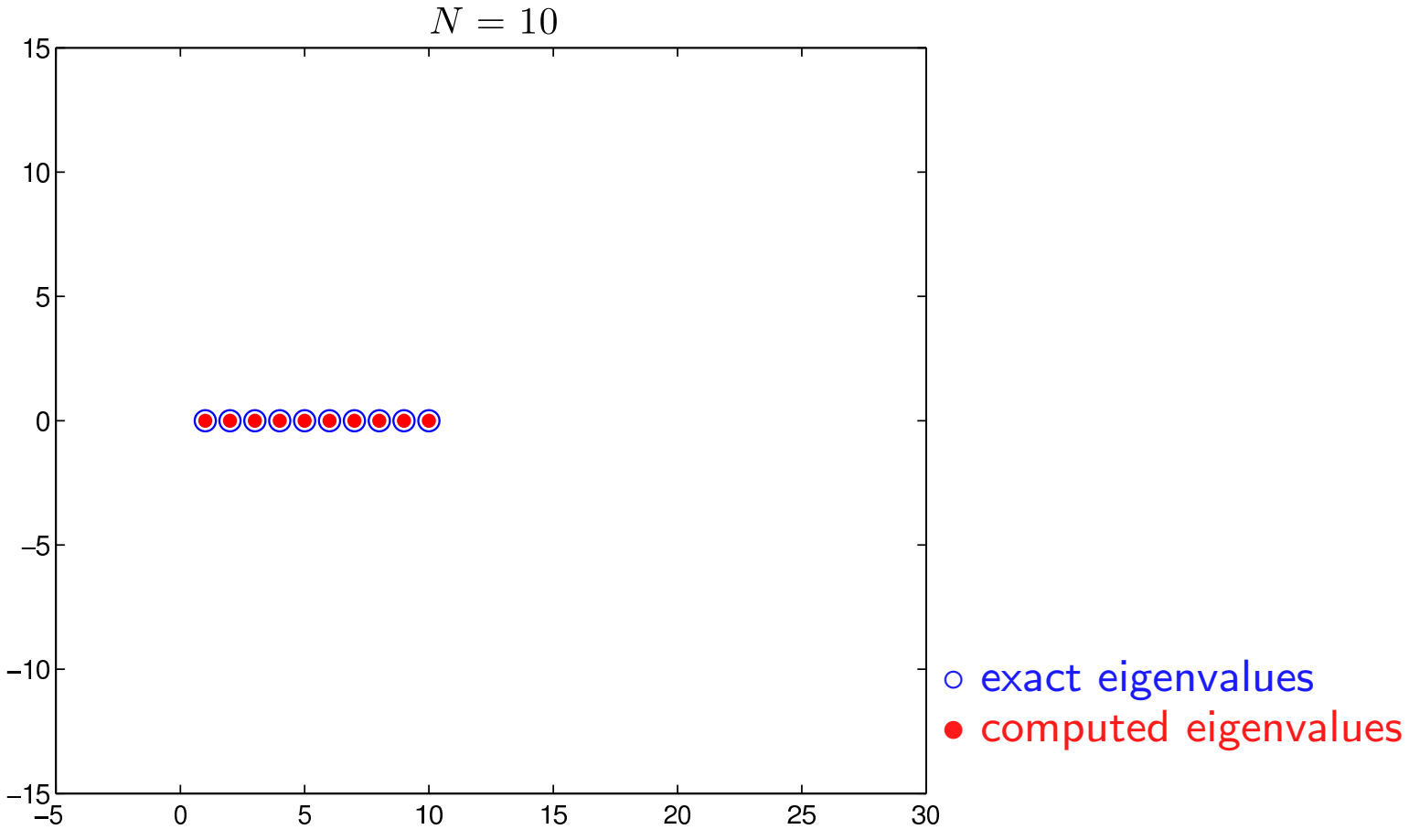
Roots of Wilkinson's Perfidious Polynomial

Computed eigenvalues of the companion matrix.



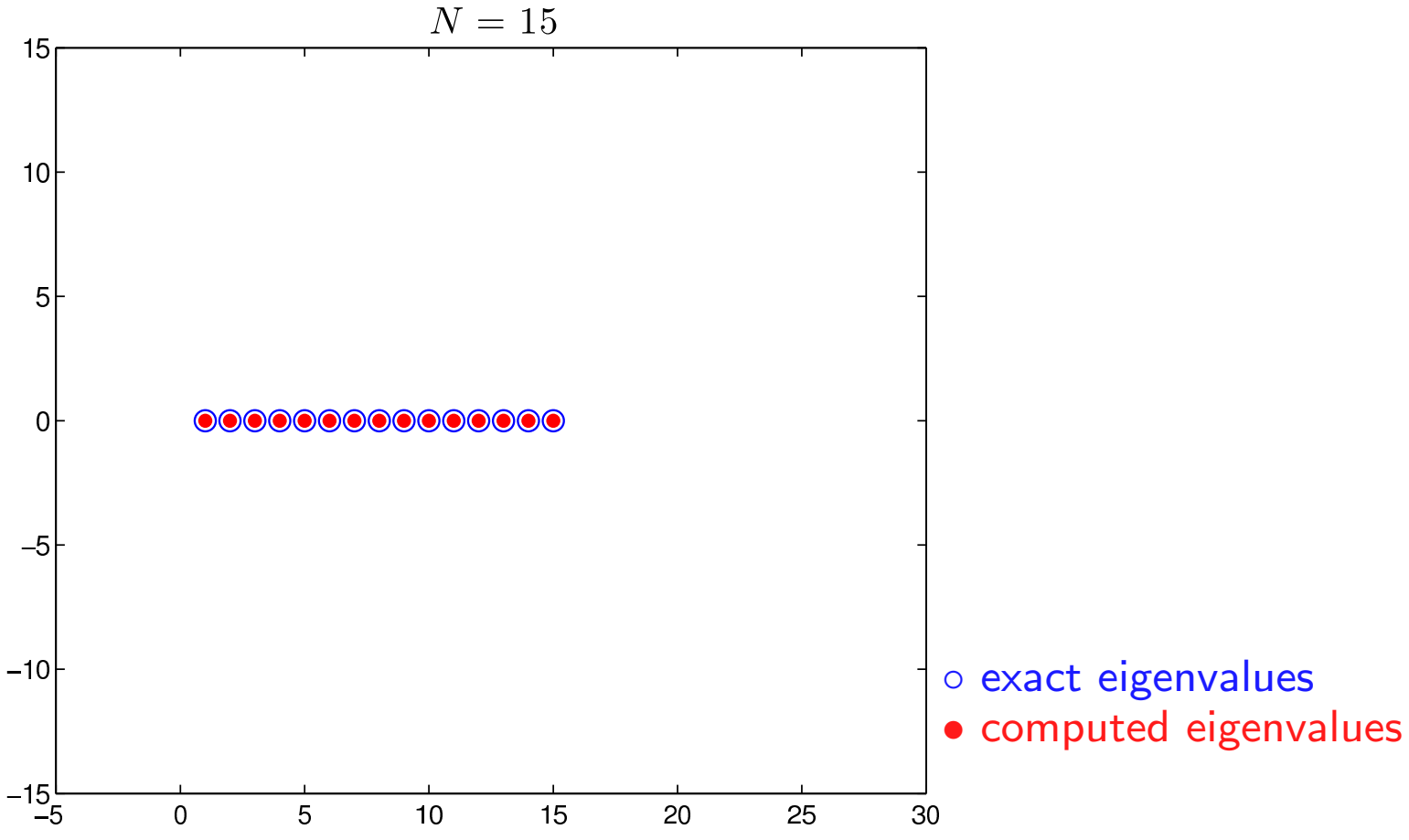
Roots of Wilkinson's Perfidious Polynomial

Computed eigenvalues of the companion matrix.



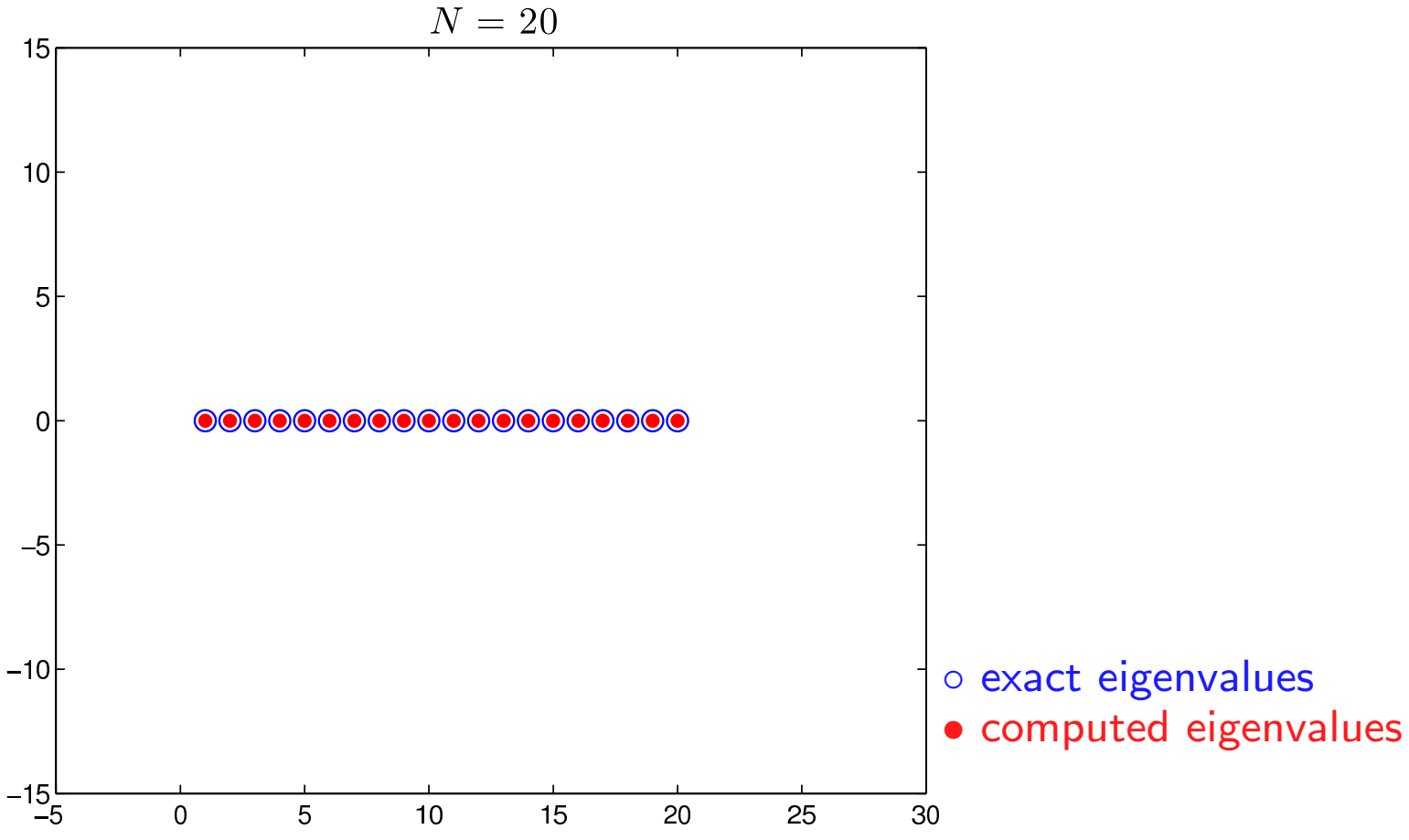
Roots of Wilkinson's Perfidious Polynomial

Computed eigenvalues of the companion matrix.



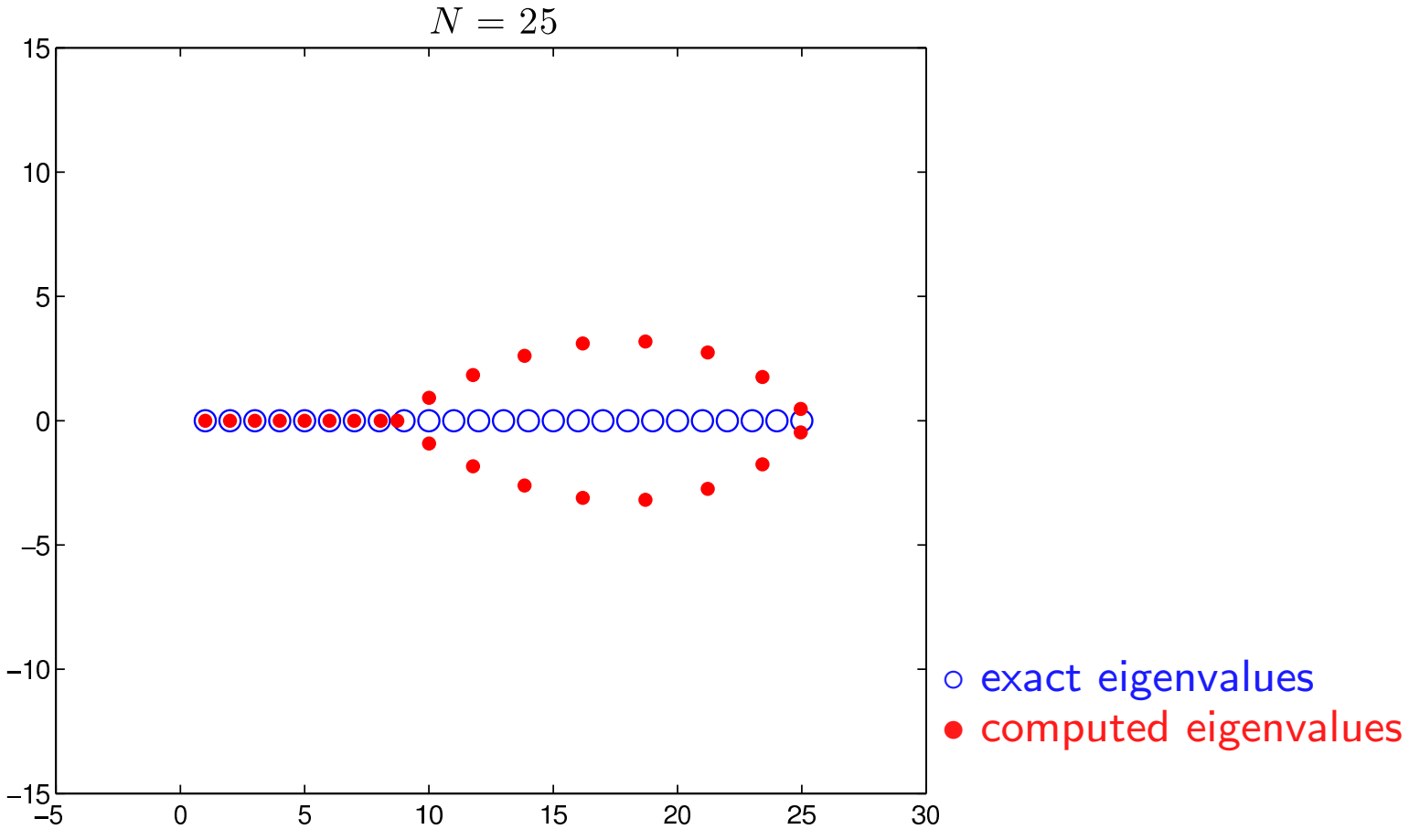
Roots of Wilkinson's Perfidious Polynomial

Computed eigenvalues of the companion matrix.



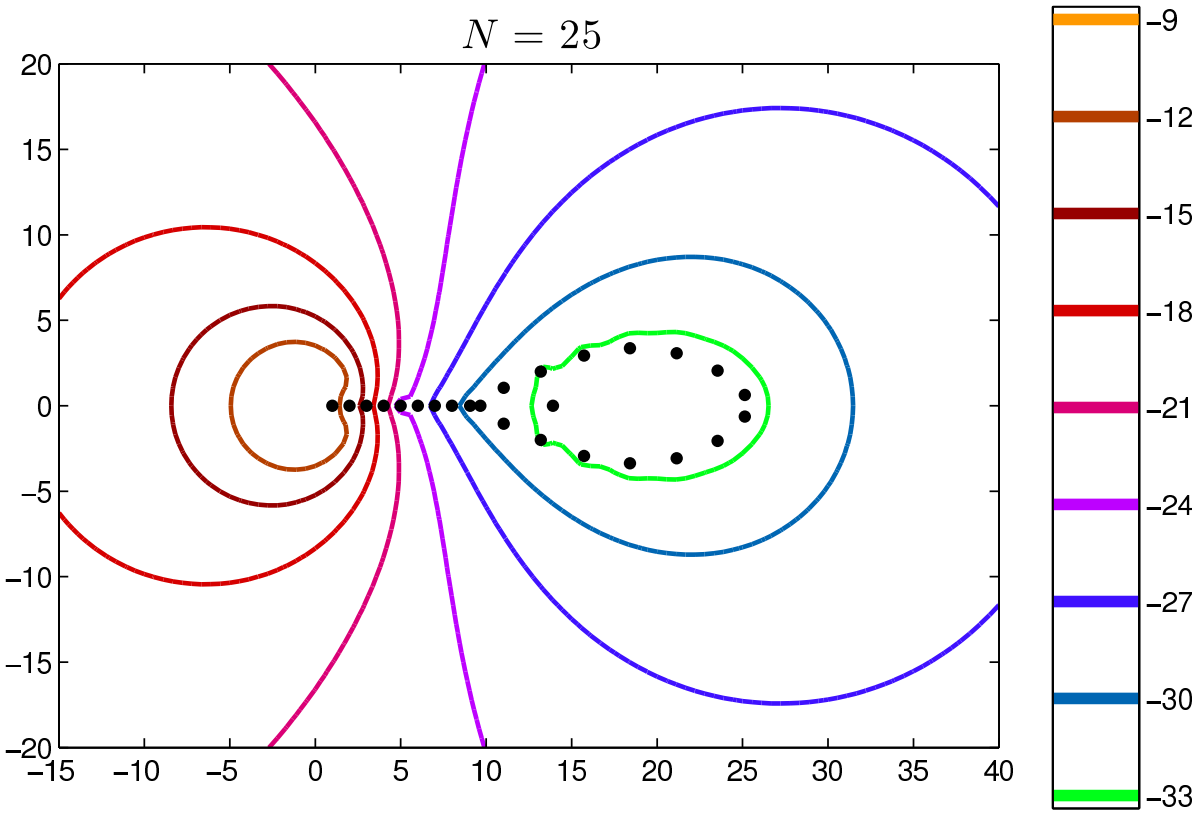
Roots of Wilkinson's Perfidious Polynomial

Computed eigenvalues of the companion matrix.



Roots of Wilkinson's Perfidious Polynomial

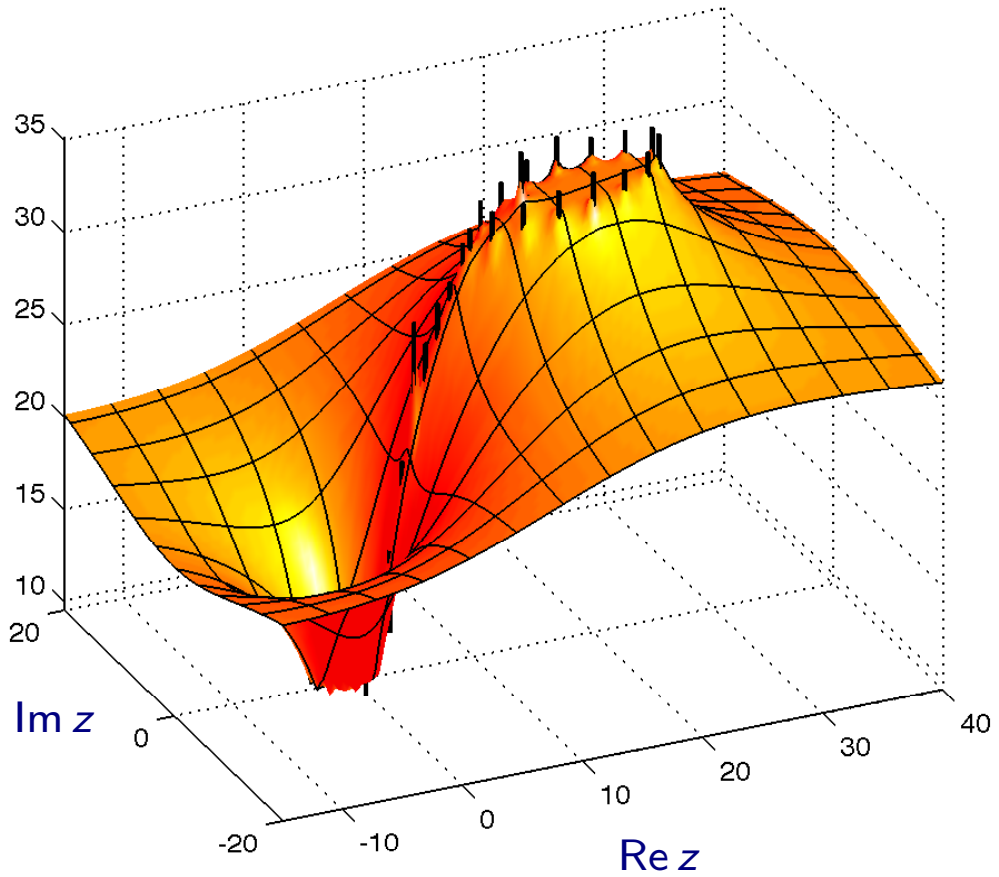
Pseudospectra for $N = 25$.



Roots of Wilkinson's Perfidious Polynomial

3d plot of the resolvent norm reveals the a local minimum.

$\|(z - \mathbf{A})^{-1}\|$



3. Bounding Norms of Functions of Matrices

Behavior of Matrices

We are primarily interested in using pseudospectra (and other tools) to study the *behavior* of a nonnormal matrix.

In particular, we seek to quantify (or bound) how nonnormality affects the value of functions of matrices.

Much research has been devoted to functions of matrices over the past decade; see the book by Nick Higham [Hig08]. We focus on functions that are analytic on the spectrum of $\mathbf{A}$.

Theorem (Spectral Mapping Theorem)

Suppose f is analytic on $\sigma(\mathbf{A})$. Then

$$\sigma(f(\mathbf{A})) = f(\sigma(\mathbf{A})).$$

Thus $\|f(\mathbf{A})\| \geq \max_{\mu \in \sigma(f(\mathbf{A}))} |\mu| = \max_{\lambda \in \sigma(\mathbf{A})} |f(\lambda)|$.

See [Huhtanen, 1999] for work on lower bounds.

Spectral Representation of a Matrix

Theorem (See Kato, 1980)

Any matrix with m distinct eigenvalues can be written in the form

$$\mathbf{A} = \sum_{j=1}^{m} \lambda_j \mathbf{P}_j + \mathbf{D}_j$$

where, for Jordan curves Γ_j surrounding λ_j and no other eigenvalues,

- $\mathbf{P}_j = \dfrac{1}{2\pi i} \displaystyle\int_{\Gamma_j} (z - \mathbf{A})^{-1} \, \mathrm{d}z$ *is a spectral projector;*
- $\mathbf{D}_j = \dfrac{1}{2\pi i} \displaystyle\int_{\Gamma_j} (z - \lambda_j)(z - \mathbf{A})^{-1} \, \mathrm{d}z$ *is nilpotent;*
- $\mathbf{P}_j\mathbf{A} = \mathbf{A}\mathbf{P}_j = \lambda_j \mathbf{P}_j$;
- $\mathbf{P}_j\mathbf{P}_k = \mathbf{0}$ *if* $j \neq k$;
- $\mathbf{D}_j = \mathbf{0}$ *if* λ_j *is not defective.*

The resolvent plays a fundamental role in the structure of the matrix $\mathbf{A}$.

Functions of a Matrix

If $\mathbf{A}$ is diagonalizable (i.e., no defective eigenvalues), then

$$\mathbf{A} = \sum_{j=1}^{m} \lambda_j \mathbf{P}_j.$$

If f is any function that is analytic on $\sigma(\mathbf{A})$ and on/inside all contours Γ_j, then

$$f(\mathbf{A}) = \sum_{j=1}^{m} f(\lambda_j) \mathbf{P}_j.$$

For all matrices, we have a more general formula.

Theorem (Cauchy Integral Formula for Matrices)

Let Γ be a finite union of Jordan curves containing $\sigma(\mathbf{A})$ in its interior, and suppose f is a function analytic on Γ and its interior. Then

$$f(\mathbf{A}) = \frac{1}{2\pi i} \int_{\Gamma} f(z)(z - \mathbf{A})^{-1} \, \mathrm{d}z.$$

Resolvent Bounds

Suppose the eigenvalues of $\mathbf{A} \in \mathbf{C}^{n\times n}$ are distinct.

Apply the previous formula to $f(\zeta) = (z-\zeta)^{-1}$:

$$f(\mathbf{A}) = (z-\mathbf{A})^{-1} = \sum_{j=1}^{n} \frac{1}{z-\lambda_j}\mathbf{P}_j.$$

If λ_j has right eigenvector $\mathbf{v}_j$ and left eigenvector $\mathbf{u}_j$, then

$$\mathbf{P}_j = \frac{\mathbf{v}_j\mathbf{u}_j^*}{\mathbf{u}_j^*\mathbf{v}_j}$$

and the norm of $\mathbf{P}_j$ is

$$\kappa(\lambda_j) := \|\mathbf{P}_j\| = \frac{\|\mathbf{v}_j\|\|\mathbf{u}_j\|}{|\mathbf{u}_j^*\mathbf{v}_j|}$$

the *condition number of the eigenvalue* λ_j.

Hence for z near λ_j,

$$\|(z-\mathbf{A})^{-1}\| \approx \frac{\kappa(\lambda_j)}{|z-\lambda_j|}.$$

Containment Regions for Pseudospectra

Hence for small $\varepsilon > 0$ and diagonalizable matrices, we can approximate

$$\sigma_\varepsilon(\mathbf{A}) \approx \cup_{j=1}^{n} \lambda_j + \Delta_{\varepsilon\kappa(\lambda_j)},$$

where $\Delta_r := \{z \in \mathbf{C} : |z| < r\}$.

Theorem (Bauer–Fike, 1963)

If $\mathbf{A} \in C^{n\times n}$ is diagonalizable, then for all $\varepsilon > 0$,

$$\sigma_\varepsilon(\mathbf{A}) \subseteq \bigcup_{j=1}^{n} \lambda_j + \Delta_{n\varepsilon\kappa(\lambda_j)}.$$

Unlike the earlier version of Bauer–Fike, the radii of the disks vary with j.

Near a defective eigenvalue of index d, Rellich's perturbation theory requires that the pseudospectrum behave like a disk whose radius scales like $\varepsilon^{1/d}$ as $\varepsilon \to 0$.

Bounds on $\|f(\mathbf{A})\|$

Suppose $\mathbf{A}$ is diagonalizable, $\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}$. Then

$$f(\mathbf{A}) = \sum_{j=1}^{n} f(\lambda_j)\mathbf{P}_j = \mathbf{V}f(\mathbf{\Lambda})\mathbf{V}^{-1}.$$

This immediately suggests several upper bounds on $\|f(\mathbf{A})\|$:

$$\begin{aligned} \|f(\mathbf{A})\| = \|\mathbf{V}f(\mathbf{\Lambda})\mathbf{V}^{-1}\| &\leq \|\mathbf{V}\|\|\mathbf{V}^{-1}\| \max_{\lambda\in\sigma(\mathbf{A})} |f(\lambda)| \\ &= \kappa(\mathbf{V}) \max_{\lambda\in\sigma(\mathbf{A})} |f(\lambda)|; \end{aligned}$$

$$\|f(\mathbf{A})\| \leq \sum_{j=1}^{n} |f(\lambda_j)|\|\mathbf{P}_j\| = \sum_{j=1}^{n} \kappa(\lambda_j)|f(\lambda_j)|.$$

We seek bounds that provide a more flexible way of handling nonnormality.

Pseudospectral Bounds on $\|f(\mathbf{A})\|$

Theorem (Cauchy Integral Formula for Matrices)

$$f(\mathbf{A}) = \frac{1}{2\pi i}\int_\Gamma f(z)(z-\mathbf{A})^{-1}\,\mathrm{d}z.$$

Take norms of the expression for $f(\mathbf{A})$:

$$\|f(\mathbf{A})\| \;=\; \left\|\frac{1}{2\pi i}\int_\Gamma f(z)(z-\mathbf{A})^{-1}\,\mathrm{d}z\right\| \;\leq\; \frac{1}{2\pi}\int_\Gamma |f(z)|\|(z-\mathbf{A})^{-1}\||dz|.$$

Pseudospectral Bounds on $\|f(\mathbf{A})\|$

Theorem (Cauchy Integral Formula for Matrices)

$$f(\mathbf{A}) = \frac{1}{2\pi i}\int_\Gamma f(z)(z-\mathbf{A})^{-1}\,\mathrm{d}z.$$

Take norms of the expression for $f(\mathbf{A})$:

$$\|f(\mathbf{A})\| \;=\; \left\|\frac{1}{2\pi i}\int_\Gamma f(z)(z-\mathbf{A})^{-1}\,\mathrm{d}z\right\| \;\le\; \frac{1}{2\pi}\int_\Gamma |f(z)|\|(z-\mathbf{A})^{-1}\||dz|.$$

Now pick Γ to be the boundary of $\sigma_\varepsilon(\mathbf{A})$

$$\begin{aligned}
\|f(\mathbf{A})\| &\le \frac{1}{2\pi}\int_{\partial\sigma_\varepsilon} |f(z)|\|(z-\mathbf{A})^{-1}\||dz| \\
&= \frac{1}{2\pi\varepsilon}\int_{\partial\sigma_\varepsilon} |f(z)||dz| \\
&\le \frac{1}{2\pi\varepsilon}\sup_{z\in\sigma_\varepsilon(\mathbf{A})}|f(z)|\int_{\partial\sigma_\varepsilon}|dz| \;\le\; \frac{L_\varepsilon}{2\pi\varepsilon}\sup_{z\in\sigma_\varepsilon(\mathbf{A})}|f(z)|
\end{aligned}$$

where L_ε denotes the arc-length of $\partial\sigma_\varepsilon(\mathbf{A})$ [Trefethen 1990].

Pseudospectral Bounds on $\|f(\mathbf{A})\|$

Theorem

Let f be analytic on $\sigma_\varepsilon(\mathbf{A})$ for some $\varepsilon > 0$. Then

$$\|f(\mathbf{A})\| \leq \frac{L_\varepsilon}{2\pi\varepsilon} \sup_{z \in \sigma_\varepsilon(\mathbf{A})} |f(z)|,$$

where L_ε denotes the countour length of the boundary of $\sigma_\varepsilon(\mathbf{A})$.

Some key observations:

- This should be regarded as a *family of bounds* that vary with ε;
- The best choice for ε will depend on the problem;
- Sometimes it is excellent; usually it is decent; on occasion it is poor;
- The choice of ε has *nothing* to do with rounding errors; do not expect the bound to be most descriptive when $\varepsilon = \varepsilon_{\text{mach}}$ or $\varepsilon = \|\mathbf{A}\|\varepsilon_{\text{mach}}$.

GMRES Convergence

Example

We wish to use the pseudospectral analysis just presented to bound the norm of the residual in the GMRES algorithm for iteratively solving $\mathbf{Ax} = \mathbf{b}$.

GMRES produces optimal iterates $\mathbf{x}_k$ whose residuals $\mathbf{r}_k = \mathbf{b} - \mathbf{Ax}_k$ satisfy

$$\|\mathbf{r}_k\| \leq \min_{\substack{p \in \mathcal{P}_k \\ p(0)=1}} \|p(\mathbf{A})\mathbf{b}\|,$$

where $\mathcal{P}_k$ denotes the set of polynomials of degree k or less.

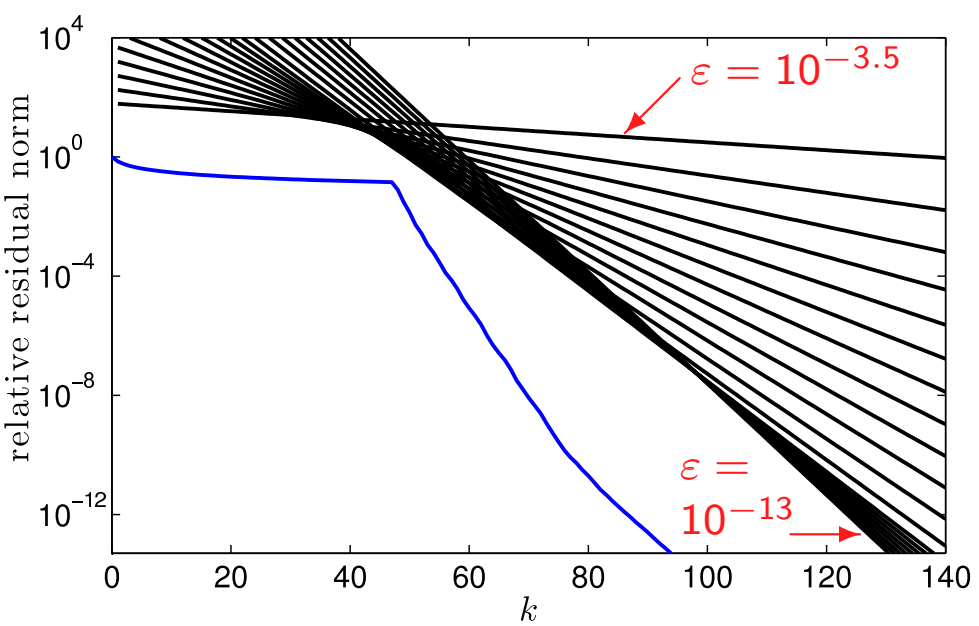
Applying the previous analysis yields

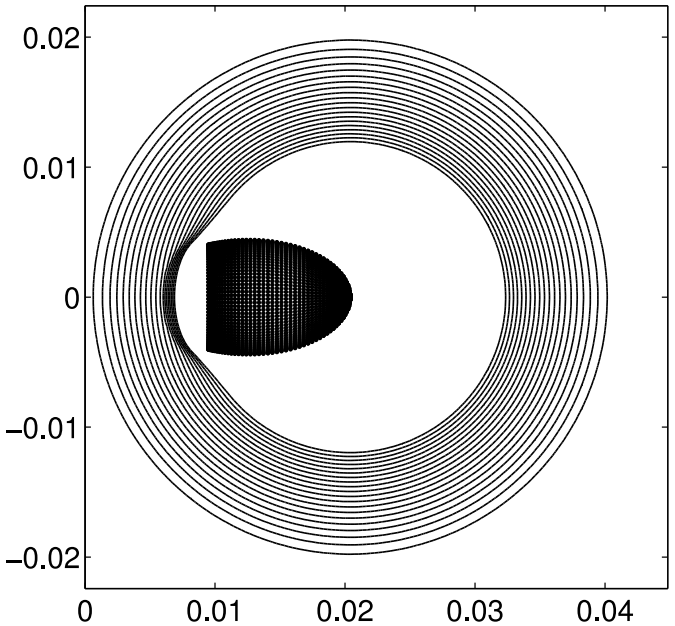
$$\frac{\|\mathbf{r}_k\|}{\|\mathbf{b}\|} \leq \frac{L_\varepsilon}{2\pi\varepsilon} \min_{\substack{p \in \mathcal{P}_k \\ p(0)=1}} \sup_{z \in \sigma_\varepsilon(\mathbf{A})} |p(z)|.$$

Different ε values give the best bounds at different stages of convergence.

Illustrations and applications: [Trefethen 1990; E. 2000; Sifuentes, E., Morgan 2011].

GMRES Convergence: Example





Pseudospectral bound for convection–diffusion problem, $n = 2304$.

$$\varepsilon = 10^{-3.5}, 10^{-3}, 10^{-3.5}, \ldots, 10^{-13}$$

Bound $\sigma_\varepsilon(\mathbf{A})$ with a circle centered at c, use $p_k(z) = (1 - z/c)^k$.

Bounds on the Matrix Exponential

To understand behavior of $\mathbf{x}'(t) = \mathbf{A}\mathbf{x}(t)$ [and LTI control systems], we wish to use pseudospectra to bound $\|\mathrm{e}^{t\mathbf{A}}\|$.

Definition

The *spectral abscissa* is the rightmost point in the spectrum:

$$\alpha(\mathbf{A}) := \max_{\lambda \in \sigma(\mathbf{A})} \operatorname{Re} z.$$

The *ε-pseudospectral abscissa* is the supremum of the real parts of $z \in \sigma_\varepsilon(\mathbf{A})$:

$$\alpha_\varepsilon(\mathbf{A}) := \sup_{z \in \sigma_\varepsilon(\mathbf{A})} \operatorname{Re} z.$$

Applying the Cauchy integral bound to $f(z) = \mathrm{e}^{tz}$ gives an upper bound.

Bounds on the Matrix Exponential

To understand behavior of $\mathbf{x}'(t) = \mathbf{A}\mathbf{x}(t)$ [and LTI control systems], we wish to use pseudospectra to bound $\|e^{t\mathbf{A}}\|$.

Definition

The *spectral abscissa* is the rightmost point in the spectrum:

$$\alpha(\mathbf{A}) := \max_{\lambda \in \sigma(\mathbf{A})} \operatorname{Re} z.$$

The *ε-pseudospectral abscissa* is the supremum of the real parts of $z \in \sigma_\varepsilon(\mathbf{A})$:

$$\alpha_\varepsilon(\mathbf{A}) := \sup_{z \in \sigma_\varepsilon(\mathbf{A})} \operatorname{Re} z.$$

Applying the Cauchy integral bound to $f(z) = e^{tz}$ gives an upper bound.

Theorem (Upper Bound)

For any $\mathbf{A} \in C^{n\times n}$ and $\varepsilon > 0$,

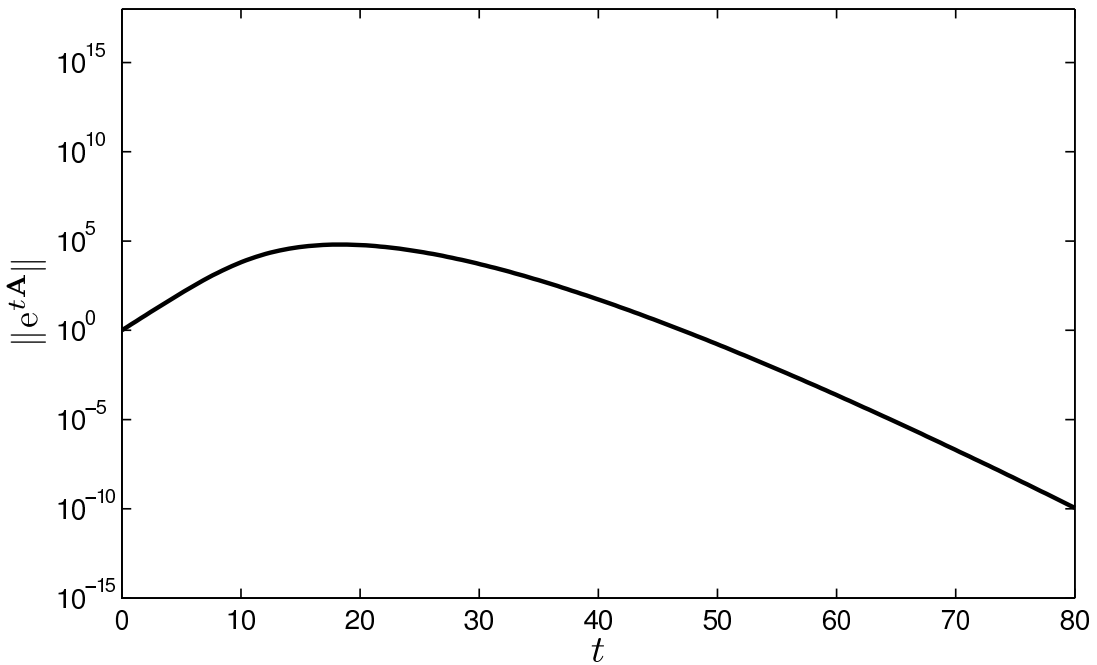
$$\|e^{t\mathbf{A}}\| \leq \frac{L_\varepsilon}{2\pi\varepsilon} e^{t\alpha_\varepsilon(\mathbf{A})},$$

where L_ε denotes the contour length of the boundary of $\sigma_\varepsilon(\mathbf{A})$.

Upper Bound on Matrix Powers

$$\mathbf{A} = \begin{bmatrix} -1 & 2 & & & \\ & -1 & \ddots & & \\ & & \ddots & 2 & \\ & & & -1 & 2 \\ & & & & -1 \end{bmatrix} \in \mathbf{C}^{20\times 20}.$$

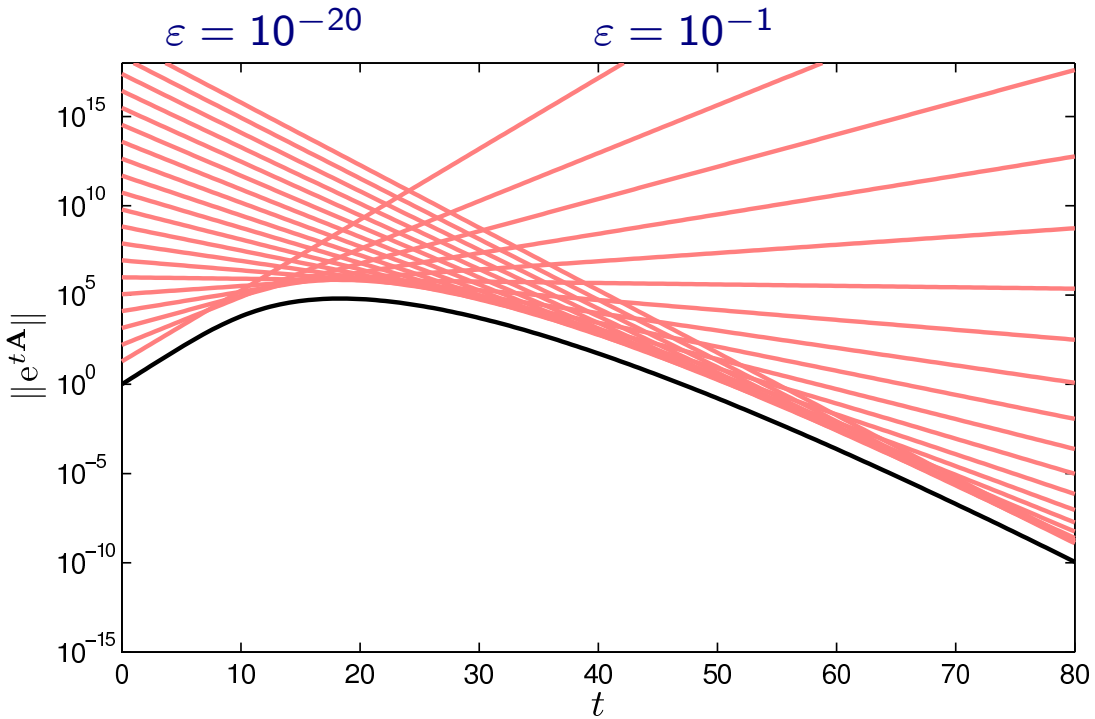
Question: can you estimate $\sigma_\varepsilon(\mathbf{A})$ for this matrix?



Upper Bound on Matrix Powers

$$\mathbf{A} = \begin{bmatrix} -1 & 2 & & & \\ & -1 & \ddots & & \\ & & \ddots & 2 & \\ & & & -1 & 2 \\ & & & & -1 \end{bmatrix} \in \mathbf{C}^{20\times 20}.$$

Question: can you estimate $\sigma_\varepsilon(\mathbf{A})$ for this matrix?



Lower Bounds on the Matrix Exponential

We would like to *guarantee* the potential for transient growth.

Theorem (Lower bound on $\mathrm{e}^{t\mathbf{A}}$)

Suppose $\alpha(\mathbf{A}) < 0$. *Then for all* $\varepsilon > 0$,

$$\sup_{t \geq 0} \|\mathrm{e}^{t\mathbf{A}}\| \geq \frac{\alpha_\varepsilon(\mathbf{A})}{\varepsilon}.$$

Lower Bounds on the Matrix Exponential

We would like to *guarantee* the potential for transient growth.

Theorem (Lower bound on $\mathrm{e}^{t\mathbf{A}}$)

Suppose $\alpha(\mathbf{A}) < 0$. *Then for all* $\varepsilon > 0$,

$$\sup_{t \geq 0} \|\mathrm{e}^{t\mathbf{A}}\| \geq \frac{\alpha_\varepsilon(\mathbf{A})}{\varepsilon}.$$

Proof. Recall the formula for the Laplace transform of $\mathrm{e}^{t\alpha}$: For $s > \alpha$,

$$\int_0^\infty \mathrm{e}^{t\alpha} \mathrm{e}^{-st} \mathrm{d}t = \frac{1}{s - \alpha}.$$

This formula generalizes for matrices, e.g., when $\alpha(\mathbf{A}) < 0$ and $\operatorname{Re} z > 0$:

$$\int_0^\infty \mathrm{e}^{t\mathbf{A}} \mathrm{e}^{-zt} \mathrm{d}t = (z - \mathbf{A})^{-1}.$$

Lower Bounds on the Matrix Exponential

We would like to *guarantee* the potential for transient growth.

Theorem (Lower bound on e^{tA})

Suppose $\alpha(\mathbf{A}) < 0$. *Then for all* $\varepsilon > 0$,

$$\sup_{t \geq 0} \|e^{t\mathbf{A}}\| \geq \frac{\alpha_\varepsilon(\mathbf{A})}{\varepsilon}.$$

Proof.

$$\int_0^\infty e^{t\mathbf{A}} e^{-zt} dt = (z - \mathbf{A})^{-1}.$$

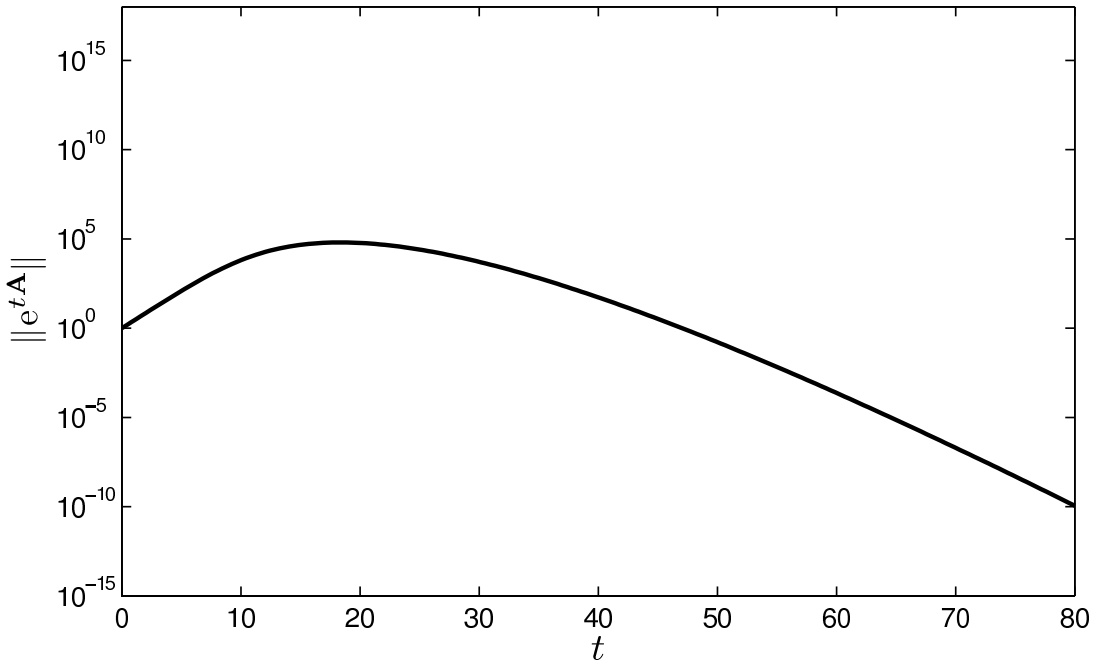
Suppose $\|e^{t\mathbf{A}}\| \leq M$ for all $t \geq 0$. Then for any $z \in \sigma_\varepsilon(\mathbf{A})$, $\operatorname{Re} z > 0$:

$$\begin{aligned} \frac{1}{\varepsilon} < \|(z - \mathbf{A})^{-1}\| &= \left\| \int_0^\infty e^{t\mathbf{A}} e^{-zt}\, dt \right\| \\ &\leq \int_0^\infty \|e^{t\mathbf{A}}\| |e^{-zt}|\, dt \leq M \int_0^\infty e^{-(\operatorname{Re} z)t}\, dt = \frac{M}{\operatorname{Re} z}. \end{aligned}$$

Hence $M \geq (\operatorname{Re} z)/\varepsilon$. Take the sup over all $z \in \sigma_\varepsilon(\mathbf{A})$ to get the bound. □

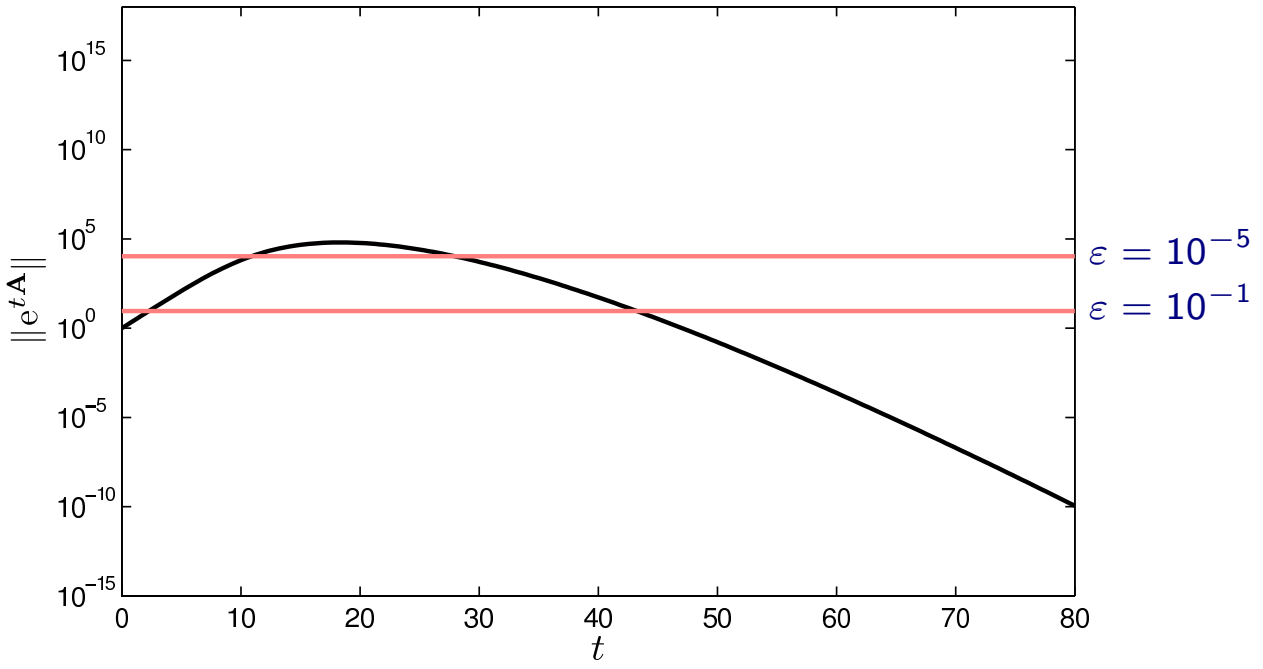
Lower Bound on the Matrix Exponential

$$\mathbf{A} = \begin{bmatrix} -1 & 2 & & & \\ & -1 & \ddots & & \\ & & \ddots & 2 & \\ & & & -1 & 2 \\ & & & & -1 \end{bmatrix} \in \mathbf{C}^{20\times 20}.$$



Lower Bound on the Matrix Exponential

$$\mathbf{A} = \begin{bmatrix} -1 & 2 & & & \\ & -1 & \ddots & & \\ & & \ddots & 2 & \\ & & & -1 & 2 \\ & & & & -1 \end{bmatrix} \in \mathbf{C}^{20\times 20}.$$



Matrix Powers

Definition

The *spectral radius* is the largest point in the spectrum:

$$\rho(\mathbf{A}) := \max_{\lambda \in \sigma(\mathbf{A})} |z|.$$

The ε-*pseudospectral radius* is the supremum of magnitudes of points in $\sigma_\varepsilon(\mathbf{A})$:

$$\rho_\varepsilon(\mathbf{A}) := \sup_{z \in \sigma_\varepsilon(\mathbf{A})} |z|.$$

Matrix Powers

Definition

The *spectral radius* is the largest point in the spectrum:

$$\rho(\mathbf{A}) := \max_{\lambda \in \sigma(\mathbf{A})} |z|.$$

The *ε-pseudospectral radius* is the supremum of magnitudes of points in $\sigma_\varepsilon(\mathbf{A})$:

$$\rho_\varepsilon(\mathbf{A}) := \sup_{z \in \sigma_\varepsilon(\mathbf{A})} |z|.$$

Theorem (Upper Bound)

For any $\mathbf{A} \in C^{n \times n}$ and $\varepsilon > 0$,

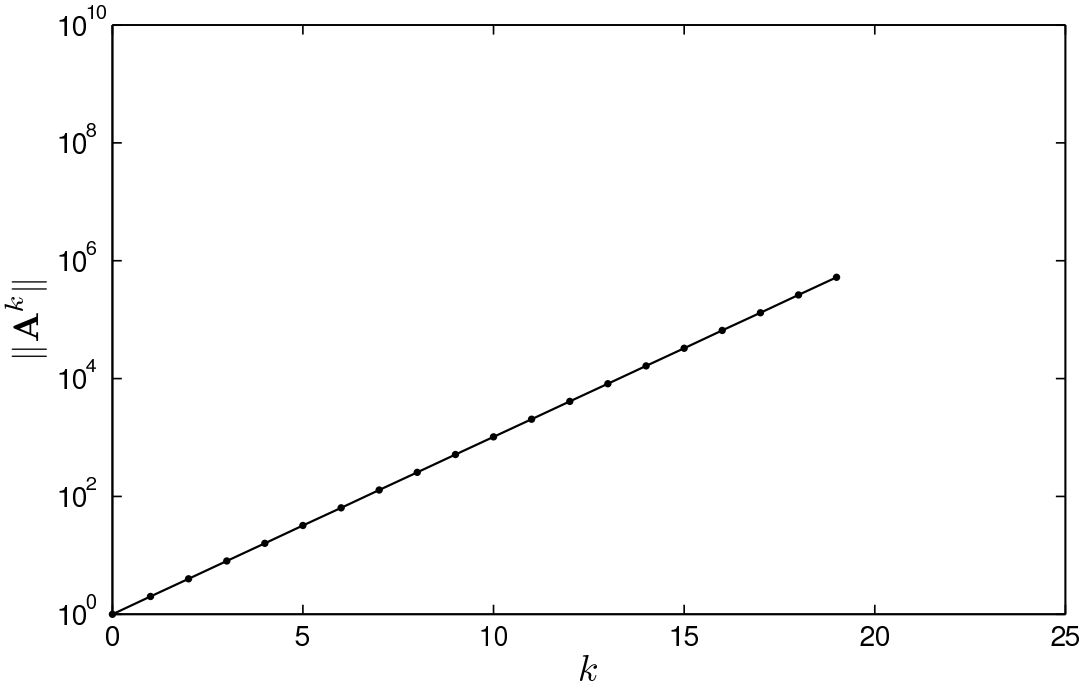
$$\|\mathbf{A}^k\| \leq \frac{\rho_\varepsilon(\mathbf{A})^{k+1}}{\varepsilon}$$

Proof: Apply the Cauchy integral bound, taking Γ to be the circle of radius $\rho_\varepsilon(\mathbf{A})$ centered at the origin. □

Upper Bound on Matrix Powers

$$\mathbf{A} = \begin{bmatrix} 0 & 2 & & & \\ & 0 & \ddots & & \\ & & \ddots & 2 & \\ & & & 0 & 2 \\ & & & & 0 \end{bmatrix} \in \mathbf{C}^{20\times 20}.$$

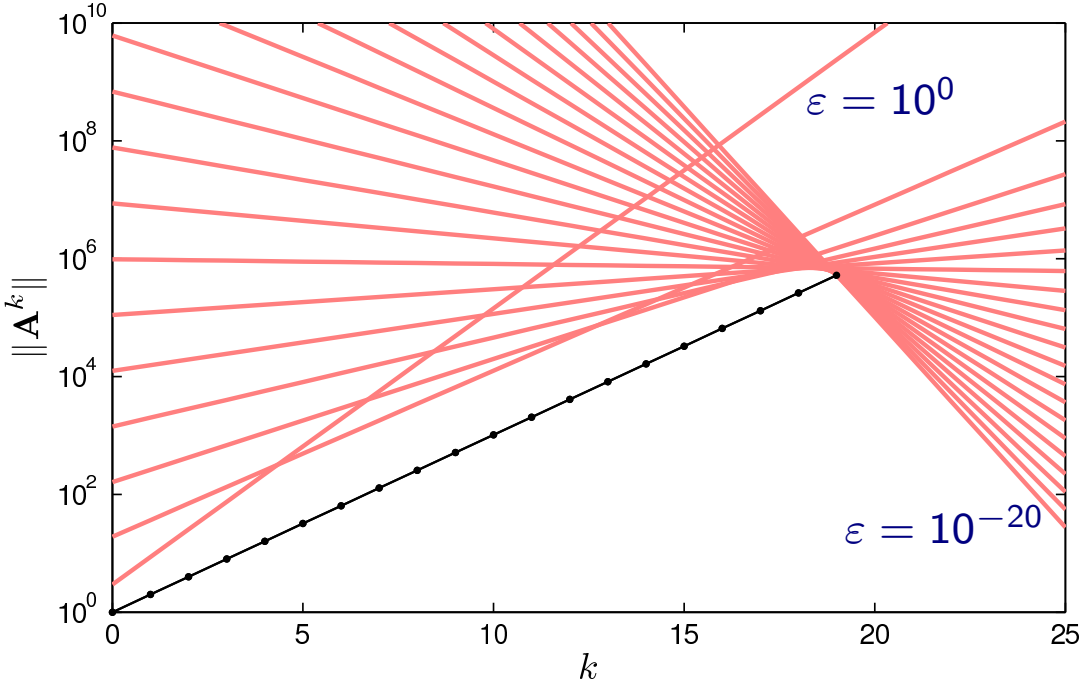
$\mathbf{A}^{20} = \mathbf{0}$, but lower powers show transient growth.



Upper Bound on Matrix Powers

$$\mathbf{A} = \begin{bmatrix} 0 & 2 & & & \\ & 0 & \ddots & & \\ & & \ddots & 2 & \\ & & & 0 & 2 \\ & & & & 0 \end{bmatrix} \in \mathbf{C}^{20\times 20}.$$

$\mathbf{A}^{20} = \mathbf{0}$, but lower powers show transient growth.



Lower Bounds on Matrix Powers

Theorem (Lower bound on power growth)

Suppose $\rho(\mathbf{A}) < 1$. *Then for all* $\varepsilon > 0$,

$$\sup_{k \geq 0} \|\mathbf{A}^k\| \geq \frac{\rho_\varepsilon(\mathbf{A}) - 1}{\varepsilon}.$$

Lower Bounds on Matrix Powers

Theorem (Lower bound on power growth)

Suppose $\rho(\mathbf{A}) < 1$. *Then for all* $\varepsilon > 0$,

$$\sup_{k \geq 0} \|\mathbf{A}^k\| \geq \frac{\rho_\varepsilon(\mathbf{A}) - 1}{\varepsilon}.$$

Proof. Since $\rho(\mathbf{A}) < 1$, $\mathbf{A}^k \to \mathbf{0}$ as $k \to \infty$. Let M denote the maximum value of $\|\mathbf{A}^k\|$, $k \geq 0$, and suppose $z \in \sigma_\varepsilon(\mathbf{A})$ for $|z| > 1$. Then

$$\frac{1}{\varepsilon} < \|(z - \mathbf{A})^{-1}\| = \left\| \frac{1}{|z|} \left(1 + \frac{1}{|z|}\mathbf{A} + \frac{1}{|z|^2}\mathbf{A}^2 + \cdots \right) \right\|$$

Lower Bounds on Matrix Powers

Theorem (Lower bound on power growth)

Suppose $\rho(\mathbf{A}) < 1$. *Then for all* $\varepsilon > 0$,

$$\sup_{k \geq 0} \|\mathbf{A}^k\| \geq \frac{\rho_\varepsilon(\mathbf{A}) - 1}{\varepsilon}.$$

Proof. Since $\rho(\mathbf{A}) < 1$, $\mathbf{A}^k \to \mathbf{0}$ as $k \to \infty$. Let M denote the maximum value of $\|\mathbf{A}^k\|$, $k \geq 0$, and suppose $z \in \sigma_\varepsilon(\mathbf{A})$ for $|z| > 1$. Then

$$\begin{aligned} \frac{1}{\varepsilon} < \|(z - \mathbf{A})^{-1}\| &= \left\| \frac{1}{|z|} \left(1 + \frac{1}{|z|}\mathbf{A} + \frac{1}{|z|^2}\mathbf{A}^2 + \cdots \right) \right\| \\ &\leq \frac{1}{|z|} \left(M + \frac{M}{|z|} + \frac{M}{|z|^2} + \cdots \right) \end{aligned}$$

Lower Bounds on Matrix Powers

Theorem (Lower bound on power growth)

Suppose $\rho(\mathbf{A}) < 1$. *Then for all* $\varepsilon > 0$,

$$\sup_{k \geq 0} \|\mathbf{A}^k\| \geq \frac{\rho_\varepsilon(\mathbf{A}) - 1}{\varepsilon}.$$

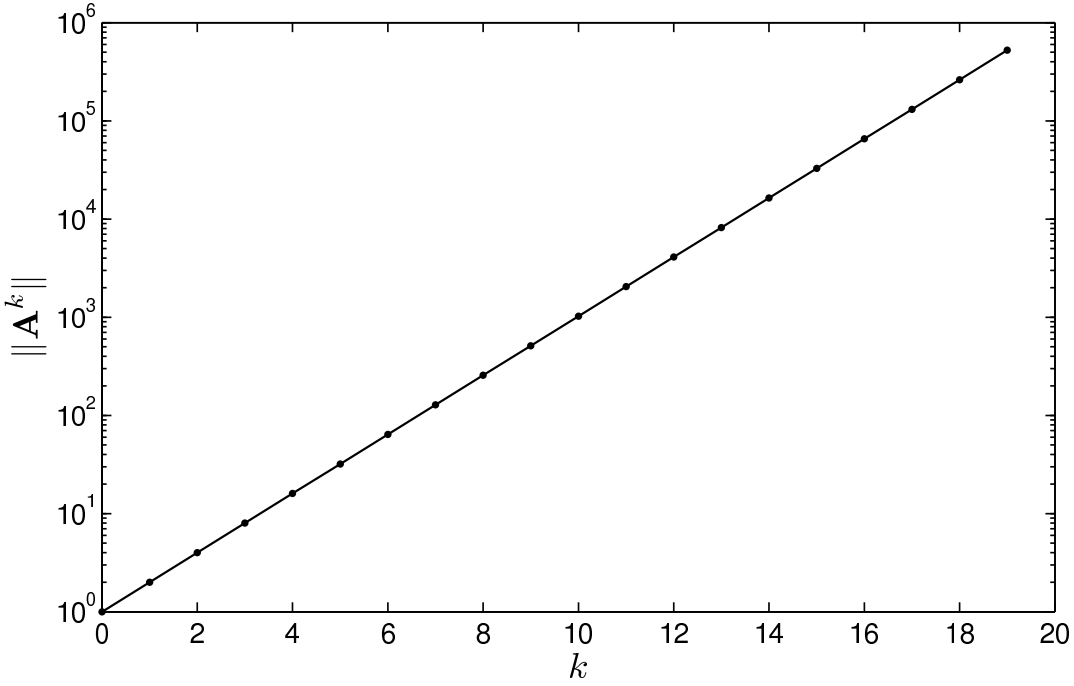
Proof. Since $\rho(\mathbf{A}) < 1$, $\mathbf{A}^k \to \mathbf{0}$ as $k \to \infty$. Let M denote the maximum value of $\|\mathbf{A}^k\|$, $k \geq 0$, and suppose $z \in \sigma_\varepsilon(\mathbf{A})$ for $|z| > 1$. Then

$$\begin{aligned} \frac{1}{\varepsilon} < \|(z - \mathbf{A})^{-1}\| &= \left\| \frac{1}{|z|} \left(1 + \frac{1}{|z|}\mathbf{A} + \frac{1}{|z|^2}\mathbf{A}^2 + \cdots \right) \right\| \\ &\leq \frac{1}{|z|} \left(M + \frac{M}{|z|} + \frac{M}{|z|^2} + \cdots \right) \\ &= \frac{M}{|z|} \frac{1}{1 - 1/|z|} = \frac{M}{|z| - 1}. \end{aligned}$$

Rearrange to obtain $M \geq (|z| - 1)/\varepsilon$ for all $z \in \sigma_\varepsilon(\mathbf{A})$ with $|z| > 1$.
Take the supremum of $|z|$ over all $z \in \sigma_\varepsilon(\mathbf{A})$ to get the result. □

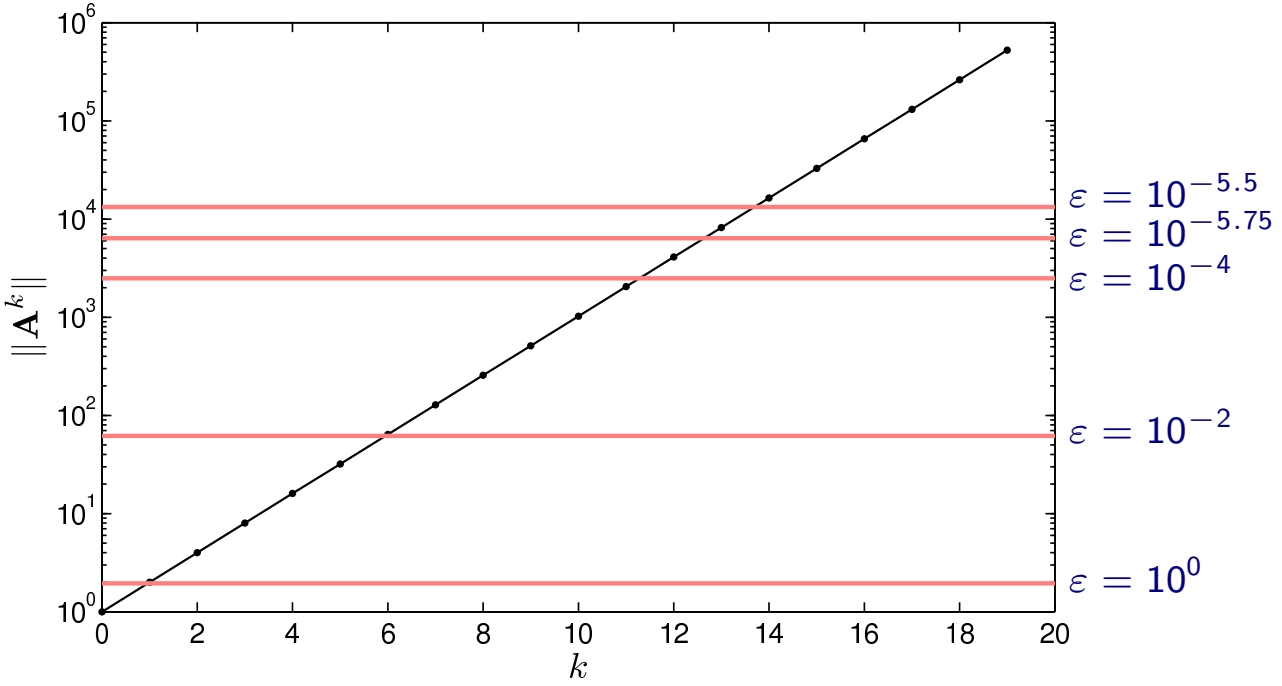
Lower Bound on Matrix Powers

$$\mathbf{A} = \begin{bmatrix} 0 & 2 & & & \\ & 0 & \ddots & & \\ & & \ddots & 2 & \\ & & & 0 & 2 \\ & & & & 0 \end{bmatrix} \in \mathbf{C}^{20\times 20}.$$



Lower Bound on Matrix Powers

$$\mathbf{A} = \begin{bmatrix} 0 & 2 & & & \\ & 0 & \ddots & & \\ & & \ddots & 2 & \\ & & & 0 & 2 \\ & & & & 0 \end{bmatrix} \in \mathbf{C}^{20\times 20}.$$



Pseudospectra are not a Panacea

Key question: "Do pseudospectra determine behavior of a matrix?" [Greenbaum & Trefethen, 1993].

Greenbaum and Trefethen define "behavior" to mean "norms of polynomials".

They prove that pseudospectra *do not* determine behavior.

$$\mathbf{A}_1 = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & & \\ & & & 0 & 0 \\ & & & & 0 \end{bmatrix}, \qquad \mathbf{A}_2 = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & & \\ & & & 0 & \alpha \\ & & & & 0 \end{bmatrix}.$$

If $\alpha \in (1, \sqrt{2}]$, then $\sigma_\varepsilon(\mathbf{A}_1) = \sigma_\varepsilon(\mathbf{A}_2)$ *for all* $\varepsilon > 0$,

$$1 = \|\mathbf{A}_1\| \neq \|\mathbf{A}_2\| = \sqrt{2}.$$

Ransford et al. [2007, 2009] have gone on to prove a host of similarly pessimistic results along these lines, e.g., relating $\|\mathbf{A}_1^k\|$ to $\|\mathbf{A}_2^k\|$, even for matrices with *superidentical* pseudospectra.

Pseudospectra are not a Panacea (in the 2-norm)

Key question: "Do pseudospectra determine behavior of a matrix?" [Greenbaum & Trefethen, 1993].

Greenbaum and Trefethen define "behavior" to mean "norms of polynomials".

They prove that pseudospectra *do not* determine behavior.

$$\mathbf{A}_1 = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & & \\ & & & 0 & 0 \\ & & & & 0 \end{bmatrix}, \qquad \mathbf{A}_2 = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & & \\ & & & 0 & \alpha \\ & & & & 0 \end{bmatrix}.$$

If $\alpha \in (1, \sqrt{2}]$, then $\sigma_\varepsilon(\mathbf{A}_1) = \sigma_\varepsilon(\mathbf{A}_2)$ *for all* $\varepsilon > 0$,

$$1 = \|\mathbf{A}_1\| \neq \|\mathbf{A}_2\| = \sqrt{2}.$$

Ransford et al. [2007, 2009] have gone on to prove a host of similarly pessimistic results along these lines, e.g., relating $\|\mathbf{A}_1^k\|$ to $\|\mathbf{A}_2^k\|$, even for matrices with *superidentical* pseudospectra.

However: Greenbaum and Trefethen [1993] do show that pseudospectra determine behavior in the Frobenius norm ($\sigma_\varepsilon(\mathbf{A})$ defined via resolvent norms).

Complex versus Real Perturbations

In many applications $\mathbf{A}$ contains only real entries; uncertainties in physical parameters will only induce *real* perturbations.

Perhaps instead of the usual definition

$$\sigma_\varepsilon(\mathbf{A}) = \{z \in \mathbf{C} : z \in \sigma(\mathbf{A} + \mathbf{E}) \text{ for some } \mathbf{E} \in \mathbf{C}^{n\times n} \text{ with } \|\mathbf{E}\| < \varepsilon\}$$

we should have

$$\sigma_\varepsilon(\mathbf{A}) = \{z \in \mathbf{C} : z \in \sigma(\mathbf{A} + \mathbf{E}) \text{ for some } \mathbf{E} \in \mathbf{R}^{n\times n} \text{ with } \|\mathbf{E}\| < \varepsilon\}$$

cf. [Hinrichsen & Pritchard, 1990, 1992, 2005].

Similarly, one can study structured perturbations that preserve other properties: stochasticity, nonnegativity, symplecticity, Toeplitz or companion structure...

For analysis involving more sophisticated structured perturbations, see the work of Michael Karow et al. [2003 – 2011...]

Complex versus Real Perturbations

The size of a complex perturbation required to make z an eigenvalue of $\mathbf{A}$ is:

$$d_{\mathbf{C}}(\mathbf{A}, z) = \left(\sigma_1\left((z - \mathbf{A})^{-1}\right)\right)^{-1}.$$

The size of a real perturbation required to make z an eigenvalue of $\mathbf{A}$ is:

$$d_{\mathbf{R}}(\mathbf{A}, z) = \left(\inf_{\gamma \in (0,1]} \sigma_2\left(\begin{bmatrix} \mathrm{Re}(z - \mathbf{A})^{-1} & -\gamma \mathrm{Im}(z - \mathbf{A})^{-1} \\ \gamma^{-1}\mathrm{Im}(z - \mathbf{A})^{-1} & \mathrm{Re}(z - \mathbf{A})^{-1} \end{bmatrix} \right) \right)^{-1}$$

[Qiu, Berhardsson, Rantzer, Davison, Young, Doyle, 1995].

The real structured pseudospectrum can be computed via the definition

$$\sigma_\varepsilon^{\mathbf{R}}(\mathbf{A}) = \{z \in \mathbf{C} : d_{\mathbf{R}}(\mathbf{A}, z) < \varepsilon\}$$

[Karow, 2003].

Complex versus Real Perturbations: Example 1

Consider the following Toeplitz matrix studied by Demmel [1987]:

$$\mathbf{A} = \begin{bmatrix} -1 & -M & -M^2 & -M^3 & -M^4 \\ 0 & -1 & -M & -M^2 & -M^3 \\ 0 & 0 & -1 & -M & -M^2 \\ 0 & 0 & 0 & -1 & -M \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix}$$

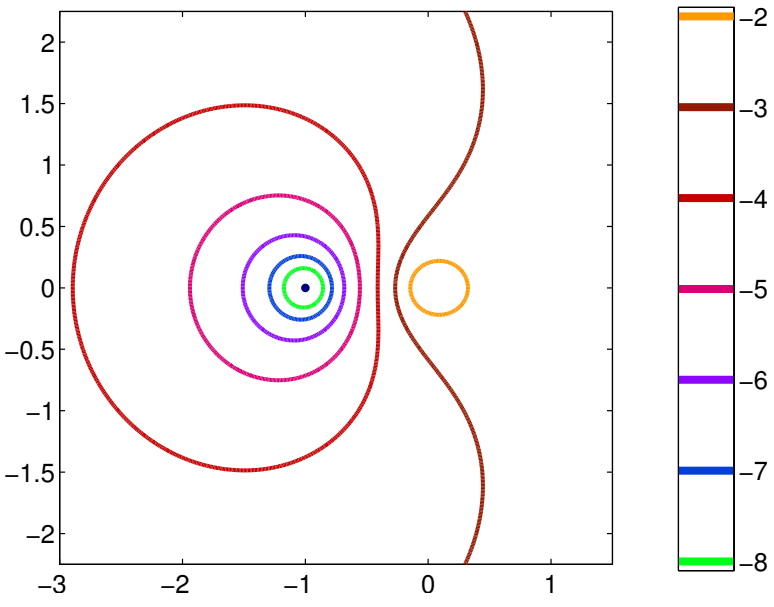
with $M = 10$.

The matrix is stable but small perturbations can move eigenvalues significantly.

How do the real structured pseudospectra compare
to the (unstructured) pseudospectra?

Complex versus Real Perturbations: Example 1

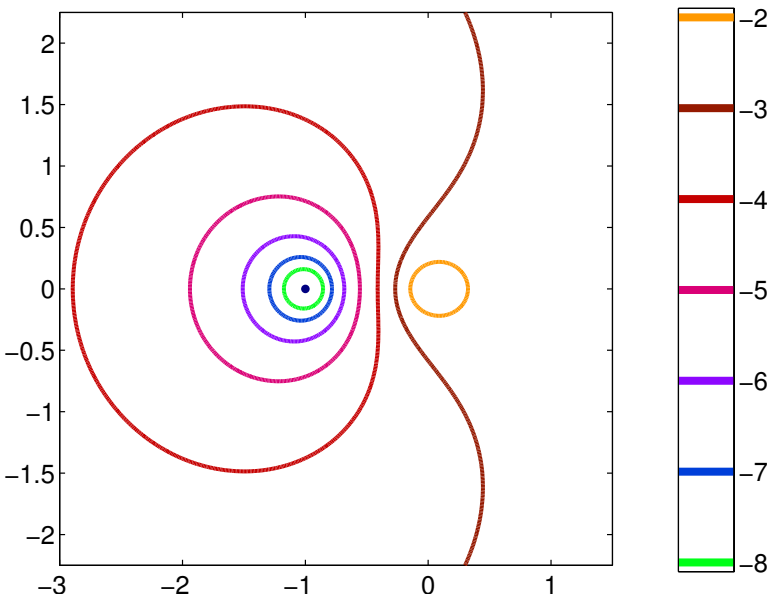
Complex and real pseudospectra for 5×5 matrix, $\varepsilon = 10^{-2}, 10^{-3}, \ldots, 10^{-8}$



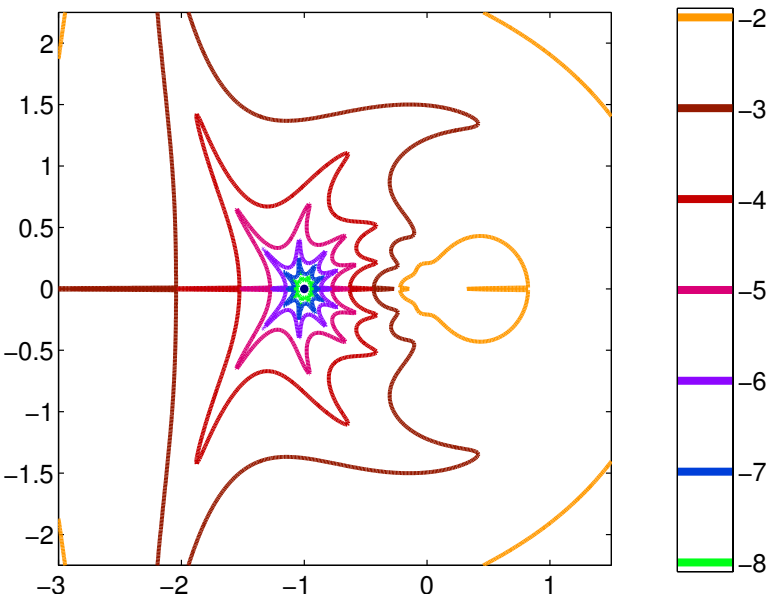
complex

Complex versus Real Perturbations: Example 1

Complex and real pseudospectra for 5×5 matrix, $\varepsilon = 10^{-2}, 10^{-3}, \ldots, 10^{-8}$



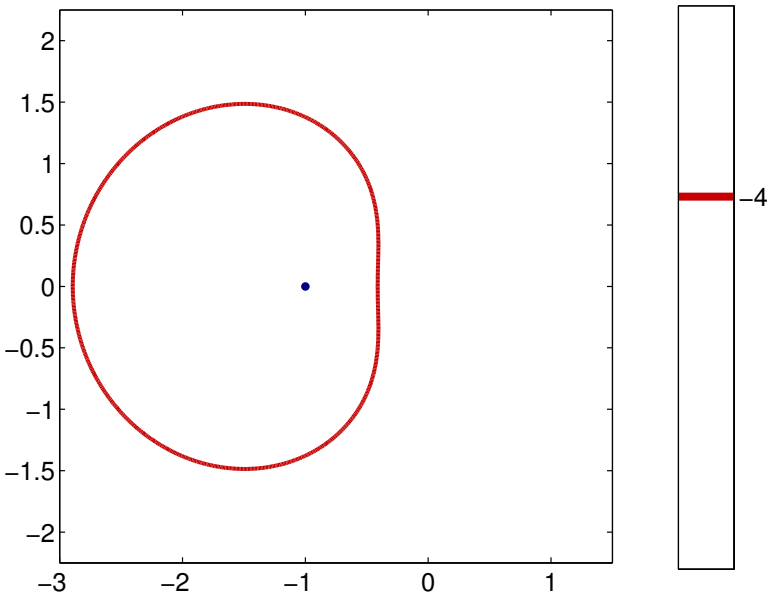
complex



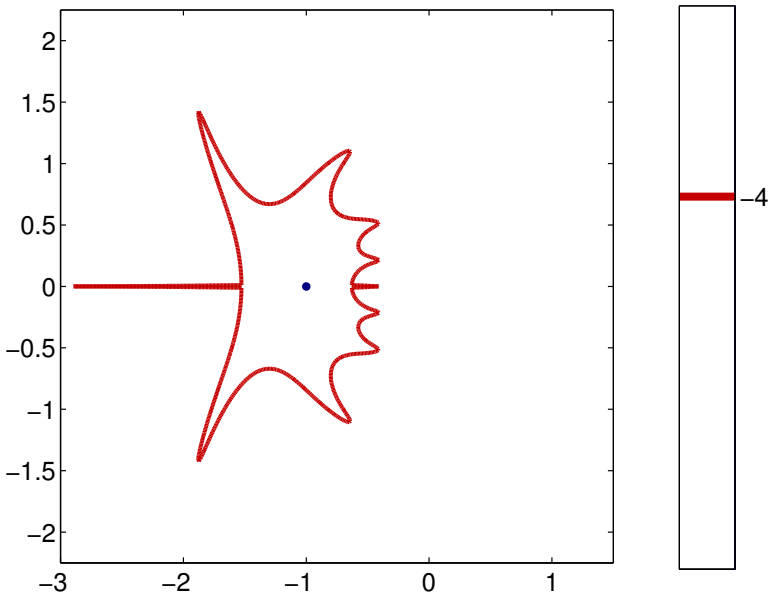
real

Complex versus Real Perturbations: Example 1

Complex and real pseudospectra for 5×5 matrix, $\varepsilon = 10^{-4}$



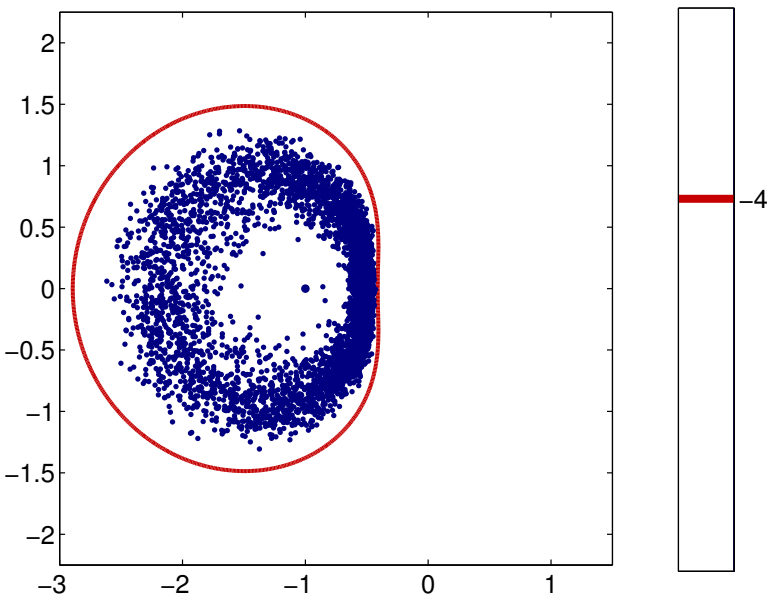
complex



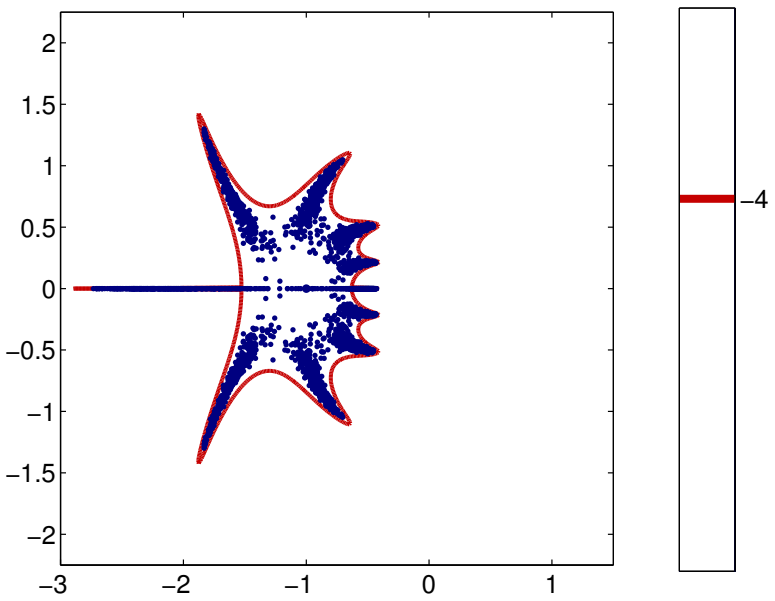
real

Complex versus Real Perturbations: Example 1

Complex and real pseudospectra for 5×5 matrix, $\varepsilon = 10^{-4}$



complex



real

eigenvalues of 1000 random perturbations of size 10^{-4}

Complex versus Real Perturbations: Example 2

Real perturbations need not describe *transient* behavior of dynamical systems.

Consider the matrix

$$\mathbf{A} = \begin{bmatrix} -1 & M^2 \\ -1 & -1 \end{bmatrix}, M \in \mathbf{R}$$

with spectrum $\sigma(\mathbf{A}) = \{-1 \pm \mathrm{i}M\}$. For $M = 100$:

Complex versus Real Perturbations: Example 2

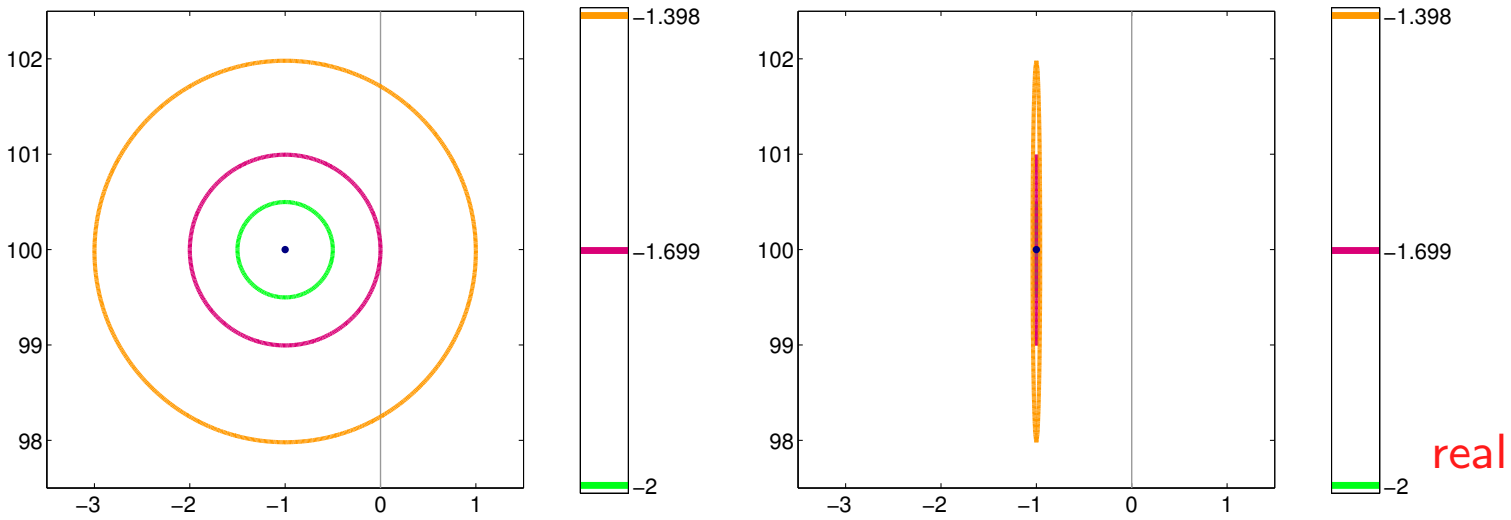
Real perturbations need not describe *transient* behavior of dynamical systems.

Consider the matrix

$$\mathbf{A} = \begin{bmatrix} -1 & M^2 \\ -1 & -1 \end{bmatrix}, M \in \mathbf{R}$$

with spectrum $\sigma(\mathbf{A}) = \{-1 \pm \mathrm{i}M\}$. For $M = 100$:

complex



real

Complex versus Real Perturbations: Example 2

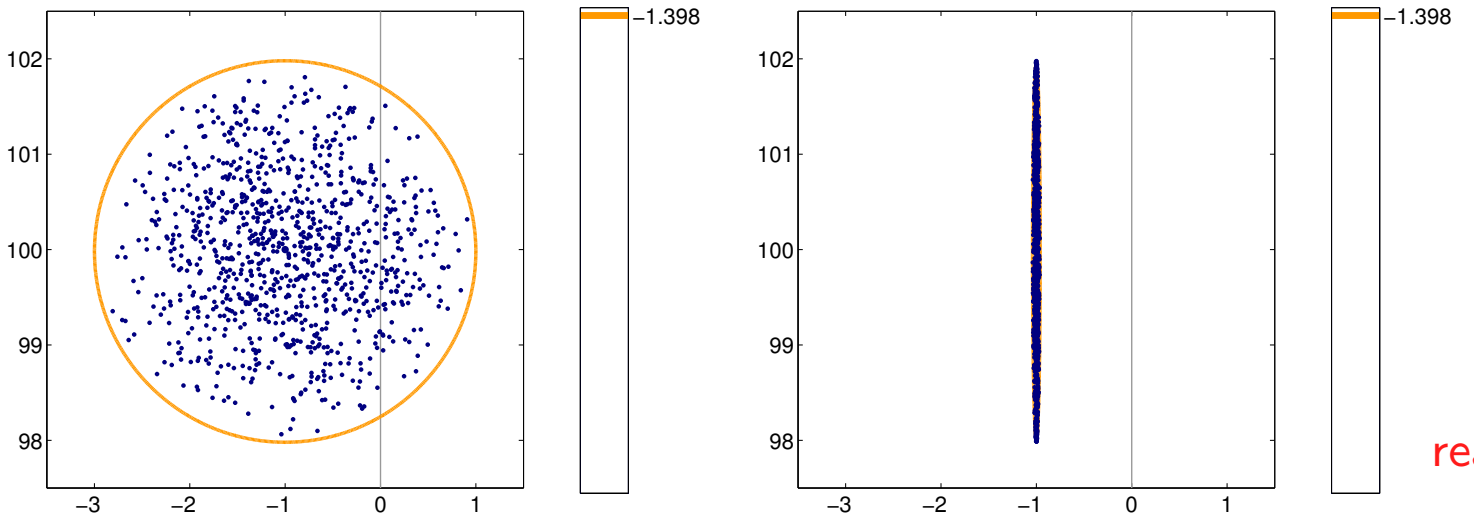
Real perturbations need not describe *transient* behavior of dynamical systems.

Consider the matrix

$$\mathbf{A} = \begin{bmatrix} -1 & M^2 \\ -1 & -1 \end{bmatrix}, M \in \mathbf{R}$$

with spectrum $\sigma(\mathbf{A}) = \{-1 \pm \mathrm{i}M\}$. For $M = 100$:

complex



real

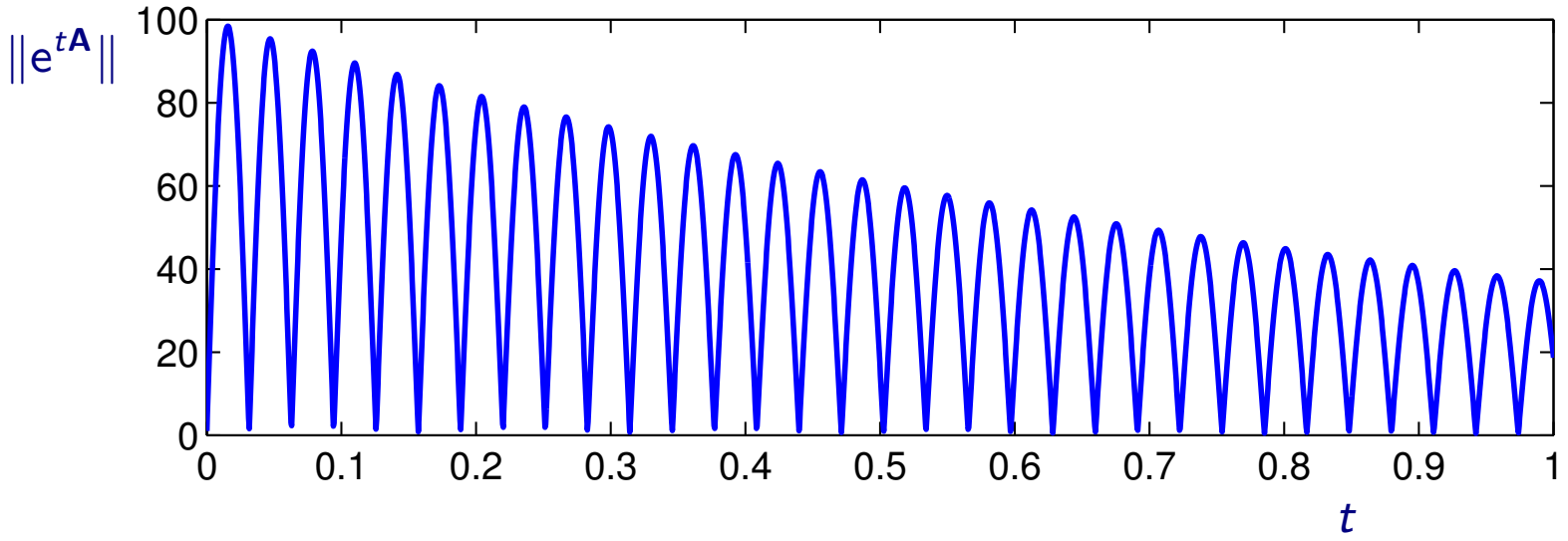
eigenvalues of 1000 random perturbations of size 400/10001

Complex versus Real Perturbations

Consider the dynamical system

$$\mathbf{x}'(t) = \begin{bmatrix} -1 & M^2 \\ -1 & -1 \end{bmatrix} \mathbf{x}(t)$$

with $M = 100$.



Real perturbations suggest this system is far from unstable,
yet transient growth on the order of M is observed.

4. Balanced Truncation, Lyapunov Equations

Balanced Truncation and Lyapunov Equations

Consider the SISO linear dynamical system:

$$\begin{aligned} \mathbf{x}'(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t) \\ \mathbf{y}(t) &= \mathbf{c}\mathbf{x}(t) + d\,u(t), \end{aligned}$$

$\mathbf{A} \in \mathbf{C}^{n\times n}$, $\mathbf{b}, \mathbf{c}^{\mathsf{T}} \in \mathbf{C}^n$, $d \in \mathbf{C}$. We assume that $\mathbf{A}$ is stable: $\alpha(\mathbf{A}) < 0$.

We wish to reduce the dimension of the dynamical system by projecting onto well-chosen subspaces.

Balanced truncation: Change basis to match states that are easy to *reach* and easy to *observe*, then project onto that prominent subspace.

Characterize how difficult it is to reach or observe a state via the infinite controllability and observability gramians $\mathbf{P}$ and $\mathbf{Q}$:

$$\mathbf{P} := \int_0^\infty \mathrm{e}^{t\mathbf{A}}\mathbf{b}\mathbf{b}^*\mathrm{e}^{t\mathbf{A}^*}\,\mathrm{d}t, \qquad \mathbf{Q} := \int_0^\infty \mathrm{e}^{t\mathbf{A}^*}\mathbf{c}^*\mathbf{c}\mathrm{e}^{t\mathbf{A}}\,\mathrm{d}t.$$

See, e.g., [Antoulas, 2005], Tuesday's lectures. . . .

Balanced Truncation and Lyapunov Equations

The gramians

$$\mathbf{P} := \int_0^\infty e^{t\mathbf{A}} \mathbf{b}\mathbf{b}^* e^{t\mathbf{A}^*} \, dt, \qquad \mathbf{Q} := \int_0^\infty e^{t\mathbf{A}^*} \mathbf{c}^* \mathbf{c} e^{t\mathbf{A}} \, dt$$

(Hermitian positive definite, for a controllable and observable stable system) can be determined by solving the Lyapunov equations

$$\mathbf{A}\mathbf{P} + \mathbf{P}\mathbf{A}^* = -\mathbf{b}\mathbf{b}^*, \qquad \mathbf{A}^*\mathbf{Q} + \mathbf{Q}\mathbf{A} = -\mathbf{c}^*\mathbf{c}.$$

If $\mathbf{x}_0 = \mathbf{0}$, the minimum energy of u required to drive $\mathbf{x}$ to state $\widehat{\mathbf{x}}$ is

$$\widehat{\mathbf{x}}^* \mathbf{P}^{-1} \widehat{\mathbf{x}}.$$

Starting from $\mathbf{x}_0 = \widehat{\mathbf{x}}$ with $u(t) \equiv 0$, the energy of output y is

$$\widehat{\mathbf{x}}^* \mathbf{Q} \widehat{\mathbf{x}}.$$

$\widehat{\mathbf{x}}^* \mathbf{P}^{-1} \widehat{\mathbf{x}}$: $\widehat{\mathbf{x}}$ is *hard to reach* if it is rich in the lowest modes of $\mathbf{P}$.
$\widehat{\mathbf{x}}^* \mathbf{Q} \widehat{\mathbf{x}}$: $\widehat{\mathbf{x}}$ is *hard to observe* if it is rich in the lowest modes of $\mathbf{Q}$.

Balanced truncation transforms the state space coordinate system to make these two gramians the same, then it truncates the lowest modes.

Balanced Truncation and Lyapunov Equations

Consider a generic coordinate transformation, for $\mathbf{S}$ invertible:

$$\begin{aligned}(\mathbf{S}\mathbf{x})'(t) &= (\mathbf{S}\mathbf{A}\mathbf{S}^{-1})(\mathbf{S}\mathbf{x}(t)) + (\mathbf{S}\mathbf{b})u(t) \\ y(t) &= (\mathbf{c}\mathbf{S}^{-1})(\mathbf{S}\mathbf{x}(t)) + du(t), \qquad (\mathbf{S}\mathbf{x})(0) = \mathbf{S}\mathbf{x}_0.\end{aligned}$$

With this transformation, the controllability and observability gramians are

$$\widehat{\mathbf{P}} = \mathbf{S}\mathbf{P}\mathbf{S}^*, \qquad \widehat{\mathbf{Q}} = \mathbf{S}^{-*}\mathbf{Q}\mathbf{S}^{-1}.$$

For balancing, we seek $\mathbf{S}$ so that $\widehat{\mathbf{P}} = \widehat{\mathbf{Q}}$ are diagonal.

Balanced Truncation and Lyapunov Equations

Consider a generic coordinate transformation, for $\mathbf{S}$ invertible:

$$\begin{aligned}(\mathbf{S}\mathbf{x})'(t) &= (\mathbf{S}\mathbf{A}\mathbf{S}^{-1})(\mathbf{S}\mathbf{x}(t)) + (\mathbf{S}\mathbf{b})u(t)\\ y(t) &= (\mathbf{c}\mathbf{S}^{-1})(\mathbf{S}\mathbf{x}(t)) + du(t), \qquad (\mathbf{S}\mathbf{x})(0) = \mathbf{S}\mathbf{x}_0.\end{aligned}$$

With this transformation, the controllability and observability gramians are

$$\widehat{\mathbf{P}} = \mathbf{S}\mathbf{P}\mathbf{S}^*, \qquad \widehat{\mathbf{Q}} = \mathbf{S}^{-*}\mathbf{Q}\mathbf{S}^{-1}.$$

For balancing, we seek $\mathbf{S}$ so that $\widehat{\mathbf{P}} = \widehat{\mathbf{Q}}$ are diagonal.

Observation (How does nonnormality affect balancing?)

- $\sigma_{\varepsilon/\kappa(\mathbf{S})}(\mathbf{S}\mathbf{A}\mathbf{S}^{-1}) \subseteq \sigma_\varepsilon(\mathbf{A}) \subseteq \sigma_{\varepsilon\kappa(\mathbf{S})}(\mathbf{S}\mathbf{A}\mathbf{S}^{-1})$.
- *The choice of internal coordinates will affect* $\mathbf{P}$, $\mathbf{Q}$, ...

Balanced Truncation and Lyapunov Equations

Consider a generic coordinate transformation, for $\mathbf{S}$ invertible:

$$\begin{aligned}(\mathbf{S}\mathbf{x})'(t) &= (\mathbf{S}\mathbf{A}\mathbf{S}^{-1})(\mathbf{S}\mathbf{x}(t)) + (\mathbf{S}\mathbf{b})u(t)\\ y(t) &= (\mathbf{c}\mathbf{S}^{-1})(\mathbf{S}\mathbf{x}(t)) + du(t), \qquad (\mathbf{S}\mathbf{x})(0) = \mathbf{S}\mathbf{x}_0.\end{aligned}$$

With this transformation, the controllability and observability gramians are

$$\widehat{\mathbf{P}} = \mathbf{S}\mathbf{P}\mathbf{S}^*, \qquad \widehat{\mathbf{Q}} = \mathbf{S}^{-*}\mathbf{Q}\mathbf{S}^{-1}.$$

For balancing, we seek $\mathbf{S}$ so that $\widehat{\mathbf{P}} = \widehat{\mathbf{Q}}$ are diagonal.

Observation (How does nonnormality affect balancing?)

- $\sigma_{\varepsilon/\kappa(\mathbf{S})}(\mathbf{S}\mathbf{A}\mathbf{S}^{-1}) \subseteq \sigma_\varepsilon(\mathbf{A}) \subseteq \sigma_{\varepsilon\kappa(\mathbf{S})}(\mathbf{S}\mathbf{A}\mathbf{S}^{-1})$.
- *The choice of internal coordinates will affect* $\mathbf{P}$, $\mathbf{Q}$, ...
- *but not the Hankel singular values:* $\widehat{\mathbf{P}}\widehat{\mathbf{Q}} = \mathbf{S}\mathbf{P}\mathbf{Q}\mathbf{S}^{-1}$,
- *and not the transfer function:*
$$d + (\mathbf{c}\mathbf{S}^{-1})(z - \mathbf{S}\mathbf{A}\mathbf{S}^{-1})^{-1}(\mathbf{S}\mathbf{b}) = d + \mathbf{c}(z - \mathbf{A})^{-1}\mathbf{b},$$
- *and not the system moments:*
$$d = d, \quad (\mathbf{c}\mathbf{S}^{-1})(\mathbf{S}\mathbf{b}) = \mathbf{c}\mathbf{b}, \quad (\mathbf{c}\mathbf{S}^{-1})(\mathbf{S}\mathbf{A}\mathbf{S}^{-1})(\mathbf{S}\mathbf{b}) = \mathbf{c}\mathbf{A}\mathbf{b}, \quad \ldots.$$

Decay of Singular Values of Lyapunov Solutions

Consider the Lyapunov equation

$$\mathbf{A}\mathbf{X} + \mathbf{X}\mathbf{A}^* = -\mathbf{b}\mathbf{b}^*.$$

for stable $\mathbf{A}$ and controllable $(\mathbf{A}, \mathbf{b})$.

In many cases (since the right-hand side $\mathbf{b}\mathbf{b}^*$ is low-rank), the singular values of $\mathbf{X}$ *decay rapidly* [Gudmundsson and Laub 1994; Penzl 2000].

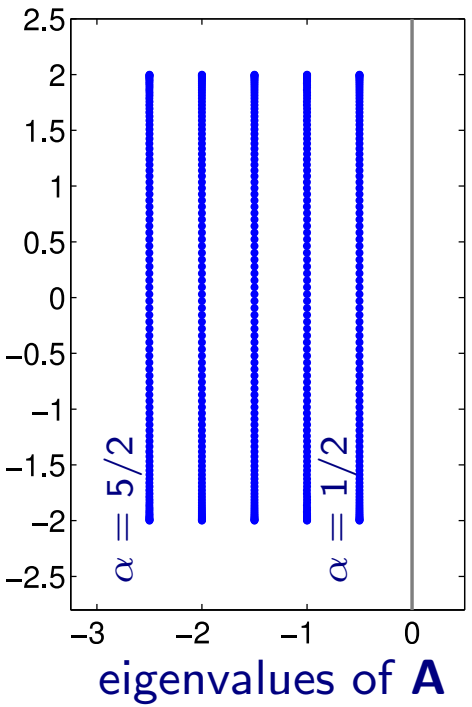
Decay of Singular Values of Lyapunov Solutions

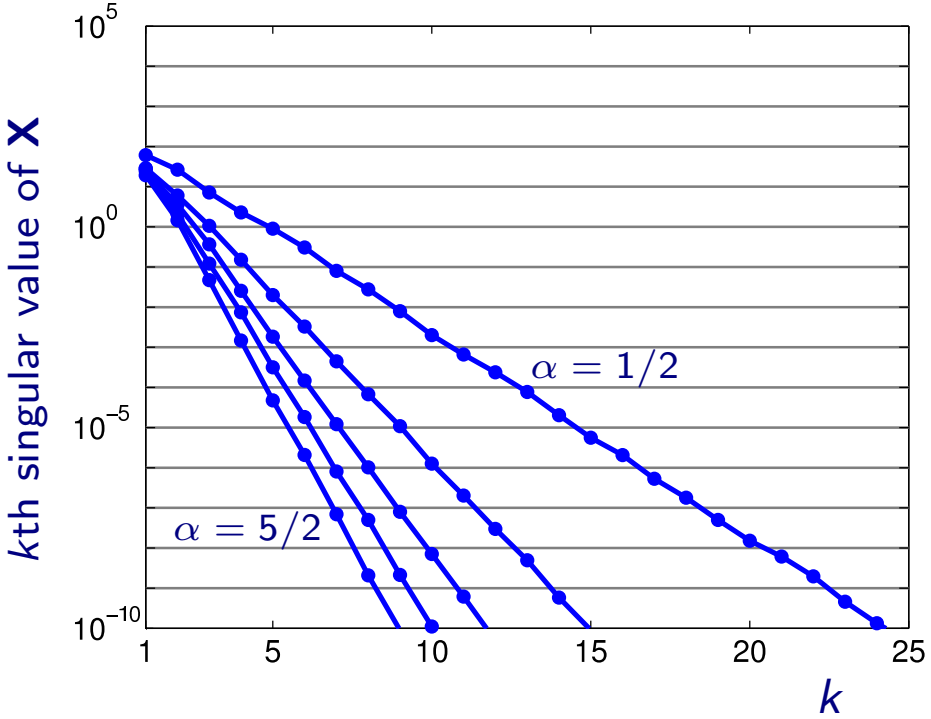
Consider the Lyapunov equation

$$\mathbf{A}\mathbf{X} + \mathbf{X}\mathbf{A}^* = -\mathbf{b}\mathbf{b}^*.$$

for stable $\mathbf{A}$ and controllable $(\mathbf{A}, \mathbf{b})$.

In many cases (since the right-hand side $\mathbf{b}\mathbf{b}^*$ is low-rank), the singular values of $\mathbf{X}$ *decay rapidly* [Gudmundsson and Laub 1994; Penzl 2000].





Normal example: $\mathbf{A} = \text{tridiag}(-1, \alpha, 1)$ with spectrum

$$\sigma(\mathbf{A}) \subseteq \{\alpha + \mathrm{i}y : y \in [-2, 2]\}.$$

Decay of Singular Values of Lyapunov Solutions

Let s_k denote the kth singular value of $\mathbf{X}$, $s_k \geq s_{k+1}$.

- For Hermitian $\mathbf{A}$, bounds on s_k/s_1 have been derived by Penzl [2000], Sabino [2006] (see Ellner & Wachspress [1991]).

- For diagonalizable $\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}$, Antoulas, Sorensen, and Zhou [2002] prove

$$s_{k+1} \leq \kappa(\mathbf{V})^2 \delta_{k+1} \|\mathbf{b}\|^2 (N-k)^2, \qquad \delta_k = \frac{-1}{2\operatorname{Re}\lambda_k} \prod_{j=1}^{k-1} \frac{|\lambda_k - \lambda_j|^2}{|\overline{\lambda}_k + \lambda_j|^2}.$$

  The $\kappa(\mathbf{V})^2$ term imposes a significant penalty for nonnormality; cf. [Truhar, Tomljanović, Li 2009].

  We seek a bound that gives a more flexible approach, by enlarging the set over which we study rational functions like δ_k.

Decay of Singular Values of Lyapunov Solutions

- $$s_{k+1} \le \kappa(\mathbf{V})^2 \delta_{k+1} \|\mathbf{b}\|^2 (N-k)^2, \qquad \delta_k = \frac{-1}{2\mathrm{Re}\,\lambda_k} \prod_{j=1}^{k-1} \frac{|\lambda_k - \lambda_j|^2}{|\overline{\lambda}_k + \lambda_j|^2}.$$

- The ADI iteration [Wachspress 1988; Penzl 2000] constructs a rank-k approximation $\mathbf{X}_k$ to $\mathbf{X}$ that satisfies

$$\mathbf{X} - \mathbf{X}_k \;=\; \phi(\mathbf{A}) \mathbf{X} \phi(\mathbf{A})^*,$$

with $\mathbf{X}_0 = \mathbf{0}$ and

$$\phi(z) \;=\; \prod_{j=1}^{k} \frac{\overline{\mu}_j - z}{\mu_j + z}.$$

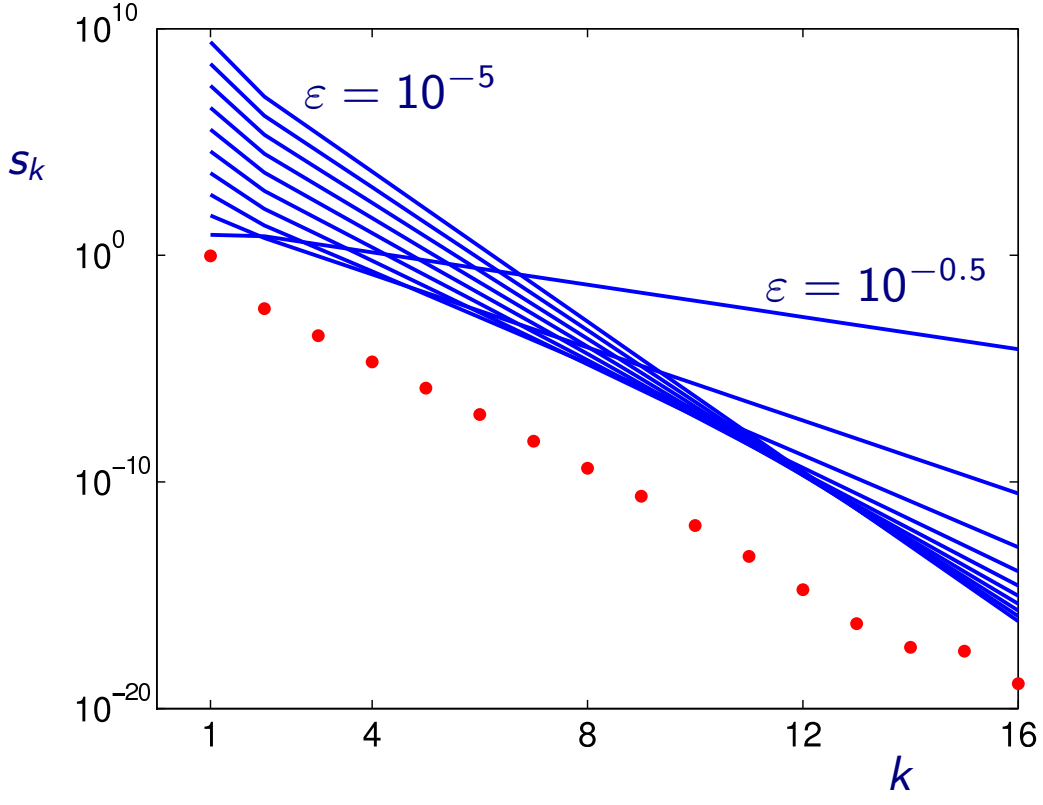
Using the fact that $\mathrm{rank}(\mathbf{X}_k) \le k$, we obtain the bound

$$\frac{s_{k+1}}{s_1} \;\le\; \min_{\mu_1,\ldots,\mu_k} \|\phi(\mathbf{A})\|_2^2 \;\le\; \left(\frac{L_\varepsilon}{2\pi\varepsilon}\right)^2 \min_{\mu_1,\ldots,\mu_k} \sup_{z \in \sigma_\varepsilon(A)} \prod_{j=1}^{k} \frac{|\overline{\mu}_j - z|}{|\mu_j + z|},$$

where L_ε is the boundary length of a contour enclosing $\sigma_\varepsilon(\mathbf{A})$ [Beattie, E., Sabino; cf. Beckermann (2004)].

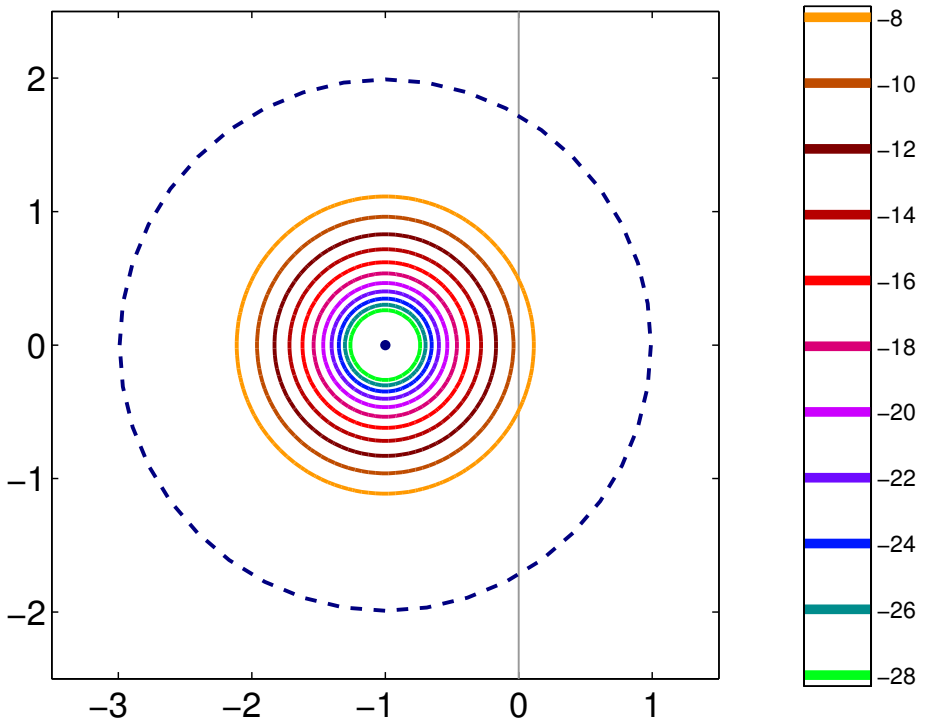
Examples of the Decay Bound Based on Pseudospectra

$$\mathbf{A} = \begin{bmatrix} -1 & 1/2 & & & \\ & -1 & \ddots & & \\ & & \ddots & 1/2 & \\ & & & -1 & 1/2 \\ & & & & -1 \end{bmatrix} \in \mathbf{C}^{16\times 16}$$



Examples of the Decay Bound Based on Pseudospectra

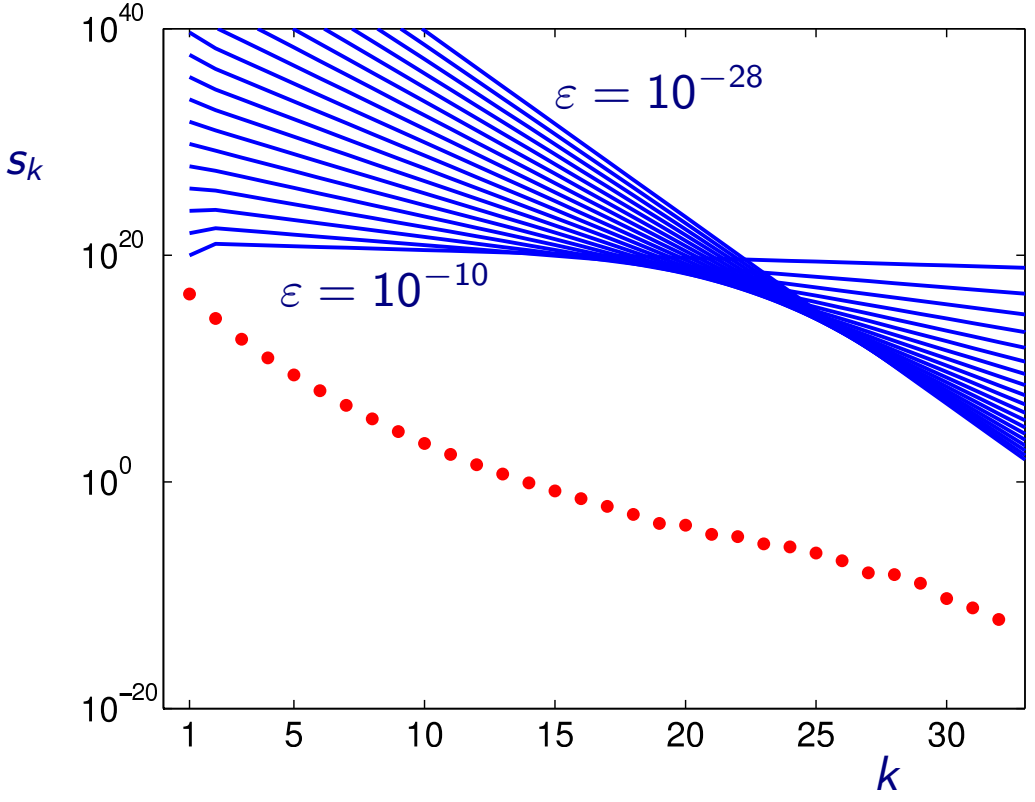
$$\mathbf{A} = \begin{bmatrix} -1 & 2 & & & \\ & -1 & \ddots & & \\ & & \ddots & 2 & \\ & & & -1 & 2 \\ & & & & -1 \end{bmatrix} \in \mathbf{C}^{32\times 32}$$



Spectrum, pseudospectra, and numerical range (boundary = dashed line)

Examples of the Decay Bound Based on Pseudospectra

$$\mathbf{A} = \begin{bmatrix} -1 & 2 & & & \\ & -1 & \ddots & & \\ & & \ddots & 2 & \\ & & & -1 & 2 \\ & & & & -1 \end{bmatrix} \in \mathbf{C}^{32\times 32}$$



Any Decay Possible for Any Spectrum

Theorem (Penzl, 2000)

Let $\mathbf{A}$ *be stable and* $\mathbf{b}$ *some vector such that* $(\mathbf{A}, \mathbf{b})$ *is controllable. Given any Hermitian positive definite* $\mathbf{X}$*, there exists some invertible matrix* $\mathbf{S}$ *such that*

$$(\mathbf{SAS}^{-1})\mathbf{X} + \mathbf{X}(\mathbf{SAS}^{-1}) = -(\mathbf{Sb})(\mathbf{Sb})^*.$$

Any prescribed singular value decay is possible for a matrix with any eigenvalues.

Proof. The proof is a construction.

- Solve $\mathbf{AY} + \mathbf{YA}^* = -\mathbf{bb}^*$ for $\mathbf{Y}$.
  ($\mathbf{Y}$ is Hermitian positive definite, since $(\mathbf{A}, \mathbf{b})$ controllable.)
- Set $\mathbf{S} := \mathbf{X}^{1/2}\mathbf{Y}^{-1/2}$.
- Notice that $\mathbf{SYS}^* = \mathbf{X}^{1/2}\mathbf{Y}^{-1/2}\mathbf{YY}^{-1/2}\mathbf{X}^{1/2} = \mathbf{X}$.
- Define $\widehat{\mathbf{A}} := \mathbf{SAS}^{-1}$, $\widehat{\mathbf{b}} := \mathbf{Sb}$.

Now it is easy to verify that $\mathbf{X}$ solves the desired Lyapunov equation:

$$\begin{aligned} \widehat{\mathbf{A}}\mathbf{X} + \mathbf{X}\widehat{\mathbf{A}}^* &= (\mathbf{SAS}^{-1})(\mathbf{SYS}^*) + (\mathbf{SYS}^*)(\mathbf{S}^{-*}\mathbf{AS}^*) \\ &= \mathbf{S}(\mathbf{AY} + \mathbf{YA})\mathbf{S}^* = -(\mathbf{Sb})(\mathbf{Sb})^* = -\widehat{\mathbf{b}}\widehat{\mathbf{b}}^*. \quad \square \end{aligned}$$

A Nonnormal Anomaly

The pseudospectral bound and the bound of Antoulas, Sorensen, and Zhou both predict that *the decay rate slows* as nonnormality increases.

However, for the solutions to Lyapunov equations this intuition can be wrong [Sabino 2006].

A Nonnormal Anomaly

The pseudospectral bound and the bound of Antoulas, Sorensen, and Zhou both predict that *the decay rate slows* as nonnormality increases.

However, for the solutions to Lyapunov equations this intuition can be wrong [Sabino 2006].

Consider

$$\mathbf{A} = \begin{bmatrix} -1 & \alpha \\ 0 & -1 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} t \\ 1 \end{bmatrix}, \quad t \in \mathbf{R}.$$

As α grows, A's departure from normality grows.
All bounds suggest that the 'decay' rate should worsen as α increases.

A Nonnormal Anomaly

The pseudospectral bound and the bound of Antoulas, Sorensen, and Zhou both predict that *the decay rate slows* as nonnormality increases.

However, for the solutions to Lyapunov equations this intuition can be wrong [Sabino 2006].

Consider

$$\mathbf{A} = \begin{bmatrix} -1 & \alpha \\ 0 & -1 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} t \\ 1 \end{bmatrix}, \quad t \in \mathbf{R}.$$

As α grows, A's departure from normality grows.
All bounds suggest that the 'decay' rate should worsen as α increases.

The Lyapunov equation $\mathbf{AX} + \mathbf{XA}^* = -\mathbf{bb}^*$ has the solution

$$\mathbf{X} = \frac{1}{4} \begin{bmatrix} 2t^2 + 2\alpha t + \alpha^2 & \alpha + 2t \\ \alpha + 2t & 2 \end{bmatrix}.$$

For each α, we wish to pick t to maximize the 'decay', i.e., s_2/s_1.

A Nonnormal Anomaly

The pseudospectral bound and the bound of Antoulas, Sorensen, and Zhou both predict that *the decay rate slows* as nonnormality increases.

However, for the solutions to Lyapunov equations this intuition can be wrong [Sabino 2006].

Consider

$$\mathbf{A} = \begin{bmatrix} -1 & \alpha \\ 0 & -1 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} t \\ 1 \end{bmatrix}, \quad t \in \mathbf{R}.$$

As α grows, A's departure from normality grows.
All bounds suggest that the 'decay' rate should worsen as α increases.

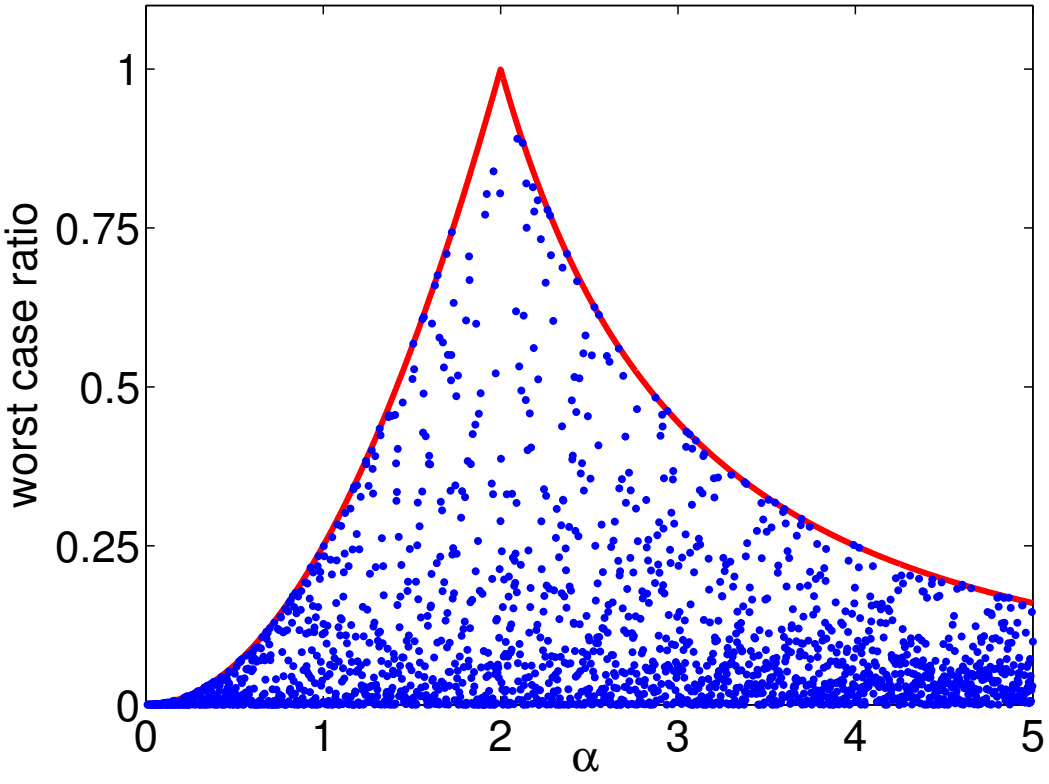
The Lyapunov equation $\mathbf{AX} + \mathbf{XA}^* = -\mathbf{bb}^*$ has the solution

$$\mathbf{X} = \frac{1}{4} \begin{bmatrix} 2t^2 + 2\alpha t + \alpha^2 & \alpha + 2t \\ \alpha + 2t & 2 \end{bmatrix}.$$

For each α, we wish to pick t to maximize the 'decay', i.e., s_2/s_1.
This is accomplished for $t = -\alpha/2$, giving

$$\frac{s_2}{s_1} = \begin{cases} \alpha^2/4, & 0 < \alpha \leq 2; \\ 4/\alpha^2, & 2 \leq \alpha. \end{cases}$$

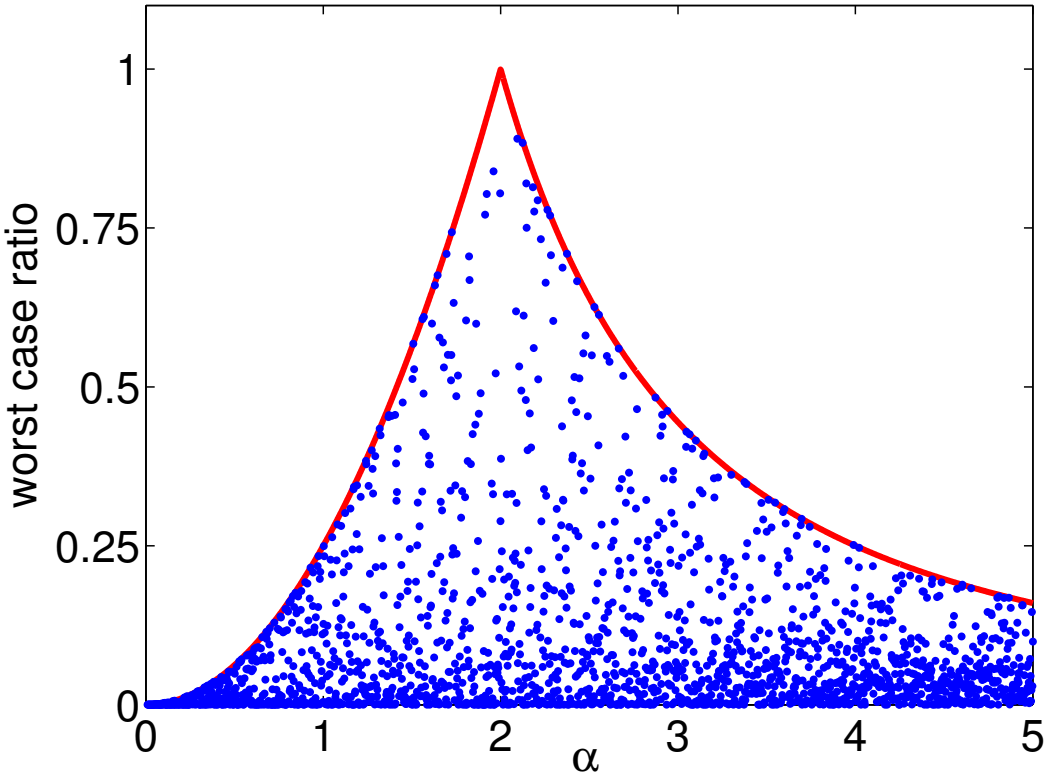
Jordan block: 2×2 case, numerical illustration



Red line: $\dfrac{s_2}{s_1} = \begin{cases} \alpha^2/4, & 0 < \alpha \leq 2; \\ 4/\alpha^2, & 2 \leq \alpha. \end{cases}$

Blue dots: s_2/s_1 for random $\mathbf{b}$ and 2000 α values.

Jordan block: 2×2 case, numerical illustration



Red line: $\dfrac{s_2}{s_1} = \begin{cases} \alpha^2/4, & 0 < \alpha \leq 2; \\ 4/\alpha^2, & 2 \leq \alpha. \end{cases}$

Blue dots: s_2/s_1 for random $\mathbf{b}$ and 2000 α values.

The effect of nonnormality on Lyapunov solutions remains only partially understood.

5. Moment Matching Model Reduction

Krylov methods for moment matching model reduction

We now turn to a model reduction approach where nonnormality plays a crucial role. Once again, begin with the SISO system

$$\begin{aligned} \mathbf{x}'(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t) \\ y(t) &= \mathbf{c}\mathbf{x}(t) + du(t), \end{aligned}$$

with $\mathbf{A} \in \mathbf{C}^{n\times n}$ and $\mathbf{b}, \mathbf{c}^{\mathsf{T}} \in \mathbf{C}^n$ and initial condition $\mathbf{x}(0) = \mathbf{x}_0$.

Compress the state space to dimension $k \ll n$ via a projection method:

$$\begin{aligned} \widehat{\mathbf{x}}'(t) &= \widehat{\mathbf{A}}\widehat{\mathbf{x}}(t) + \widehat{\mathbf{b}}u(t) \\ \widehat{y}(t) &= \widehat{\mathbf{c}}\widehat{\mathbf{x}}(t) + du(t), \end{aligned}$$

where

$$\widehat{\mathbf{A}} = \mathbf{W}^*\mathbf{A}\mathbf{V} \in \mathbf{C}^{k\times k}, \qquad \widehat{\mathbf{b}} = \mathbf{W}^*\mathbf{b} \in \mathbf{C}^{k\times 1}, \qquad \widehat{\mathbf{c}} = \mathbf{c}\mathbf{V} \in \mathbf{C}^{1\times k}$$

for some $\mathbf{V}, \mathbf{W} \in \mathbf{C}^{n\times k}$ with $\mathbf{W}^*\mathbf{V} = \mathbf{I}$.

The matrices $\mathbf{V}$ and $\mathbf{W}$ are constructed by a Krylov subspace method.

Krylov methods for moment matching model reduction

$$\widehat{\mathbf{A}} = \mathbf{W}^*\mathbf{A}\mathbf{V} \in \mathbf{C}^{k\times k}, \qquad \widehat{\mathbf{b}} = \mathbf{W}^*\mathbf{b} \in \mathbf{C}^{k\times 1}, \qquad \widehat{\mathbf{c}} = \mathbf{c}\mathbf{V} \in \mathbf{C}^{1\times k}$$

Arnoldi Reduction

If $\mathbf{V} = \mathbf{W}$ *and the columns of* $\mathbf{V}$ *span the* k*th Krylov subspace,*

$$\mathrm{Ran}(\mathbf{V}) = \mathcal{K}_k(\mathbf{A}, \mathbf{b}) = \mathrm{span}\{\mathbf{b}, \mathbf{A}\mathbf{b}, \ldots, \mathbf{A}^{k-1}\mathbf{b}\},$$

then the reduced model matches k *moments of the system:*

$$\widehat{\mathbf{c}}\widehat{\mathbf{A}}^j\widehat{\mathbf{b}} = \mathbf{c}\mathbf{A}^j\mathbf{b}, \qquad j = 0, \ldots, k-1.$$

Bi-Lanczos Reduction

If the columns and $\mathbf{V}$ *and* $\mathbf{W}$ *span the Krylov subspaces*

$$\begin{aligned}\mathrm{Ran}(\mathbf{V}) &= \mathcal{K}_k(\mathbf{A}, \mathbf{b}) = \mathrm{span}\{\mathbf{b}, \mathbf{A}\mathbf{b}, \ldots, \mathbf{A}^{k-1}\mathbf{b}\}\\ \mathrm{Ran}(\mathbf{W}) &= \mathcal{K}_k(\mathbf{A}^*, \mathbf{c}^*) = \mathrm{span}\{\mathbf{c}^*, \mathbf{A}^*\mathbf{c}^*, \ldots, (\mathbf{A}^*)^{k-1}\mathbf{c}^*\},\end{aligned}$$

then the reduced model matches $2k$ *moments of the system:*

$$\widehat{\mathbf{c}}\widehat{\mathbf{A}}^j\widehat{\mathbf{b}} = \mathbf{c}\mathbf{A}^j\mathbf{b}, \qquad j = 0, \ldots, 2k-1.$$

Stability of Reduced Order Models

Question

Does the reduced model inherit properties of the original system?

Properties include stability, passivity, second-order structure, etc.

In this lecture we are concerned with *stability* – and, more generally, the behavior of eigenvalues of the reduced matrix $\widehat{\mathbf{A}}$.

Stability of Reduced Order Models

Question

Does the reduced model inherit properties of the original system?

Properties include stability, passivity, second-order structure, etc.

In this lecture we are concerned with *stability* – and, more generally, the behavior of eigenvalues of the reduced matrix $\widehat{\mathbf{A}}$.

Observation

For Arnoldi reduction, the eigenvalues of $\widehat{\mathbf{A}}$ are contained in the numerical range

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

Proof: If $\widehat{\mathbf{A}}\mathbf{z} = \theta\mathbf{z}$ for $\|\mathbf{z}\| = 1$, then $\theta = \mathbf{z}^*\mathbf{V}^*\mathbf{A}\mathbf{V}^*\mathbf{z} = (\mathbf{V}\mathbf{z})^*\mathbf{A}(\mathbf{V}\mathbf{z})$, where $\|\mathbf{V}\mathbf{z}\| = \|\mathbf{z}\| = 1$ □

Stability of Reduced Order Models

Question

Does the reduced model inherit properties of the original system?

Properties include stability, passivity, second-order structure, etc.

In this lecture we are concerned with *stability* – and, more generally, the behavior of eigenvalues of the reduced matrix $\widehat{\mathbf{A}}$.

Observation

For Arnoldi reduction, the eigenvalues of $\widehat{\mathbf{A}}$ are contained in the numerical range

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

Proof: If $\widehat{\mathbf{A}}\mathbf{z} = \theta\mathbf{z}$ for $\|\mathbf{z}\| = 1$, then $\theta = \mathbf{z}^*\mathbf{V}^*\mathbf{A}\mathbf{V}^*\mathbf{z} = (\mathbf{V}\mathbf{z})^*\mathbf{A}(\mathbf{V}\mathbf{z})$, where $\|\mathbf{V}\mathbf{z}\| = \|\mathbf{z}\| = 1$ □

No such bound for Bi-Lanczos: the decomposition may not even exist.
For the famous 'CD Player' model, $\mathbf{cb} = \mathbf{0}$: method breaks down at first step.

Stability and Bi-Lanczos

For Bi-Lanczos, the stability question is rather more subtle.

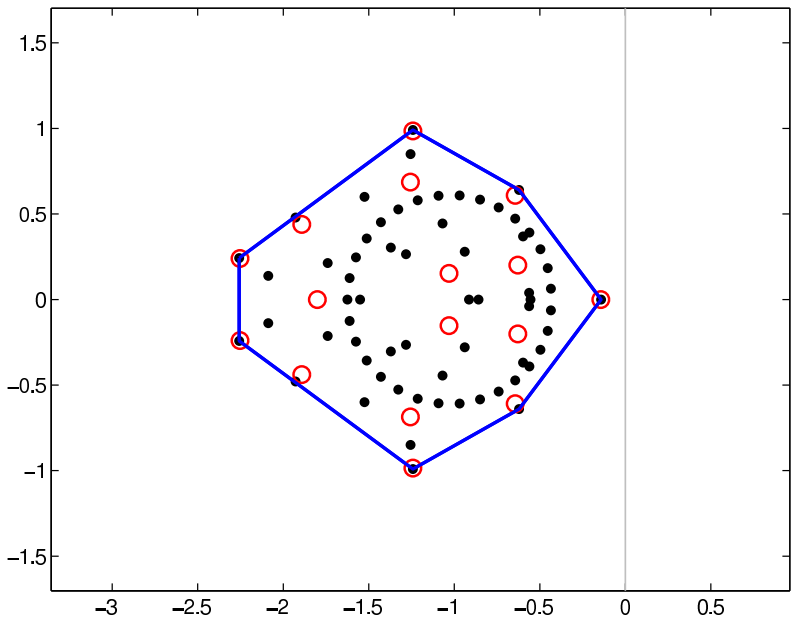
For example, if $\mathbf{b}$ and $\mathbf{c}$ are nearly orthogonal,
the (1,1) entry in $\widehat{\mathbf{A}} = \mathbf{W}^* \mathbf{A} \mathbf{V}$ will generally be very large.

Given a fixed $\mathbf{b}$, one has much freedom to rig poor results via $\mathbf{c}$:

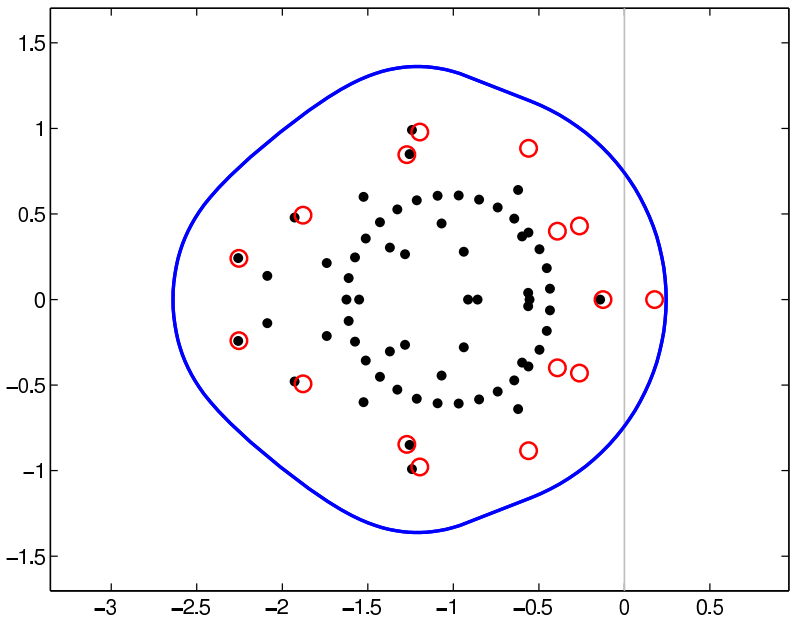
> [A]ny three-term recurrence (run for no more than $(n+2)/2$ steps, where n is the size of the matrix) *is* the two-sided Lanczos algorithm for some left starting vector. [Greenbaum, 1998]

Example: Arnoldi Reduction for Normal versus Nonnormal

The eigenvalues of $\widehat{\mathbf{A}}$ are known as *Ritz values*.



normal matrix



nonnormal matrix

Eigenvalues (•), Ritz values (○), and numerical range for isospectral matrices.

A Remedy for Unstable Arnoldi Models?

One can counteract instability by *restarting* the Arnoldi algorithm to shift out unstable eigenvalues [Grimme, Sorensen, Van Dooren, 1994]; cf. [Jaimoukha and Kasenally, 1997].

- $\widehat{\mathbf{A}} := \mathbf{V}^*\mathbf{A}\mathbf{V}$ has eigenvalues $\theta_1, \ldots, \theta_k$ (Ritz values for $\mathbf{A}$)
- Suppose $\theta_1, \ldots, \theta_p$ are in the right half plane.
- Replace the starting vector $\mathbf{b}$ by the filtered vector

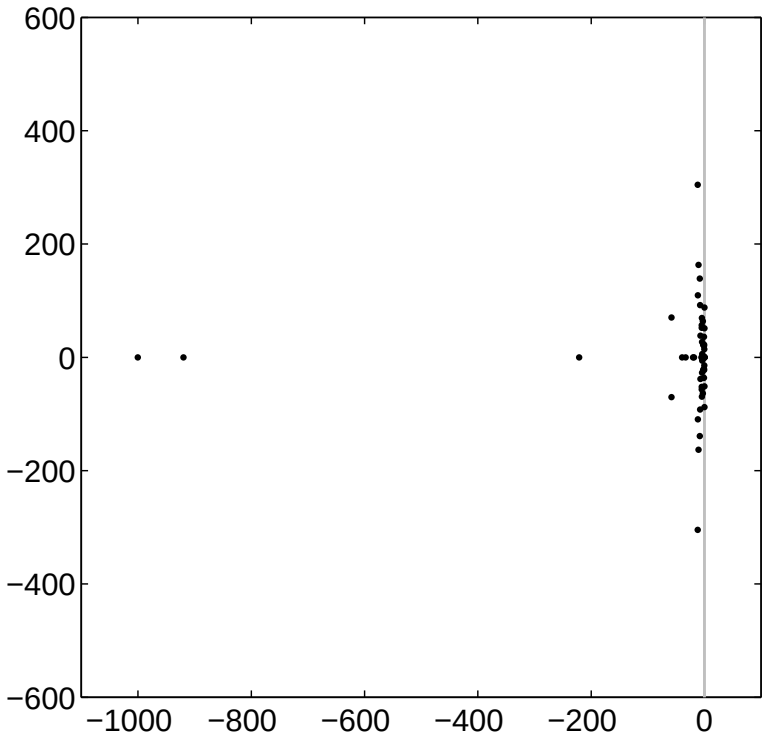
$$\mathbf{b}_+ = \psi(\mathbf{A})\mathbf{b}, \qquad \psi(z) = \prod_{j=1}^{p}(z - \theta_j),$$

  where the *filter polynomial* ψ "discourages" future Ritz values near the *shifts* $\theta_1, \ldots, \theta_p$.
- Build new matrices $\mathbf{V}$, $\widehat{\mathbf{A}}$ with starting vector $\mathbf{b}_+$ (implicit restart).
- Now *modified* moments, $\mathbf{b}^*\psi(\mathbf{A})^*\mathbf{A}^j\psi(\mathbf{A})\mathbf{b}$, will be matched.

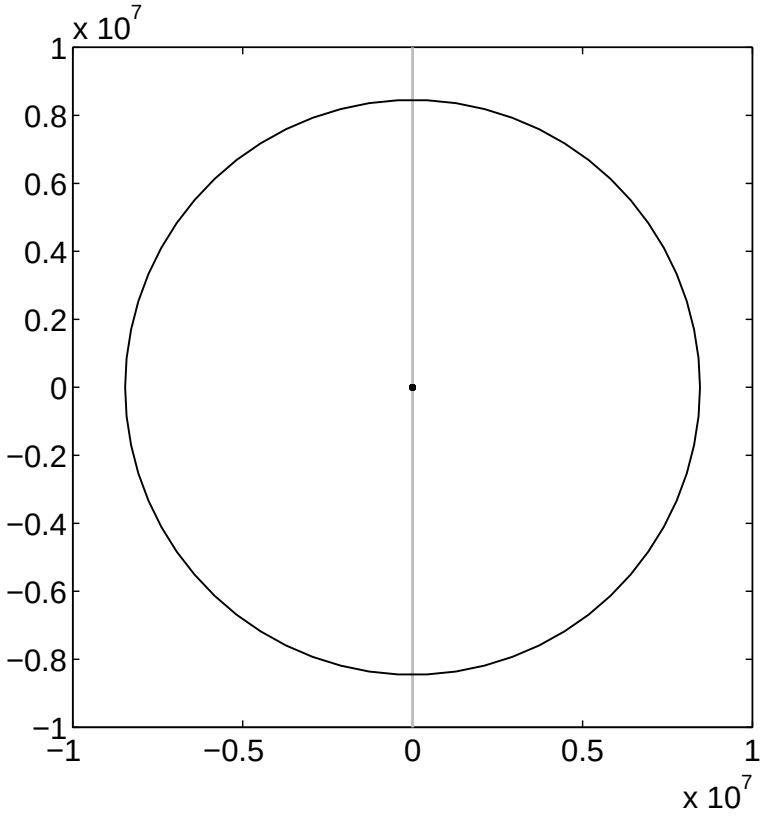
Repeat this process until $\widehat{\mathbf{A}}$ has no unstable modes.

Matching the Moments of a Nonnormal Matrix

Model of flutter in a Boeing 767 from SLICOT ($n = 55$), stabilized by Burke, Lewis, Overton [2003].



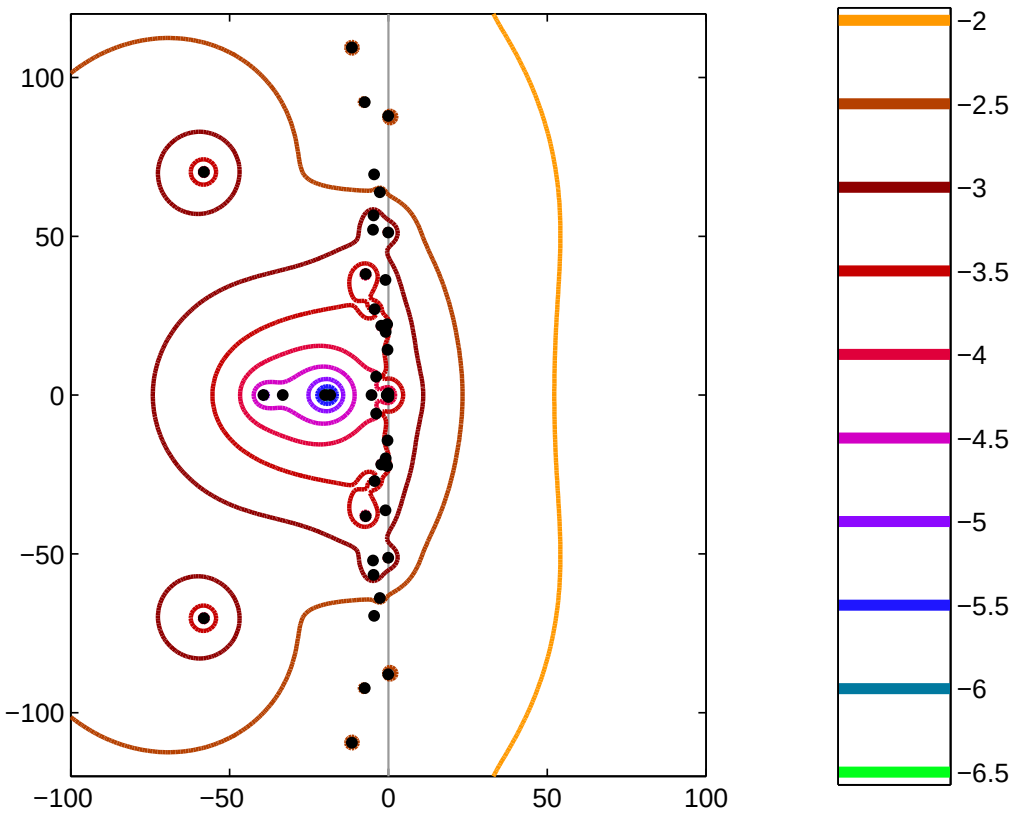
$\sigma(\mathbf{A})$



$W(\mathbf{A})$

Matching the Moments of a Nonnormal Matrix

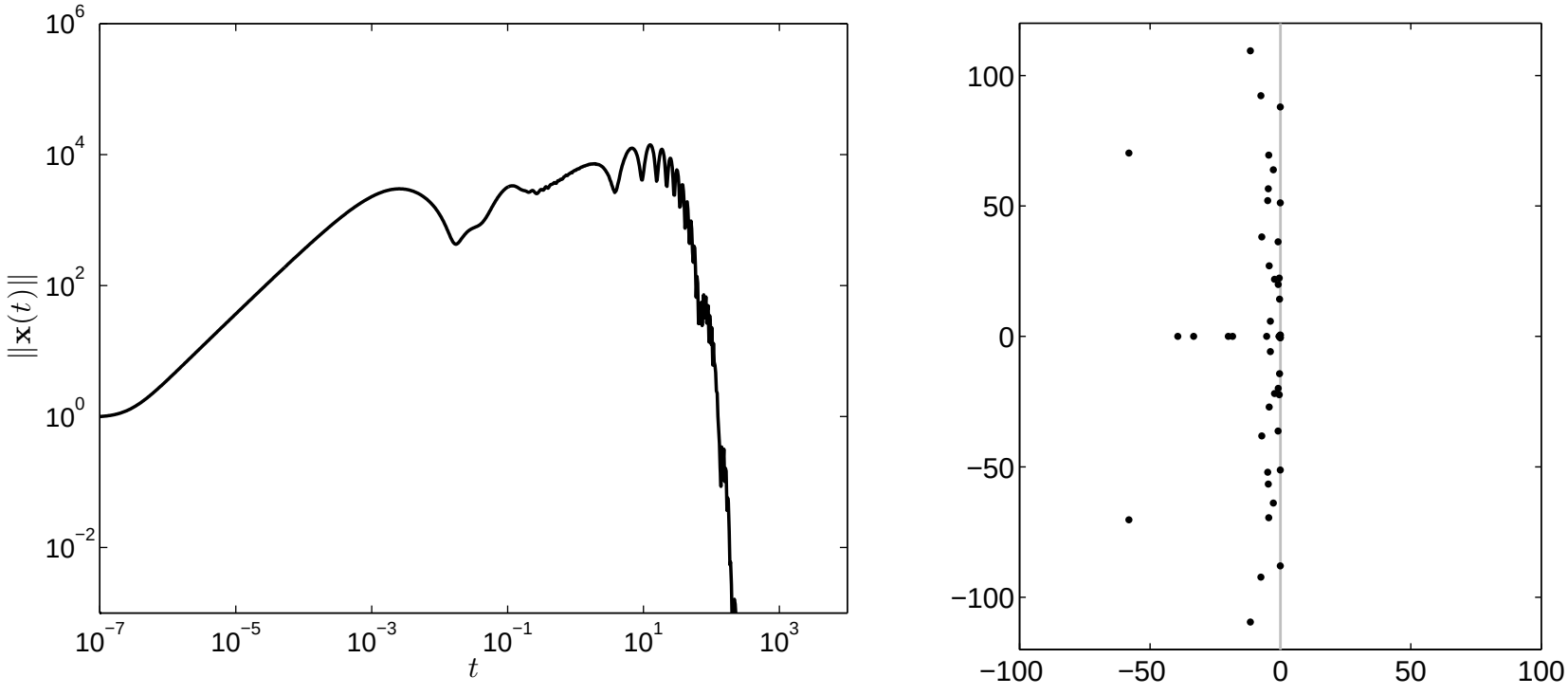
Model of flutter in a Boeing 767 from SLICOT ($n = 55$), stabilized by Burke, Lewis, Overton [2003].



$$\sigma_\varepsilon(\mathbf{A}) = \{z \in \mathbf{C} : \|(z - \mathbf{A})^{-1}\| > 1/\varepsilon\}$$

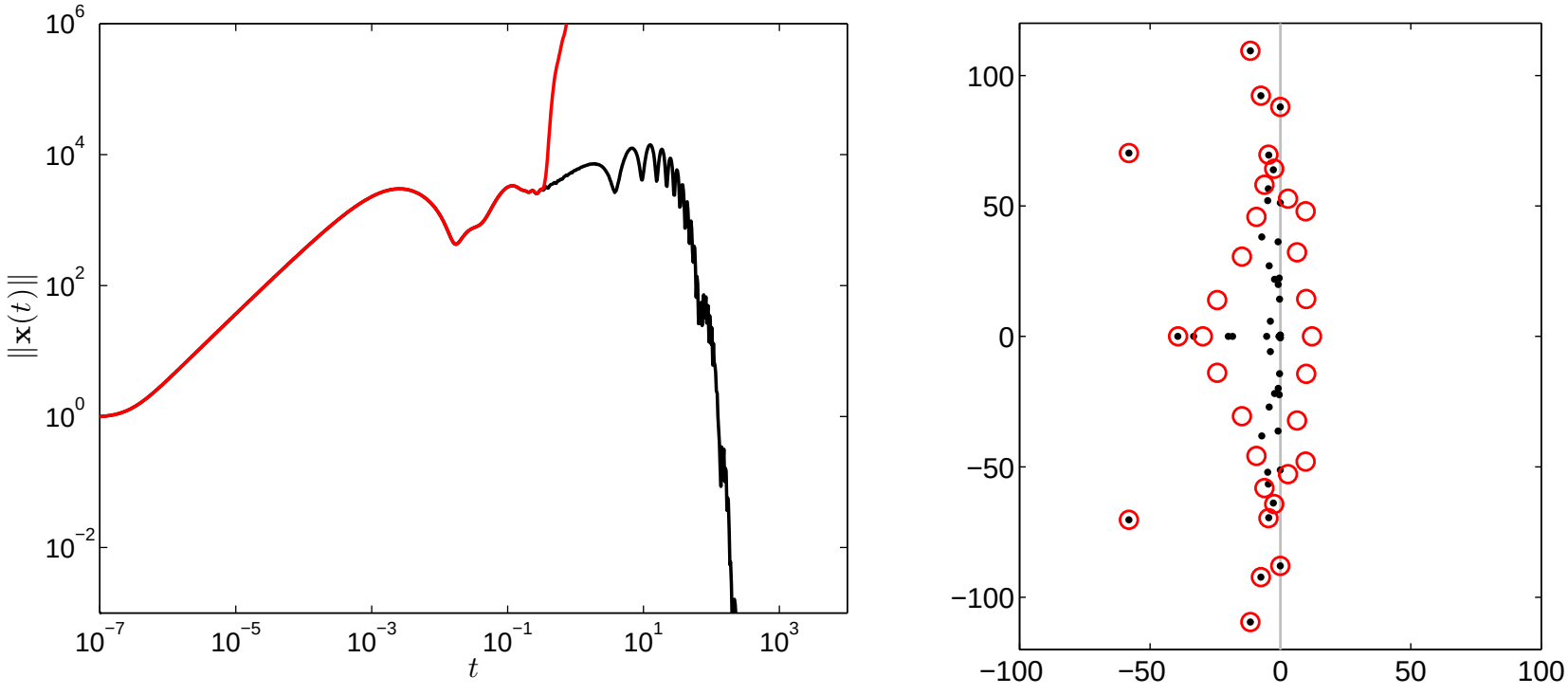
see [Trefethen & E. 2005]

Reduction via Restarted Arnoldi



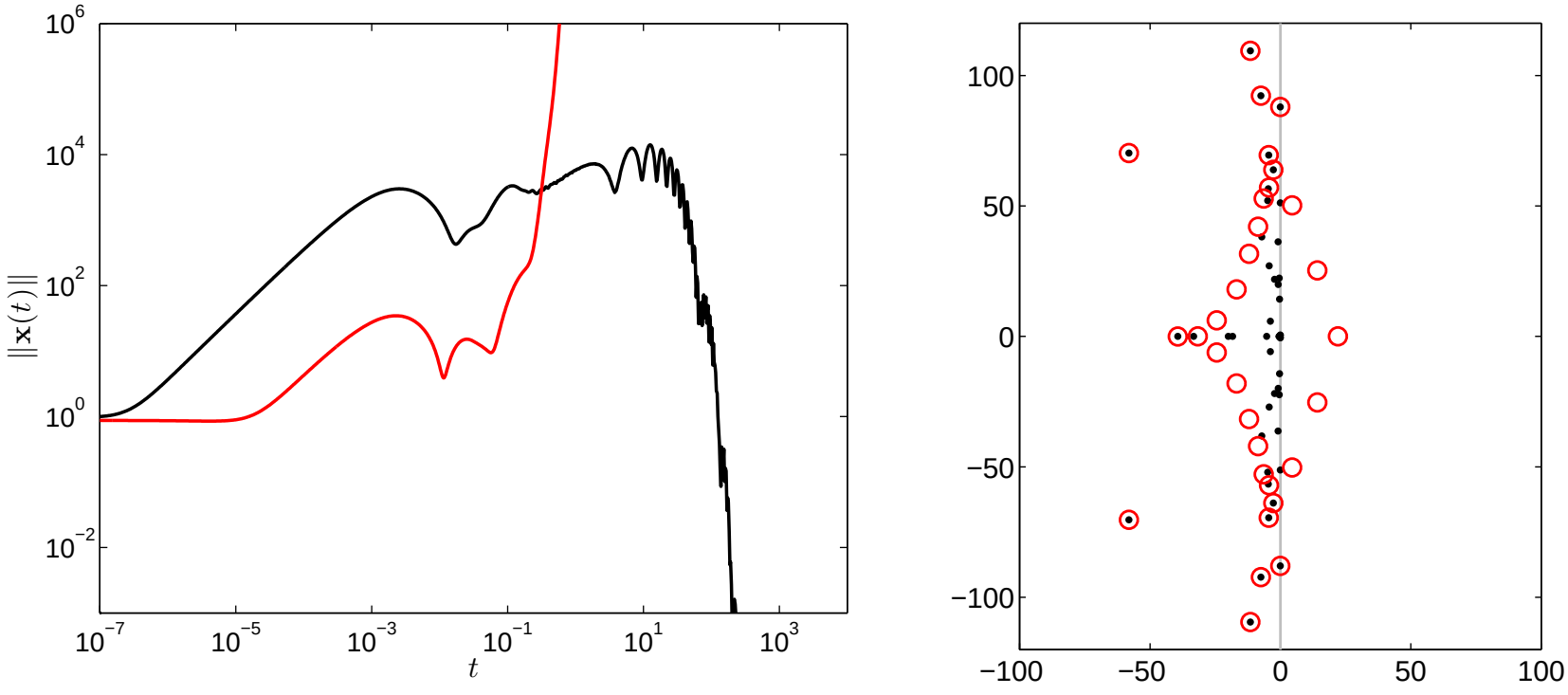
Transient behavior and spectrum of the original system

Reduction via Restarted Arnoldi



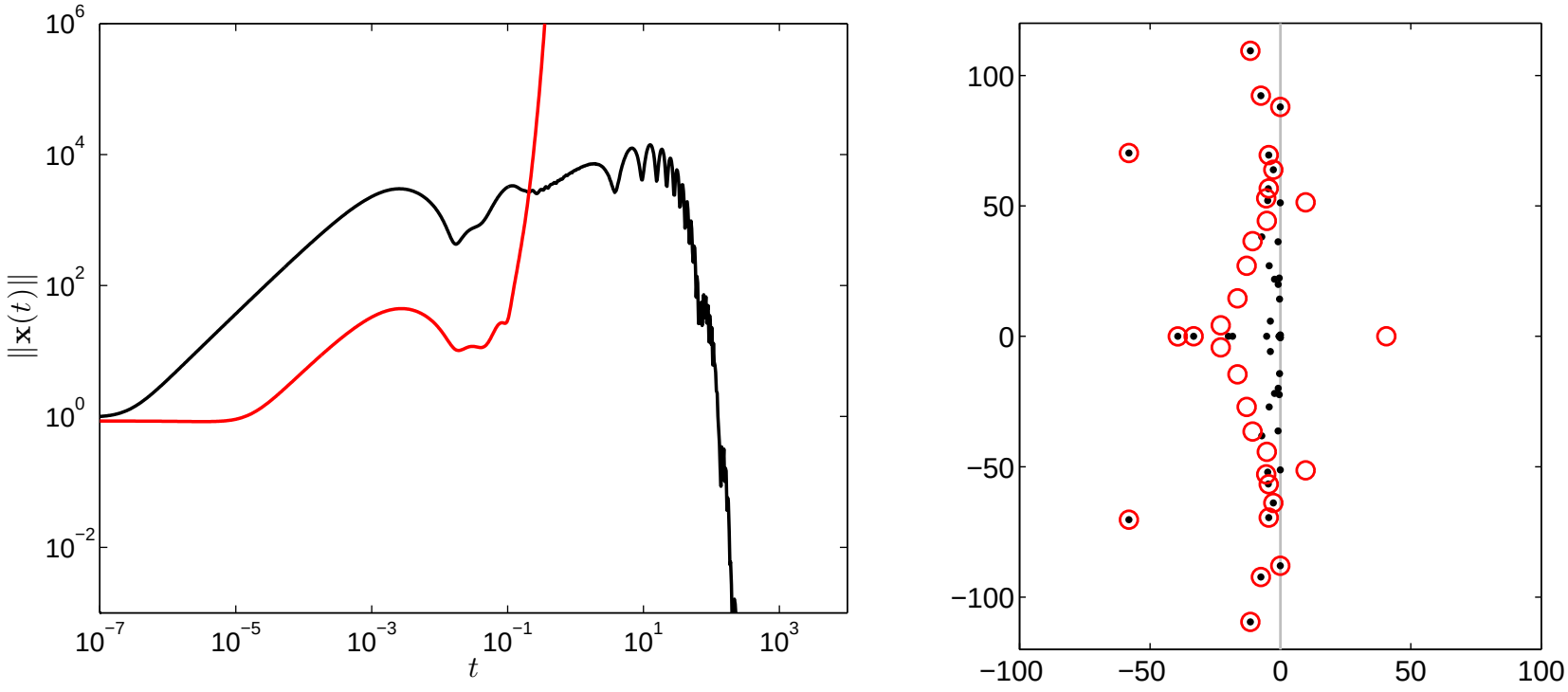
Transient behavior and spectrum for reduced system, $m = 40$

Reduction via Restarted Arnoldi



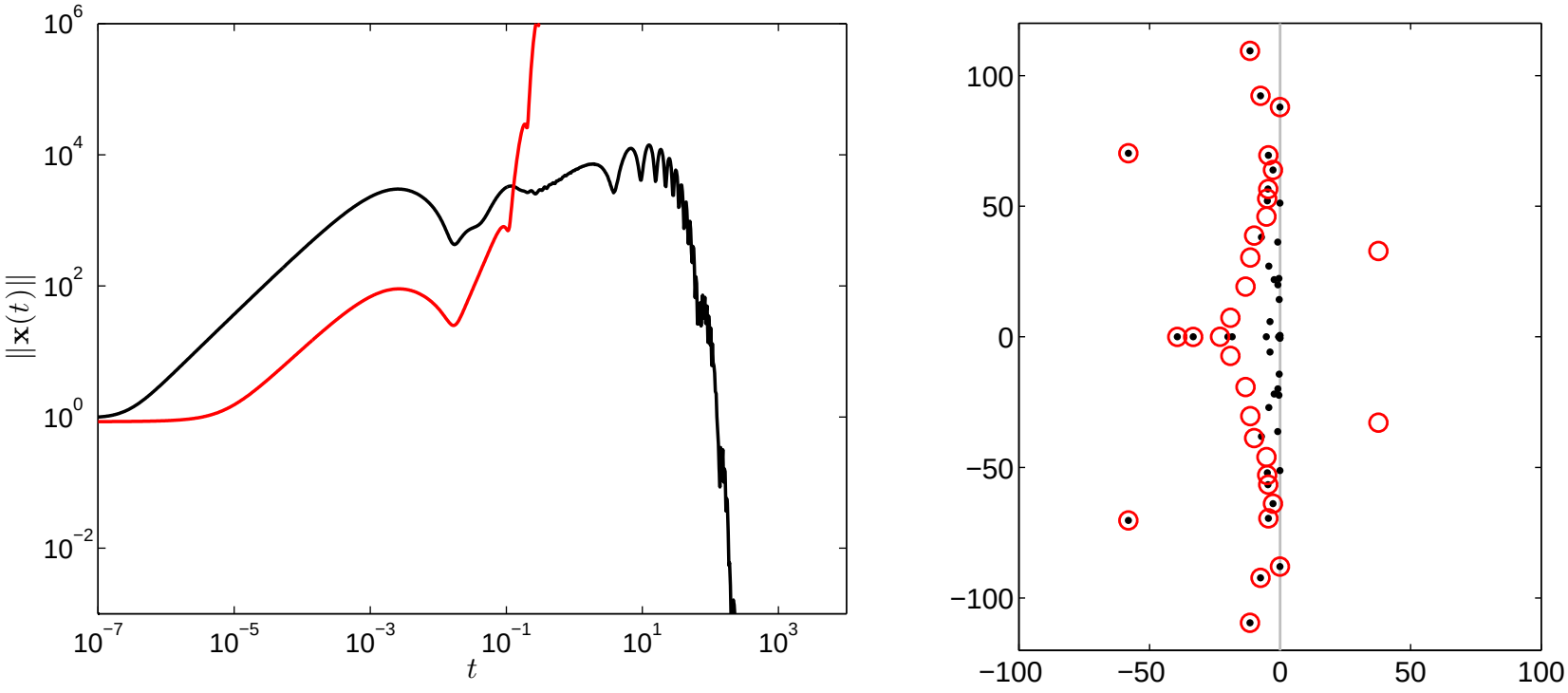
Transient behavior and spectrum for reduced system, $m = 40$,
after one implicit restart

Reduction via Restarted Arnoldi



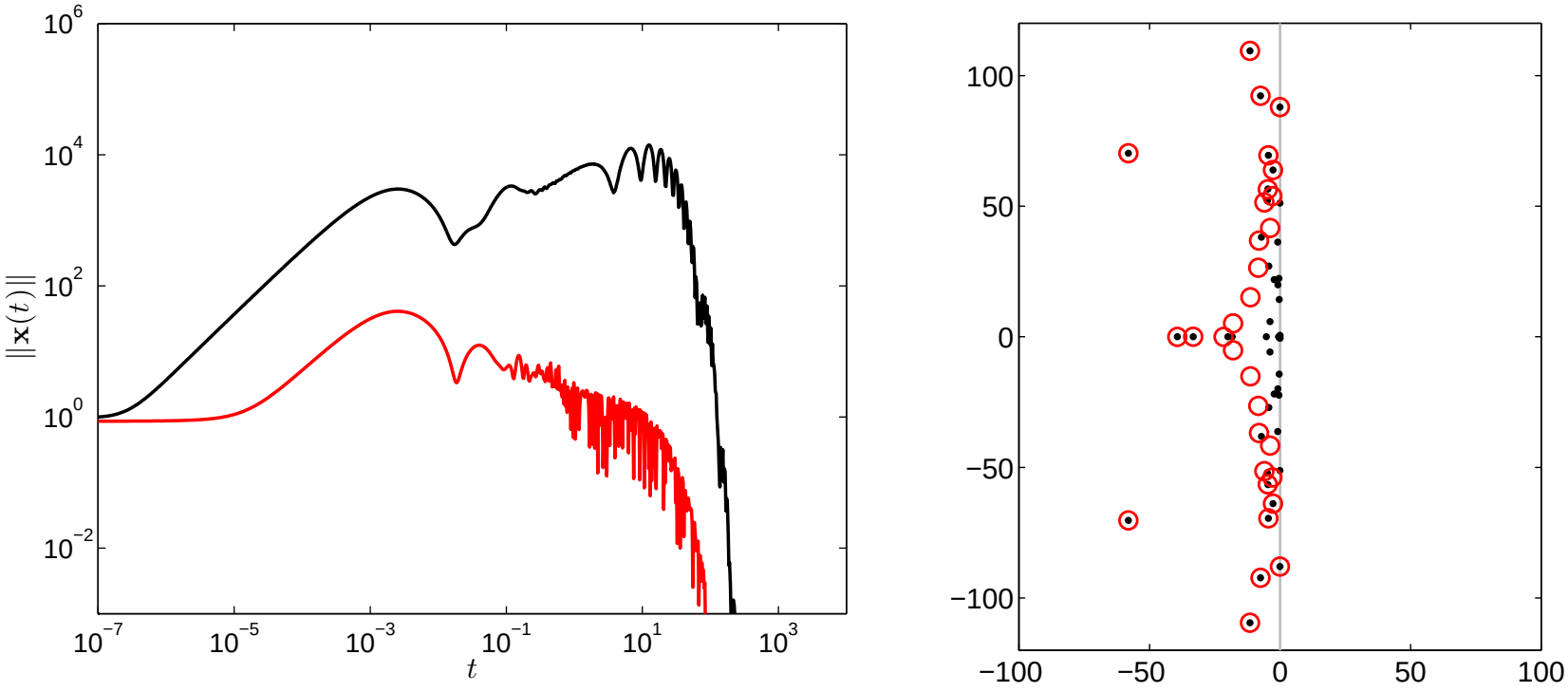
Transient behavior and spectrum for reduced system, $m = 40$,
after two implicit restarts

Reduction via Restarted Arnoldi



Transient behavior and spectrum for reduced system, $m = 40$,
after three implicit restarts

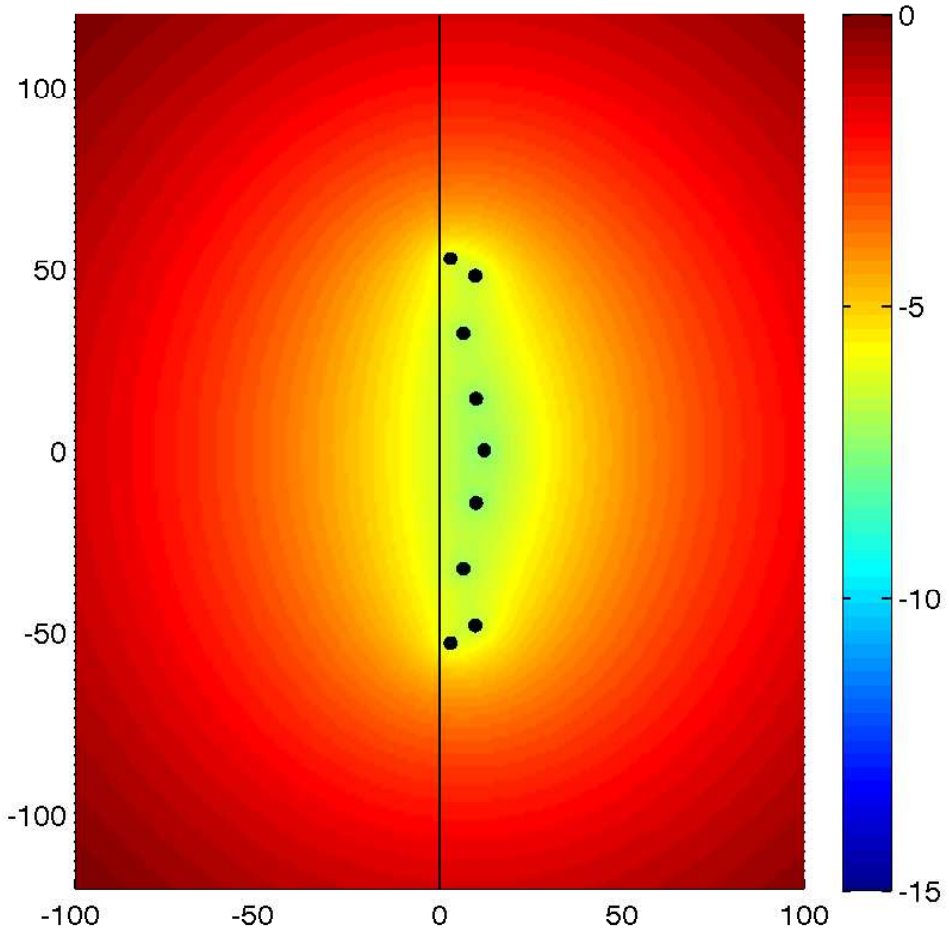
Reduction via Restarted Arnoldi



Transient behavior and spectrum for reduced system, $m = 40$, after four implicit restarts

Reduced system is stable, but underestimates transient growth by several orders of magnitude.

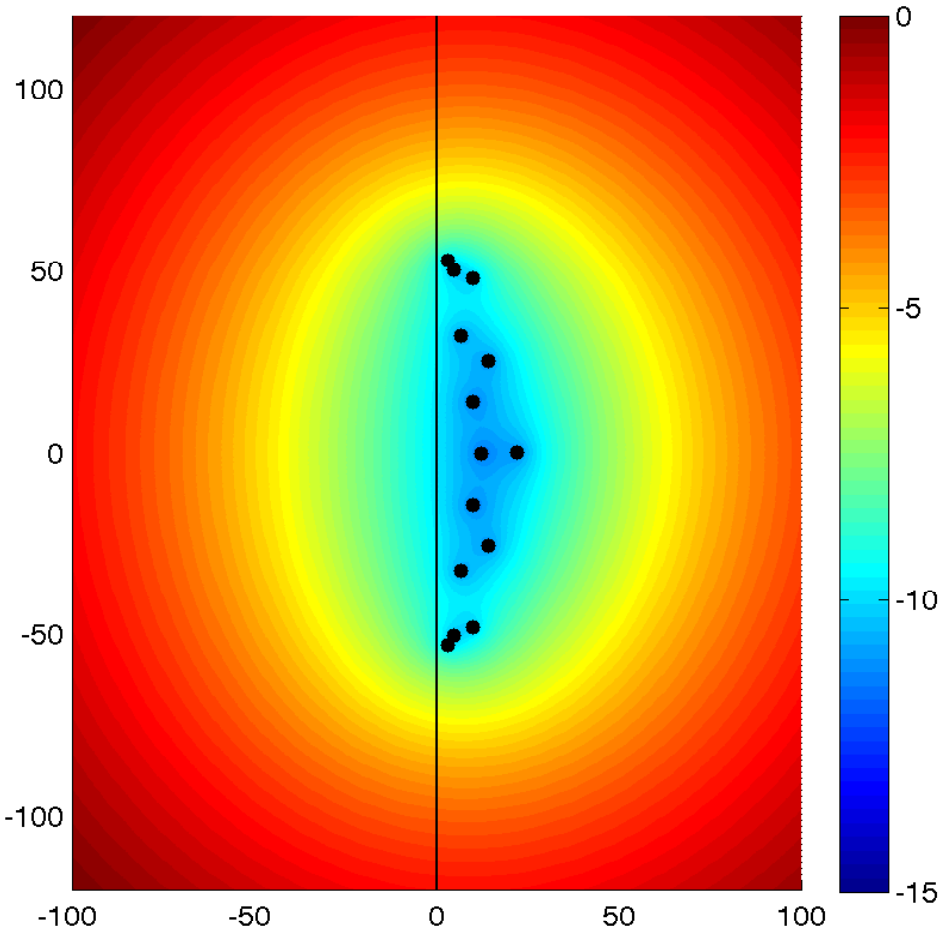
Aggregate Filter Polynomials



Color indicates relative size of $\log_{10} |\psi(z)|$.

first restart

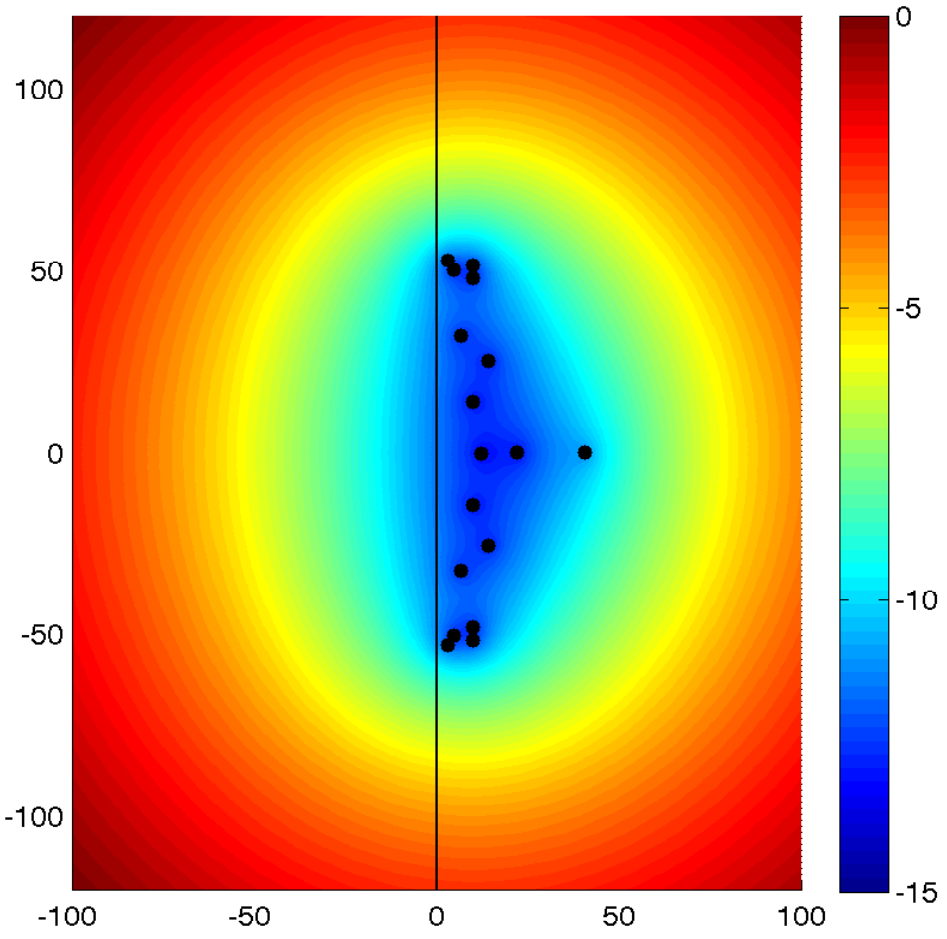
Aggregate Filter Polynomials



Color indicates relative size of $\log_{10} |\psi(z)|$.

second restart

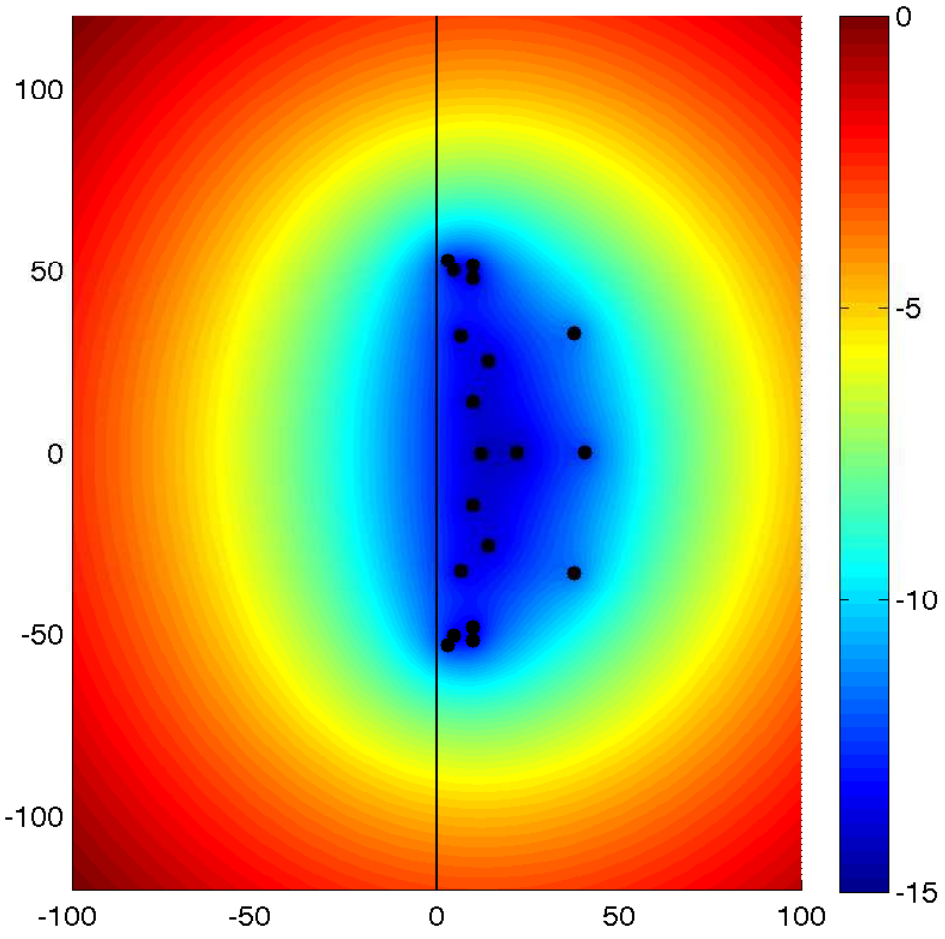
Aggregate Filter Polynomials



Color indicates relative size of $\log_{10} |\psi(z)|$.

third restart

Aggregate Filter Polynomials



Color indicates relative size of $\log_{10} |\psi(z)|$.

fourth restart

Another Example: a Nonlinear Heat Equation

Linear models often arise as linearizations of nonlinear equations.

Consider the nonlinear heat equation on $x \in [-1, 1]$ with $u(-1, t) = u(1, t) = 0$

$$u_t(x, t) = \nu u_{xx}(x, t)$$

with $\nu > 0$

Another Example: a Nonlinear Heat Equation

Linear models often arise as linearizations of nonlinear equations.

Consider the nonlinear heat equation on $x \in [-1, 1]$ with $u(-1, t) = u(1, t) = 0$

$$u_t(x, t) = \nu u_{xx}(x, t) + \sqrt{\nu} u_x(x, t)$$

with $\nu > 0$

Another Example: a Nonlinear Heat Equation

Linear models often arise as linearizations of nonlinear equations.

Consider the nonlinear heat equation on $x \in [-1, 1]$ with $u(-1, t) = u(1, t) = 0$

$$u_t(x, t) = \nu u_{xx}(x, t) + \sqrt{\nu} u_x(x, t) + \tfrac{1}{8} u(x, t)$$

with $\nu > 0$

Another Example: a Nonlinear Heat Equation

Linear models often arise as linearizations of nonlinear equations.

Consider the nonlinear heat equation on $x \in [-1, 1]$ with $u(-1, t) = u(1, t) = 0$

$$u_t(x, t) = \nu u_{xx}(x, t) + \sqrt{\nu} u_x(x, t) + \tfrac{1}{8} u(x, t) + u(x, t)^p$$

with $\nu > 0$ and $p > 1$ [Demanet, Holmer, Zworski].

Another Example: a Nonlinear Heat Equation

Linear models often arise as linearizations of nonlinear equations.

Consider the nonlinear heat equation on $x \in [-1, 1]$ with $u(-1, t) = u(1, t) = 0$

$$u_t(x, t) = \nu u_{xx}(x, t) + \sqrt{\nu} u_x(x, t) + \tfrac{1}{8} u(x, t) + u(x, t)^p$$

with $\nu > 0$ and $p > 1$ [Demanet, Holmer, Zworski].

The linearization L, an advection–diffusion operator,

$$Lu = \nu u_{xx} + \sqrt{\nu} u_x + \tfrac{1}{8} u$$

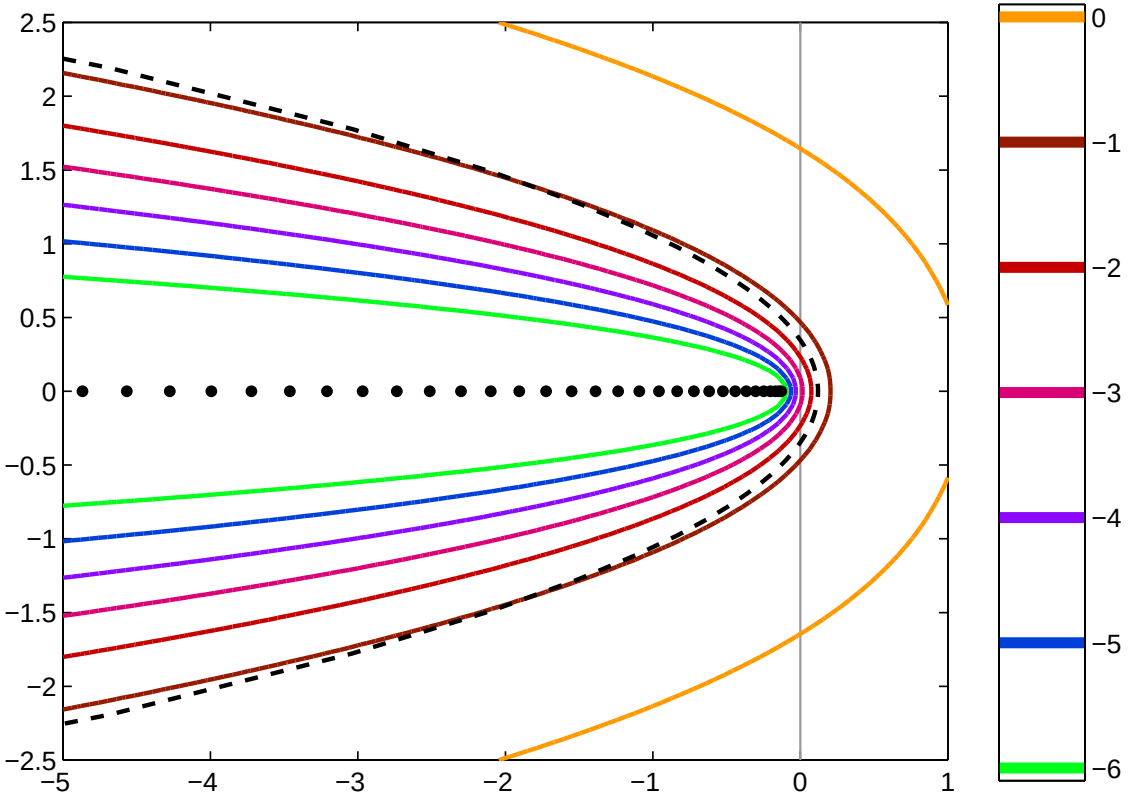
has eigenvalues and eigenfunctions

$$\lambda_n = -\frac{1}{8} - \frac{n^2 \pi^2 \nu}{4}, \qquad u_n(x) = e^{-x/(2\sqrt{\nu})} \sin(n\pi x/2);$$

see, e.g., [Reddy & Trefethen 1994].

The linearized operator is stable for all $\nu > 0$, but has interesting transients . . . .

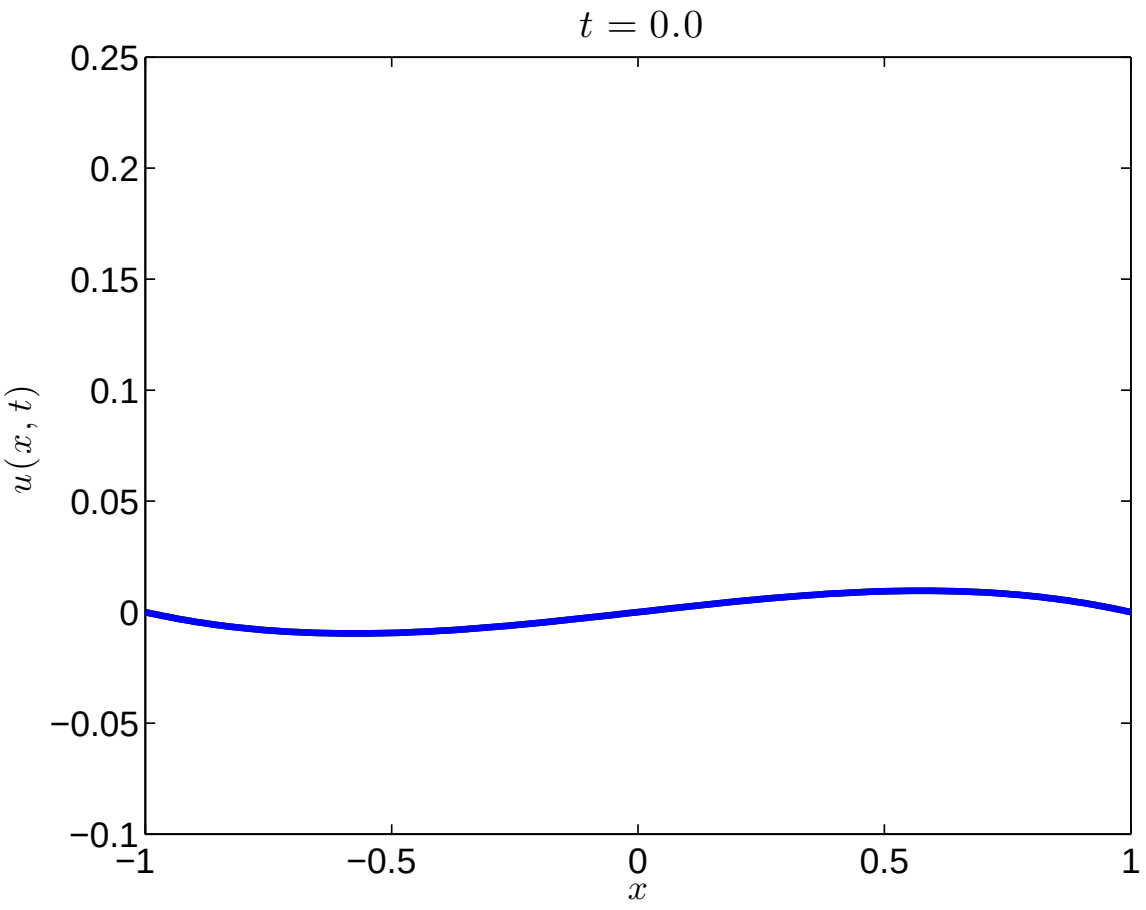
Nonnormality in the Linearization



Spectrum, pseudospectra, and numerical range (L^2 norm, $\nu = 0.02$)

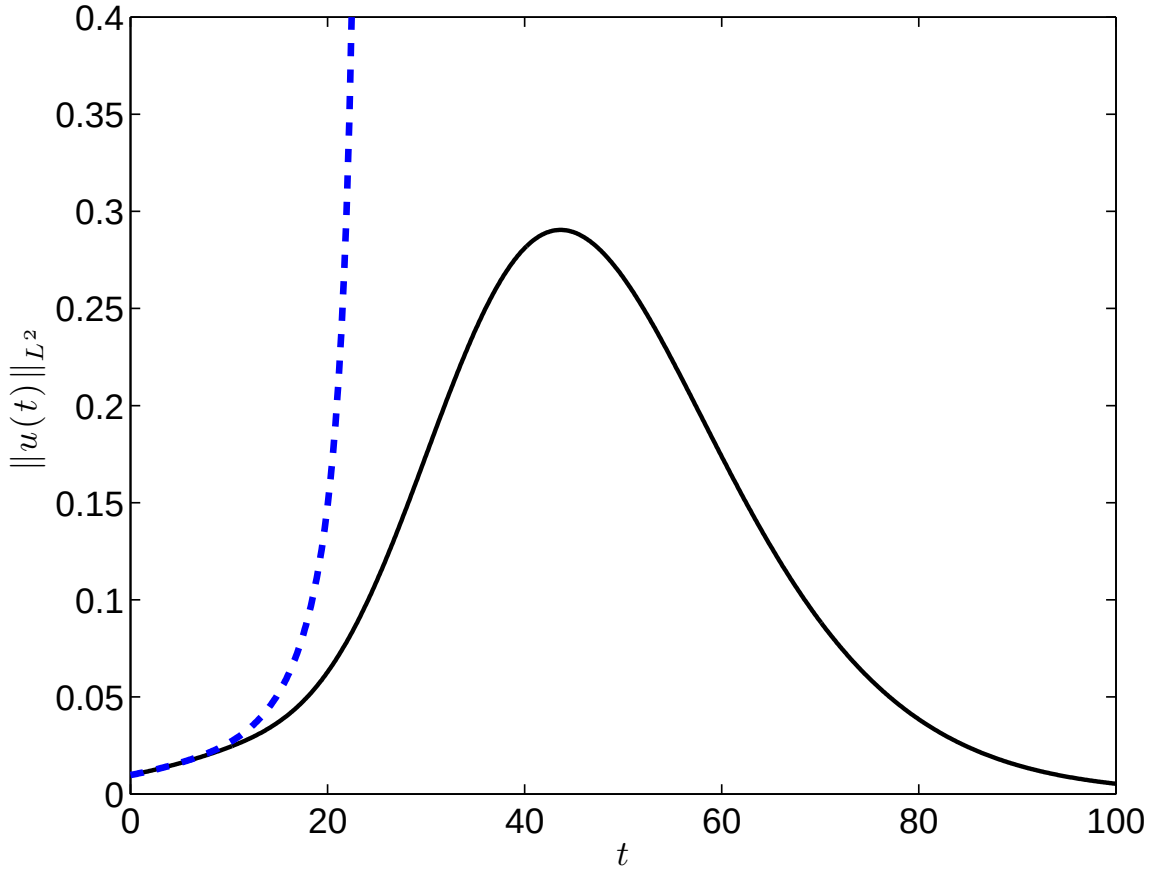
Transient growth can feed the nonlinearity
cf. [Trefethen, Trefethen, Reddy, Driscoll 1993], ...

Evolution of a Small Initial Condition



Nonlinear model (blue) and linearization (black)

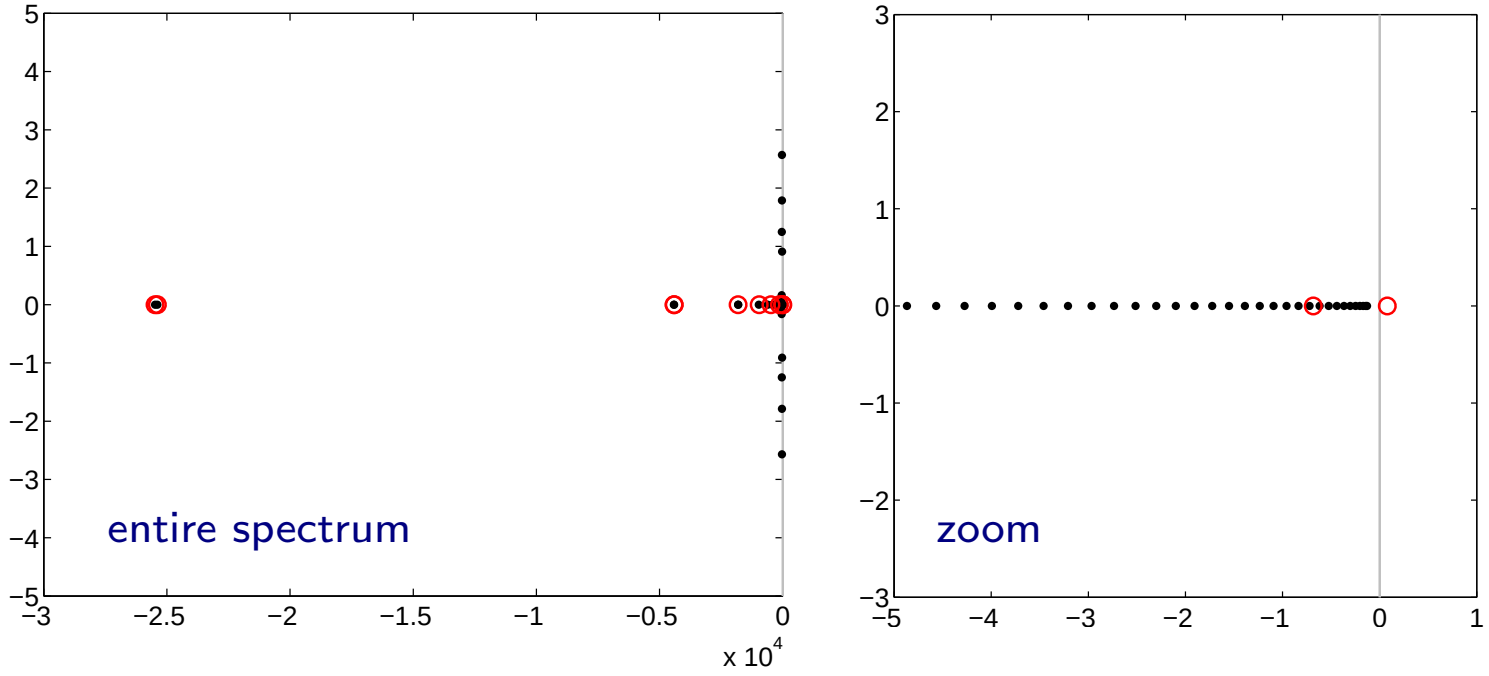
Transient Behavior



Linearized system (black) and nonlinear system (dashed blue)

Nonnormal growth feeds the nonlinear instability.

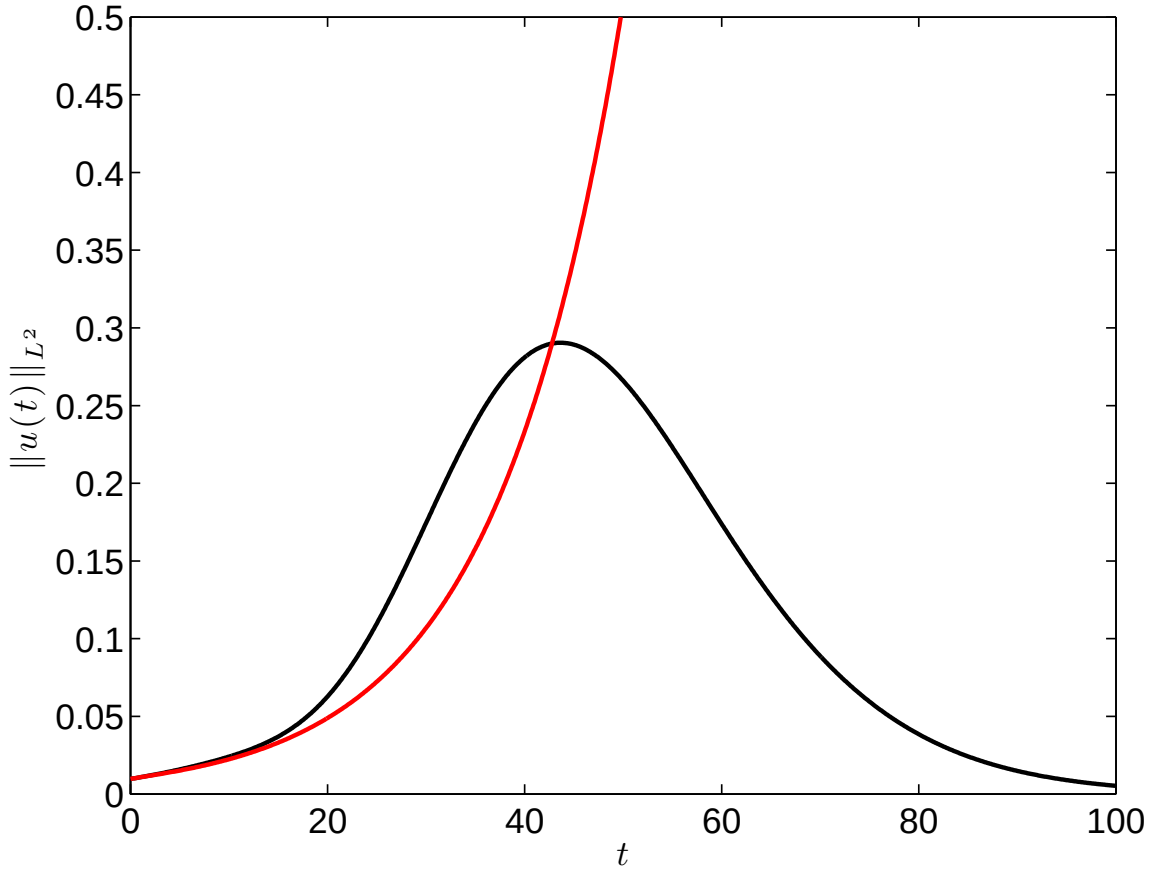
Transient Behavior: Reduction of the Linearized Model



Spectral discretization, $n = 128$ (black) and Arnoldi reduction, $m = 10$ (red).

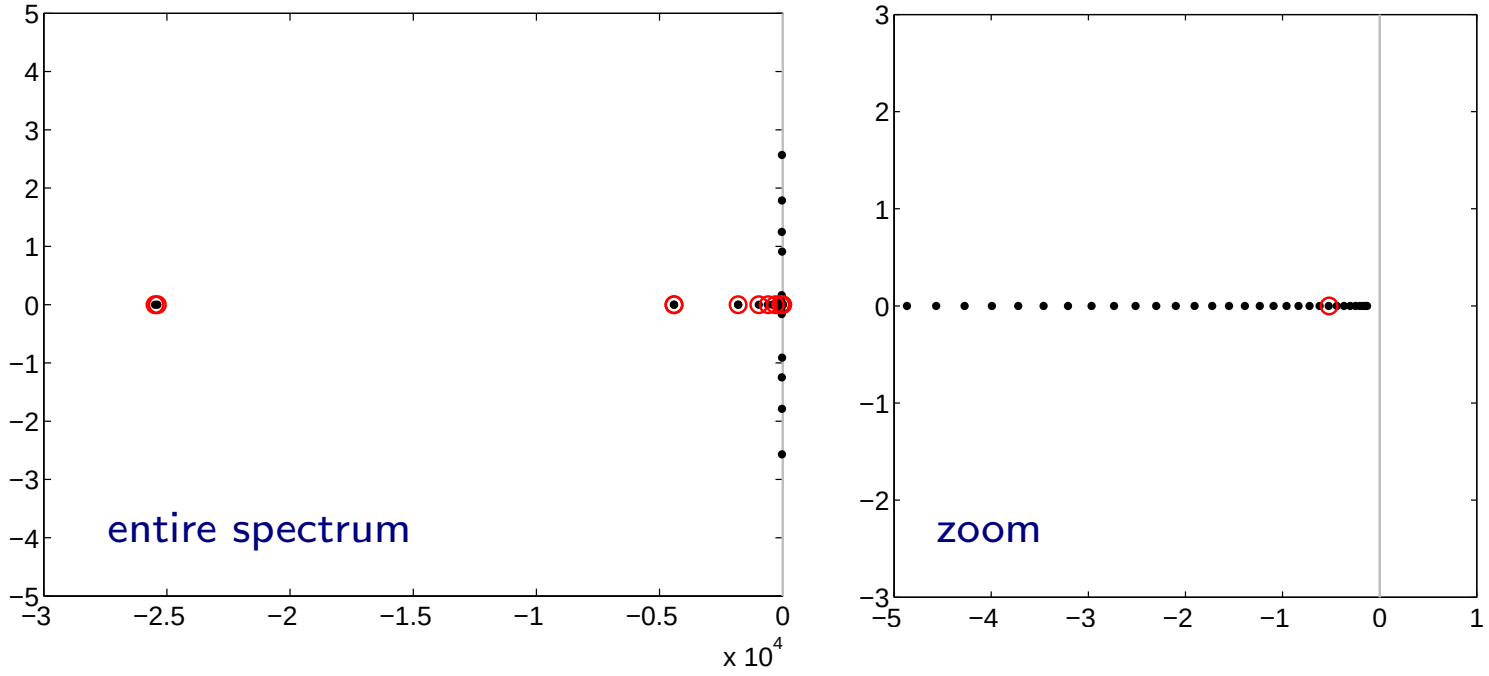
[Many Ritz values capture *spurious* eigenvalues of the discretization of the left.]

Transient Behavior: Reduction of the Linearized Model



Spectral discretization, $n = 128$ (black) and Arnoldi reduction, $m = 10$ (red).

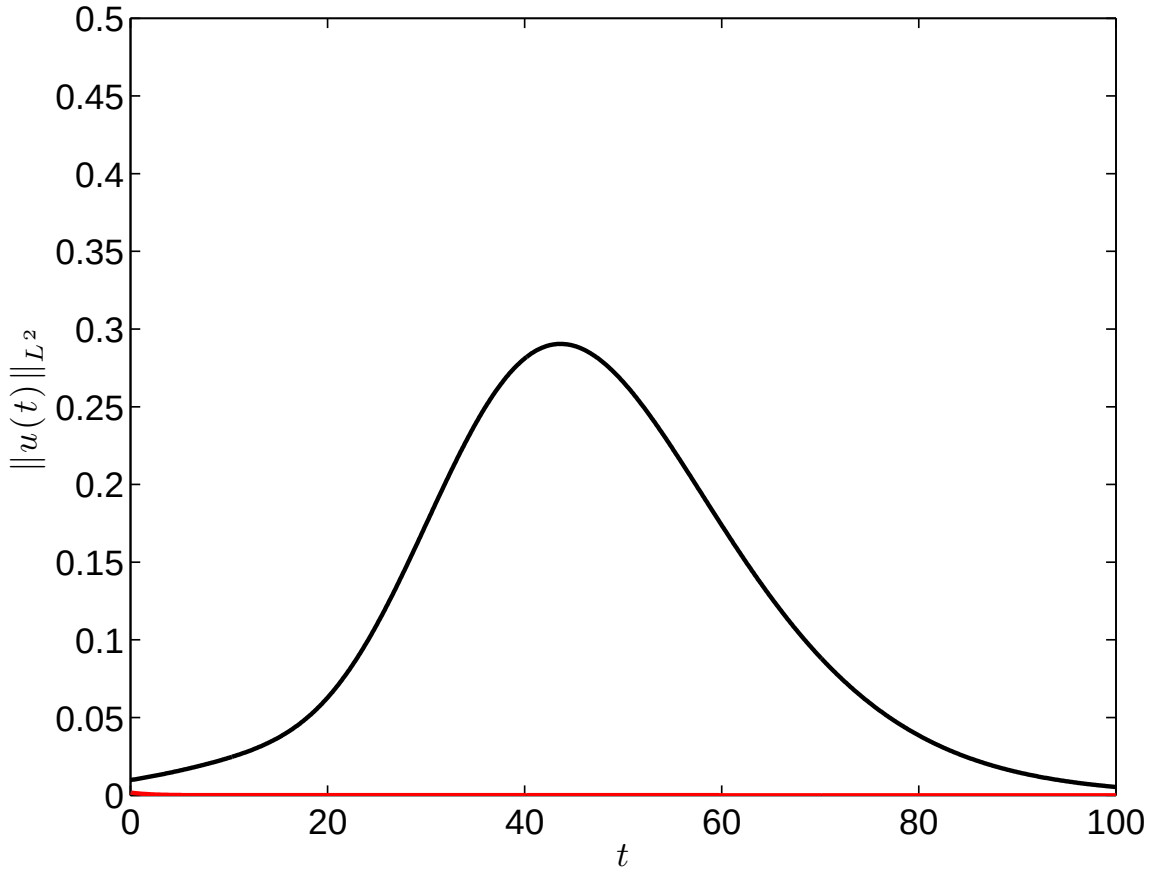
Transient Behavior: Reduction of the Linearized Model



Spectral discretization, $n = 128$ (black) and Arnoldi reduction, $m = 10$ (red) after a restart to remove the spurious eigenvalue.

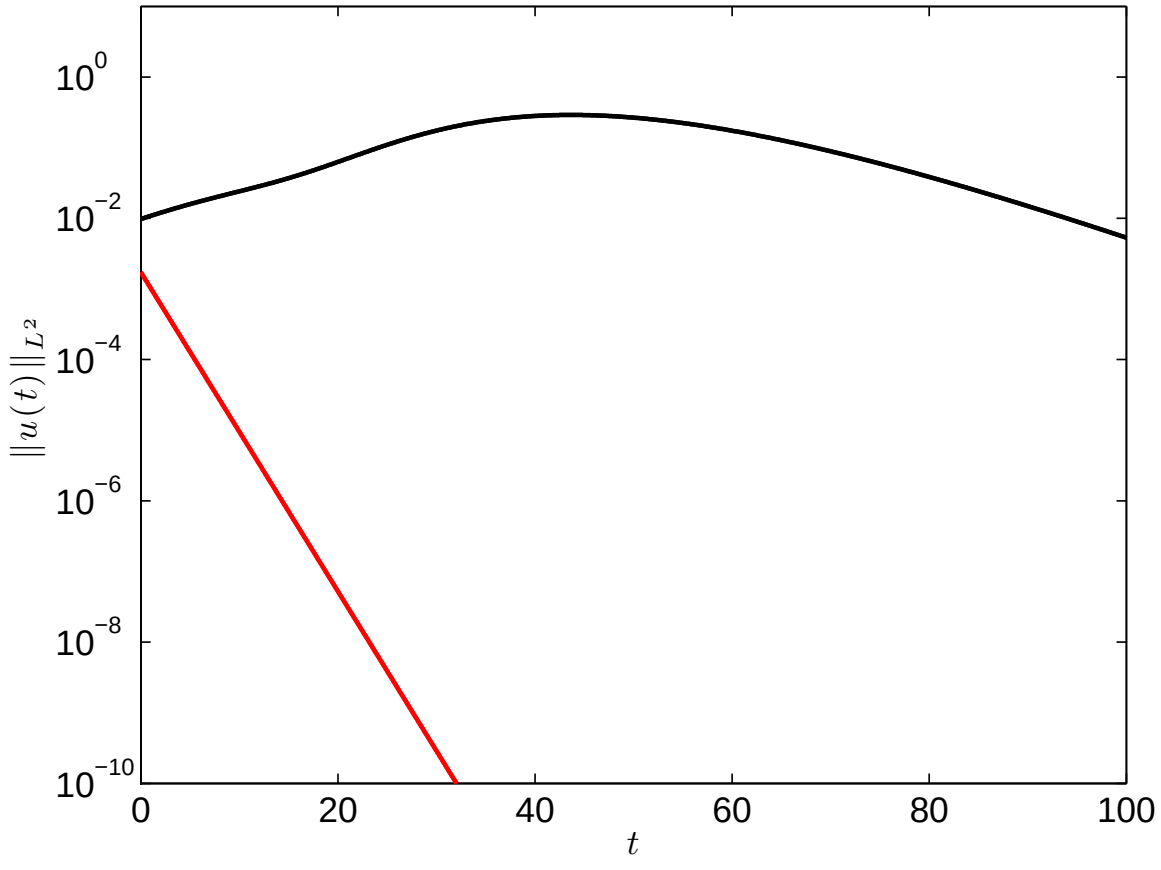
[This effectively pushes Ritz values to the left.]

Transient Behavior: Reduction of the Linearized Model



Spectral discretization, $n = 128$ (black) and Arnoldi reduction, $m = 10$ (red) after one restart to remove the spurious eigenvalue.

Transient Behavior: Reduction of the Linearized Model



Spectral discretization, $n = 128$ (black) and Arnoldi reduction, $m = 10$ (red) after one restart to remove the spurious eigenvalue.

6. GEPs and DAEs

Generalized Eigenvalue Problems

Problem

How should one adapt the definition of the ε-pseudospectrum to the generalized eigenvalue problem

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{B}\mathbf{x} \ ?$$

Generalized Eigenvalue Problems

Problem

How should one adapt the definition of the ε-pseudospectrum to the generalized eigenvalue problem

$$\mathbf{A}\mathbf{x} = \lambda \mathbf{B}\mathbf{x} \ ?$$

Equivalent definitions of $\sigma_\varepsilon(\mathbf{A})$ lead to different meanings for $\sigma_\varepsilon(\mathbf{A}, \mathbf{B})$.

- Approach 1: eigenvalues of perturbations
  $\sigma_\varepsilon(\mathbf{A}) = \{z \in \mathbf{C} : z \in \sigma(\mathbf{A} + \mathbf{E}) \text{ for some } \mathbf{E} \in \mathbf{C}^{n \times n} \text{ with } \|\mathbf{E}\| < \varepsilon\}$
- Approach 2: matrix behavior
  $\sigma_\varepsilon(\mathbf{A}) = \{z \in \mathbf{C} : \|(z - \mathbf{A})^{-1}\| > 1/\varepsilon\}$

GEPs: Eigenvalue Perturbation Approach

Approach 1: eigenvalues of perturbations

- $\sigma_\varepsilon(\mathbf{A}) = \{z \in \mathbf{C} : z \in \sigma(\mathbf{A}+\mathbf{E}) \text{ for some } \mathbf{E} \in \mathbf{C}^{n\times n} \text{ with } \|\mathbf{E}\| < \varepsilon\}$

  Frayssé, Gueury, Nicoud, Toumazou [1996] proposed:

$$\begin{aligned}\sigma_\varepsilon(\mathbf{A},\mathbf{B}) \;=\; \{z \in \mathbf{C} : (\mathbf{A}+\mathbf{E}_0)\mathbf{x} = z(\mathbf{B}+\mathbf{E}_1)\mathbf{x} \text{ for some} \\ \mathbf{x} \neq \mathbf{0} \text{ and } \mathbf{E}_0,\, \mathbf{E}_1 \text{ with } \|\mathbf{E}_0\| < \varepsilon\alpha_0,\, \|\mathbf{E}_1\| < \varepsilon\alpha_1\},\end{aligned}$$

  where, e.g., either $\alpha_0 = \alpha_1 = 1$, or $\alpha_0 = \|\mathbf{A}\|$ and $\alpha_1 = \|\mathbf{B}\|$.

GEPs: Eigenvalue Perturbation Approach

Approach 1: eigenvalues of perturbations

- $\sigma_\varepsilon(\mathbf{A}) = \{z \in \mathbf{C} : z \in \sigma(\mathbf{A}+\mathbf{E}) \text{ for some } \mathbf{E} \in \mathbf{C}^{n\times n} \text{ with } \|\mathbf{E}\| < \varepsilon\}$

  Frayssé, Gueury, Nicoud, Toumazou [1996] proposed:

$$\begin{aligned}\sigma_\varepsilon(\mathbf{A},\mathbf{B}) \;=\; \{z \in \mathbf{C} : (\mathbf{A}+\mathbf{E}_0)\mathbf{x} = z(\mathbf{B}+\mathbf{E}_1)\mathbf{x} \text{ for some} \\ \mathbf{x} \neq \mathbf{0} \text{ and } \mathbf{E}_0,\, \mathbf{E}_1 \text{ with } \|\mathbf{E}_0\| < \varepsilon\alpha_0,\ \|\mathbf{E}_1\| < \varepsilon\alpha_1\},\end{aligned}$$

  where, e.g., either $\alpha_0 = \alpha_1 = 1$, or $\alpha_0 = \|\mathbf{A}\|$ and $\alpha_1 = \|\mathbf{B}\|$.

- This has an equivalent resolvent-like formulation:

$$\sigma_\varepsilon(\mathbf{A},\mathbf{B}) = \{z \in \mathbf{C} : \|(\mathbf{B}z-\mathbf{A})^{-1}\|(\alpha_0 + \alpha_1|z|) > 1/\varepsilon\}.$$

GEPs: Eigenvalue Perturbation Approach

Approach 1: eigenvalues of perturbations

- $\sigma_\varepsilon(\mathbf{A}) = \{z \in \mathbf{C} : z \in \sigma(\mathbf{A}+\mathbf{E}) \text{ for some } \mathbf{E} \in \mathbf{C}^{n\times n} \text{ with } \|\mathbf{E}\| < \varepsilon\}$

  Frayssé, Gueury, Nicoud, Toumazou [1996] proposed:

$$\begin{aligned}\sigma_\varepsilon(\mathbf{A},\mathbf{B}) \;=\; \{z \in \mathbf{C} : (\mathbf{A}+\mathbf{E}_0)\mathbf{x} = z(\mathbf{B}+\mathbf{E}_1)\mathbf{x} \text{ for some} \\ \mathbf{x} \neq \mathbf{0} \text{ and } \mathbf{E}_0,\, \mathbf{E}_1 \text{ with } \|\mathbf{E}_0\| < \varepsilon\alpha_0,\ \|\mathbf{E}_1\| < \varepsilon\alpha_1\},\end{aligned}$$

  where, e.g., either $\alpha_0 = \alpha_1 = 1$, or $\alpha_0 = \|\mathbf{A}\|$ and $\alpha_1 = \|\mathbf{B}\|$.

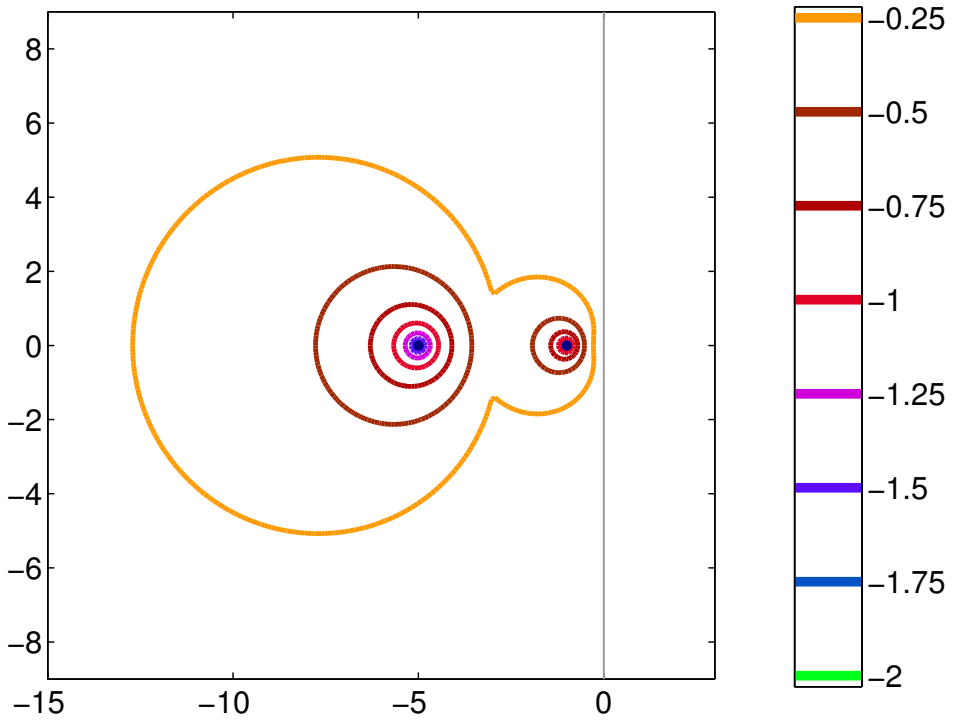
- This has an equivalent resolvent-like formulation:

$$\sigma_\varepsilon(\mathbf{A},\mathbf{B}) = \{z \in \mathbf{C} : \|(\mathbf{B}z - \mathbf{A})^{-1}\|(\alpha_0 + \alpha_1|z|) > 1/\varepsilon\}.$$

- Generalized to matrix polynomials by Tisseur & N. Higham [2001, 2002]; see also Lancaster & Psarrakos [2005].
- Cf. [Boutry, Elad, Golub, Milanfar, 2005] for rectangular pencils.

GEPs: Eigenvalue Perturbation Approach

$$\mathbf{A} = \begin{bmatrix} -1 & -5M \\ 0 & -5 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 1 & M \\ 0 & 1 \end{bmatrix}$$

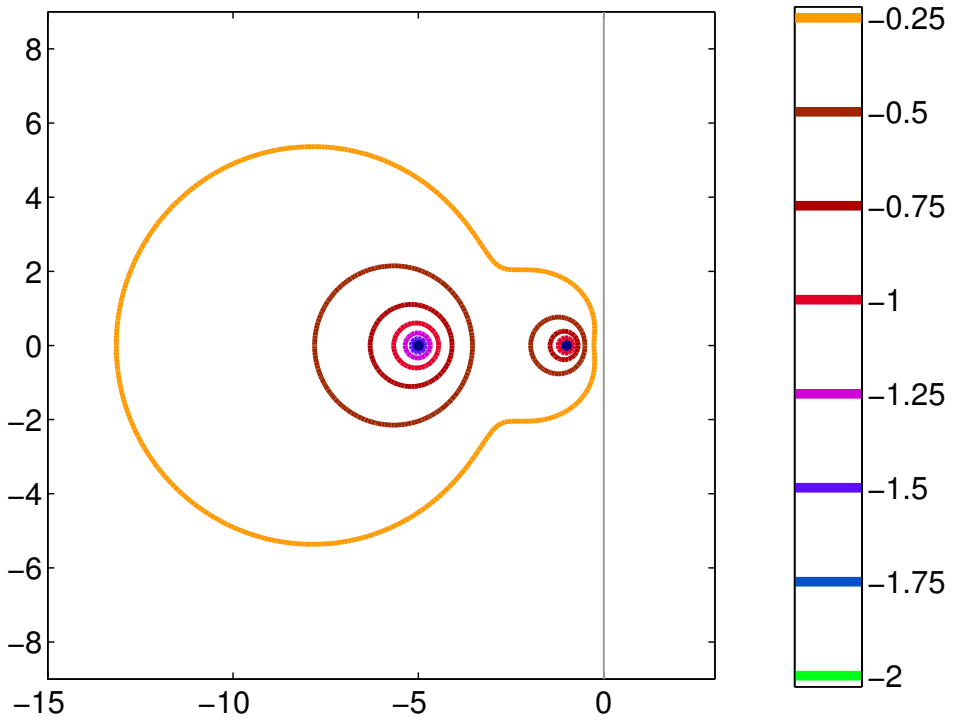


$M = 0$

$$\sigma(A, B) = \{-1, -5\}$$

GEPs: Eigenvalue Perturbation Approach

$$\mathbf{A} = \begin{bmatrix} -1 & -5M \\ 0 & -5 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 1 & M \\ 0 & 1 \end{bmatrix}$$

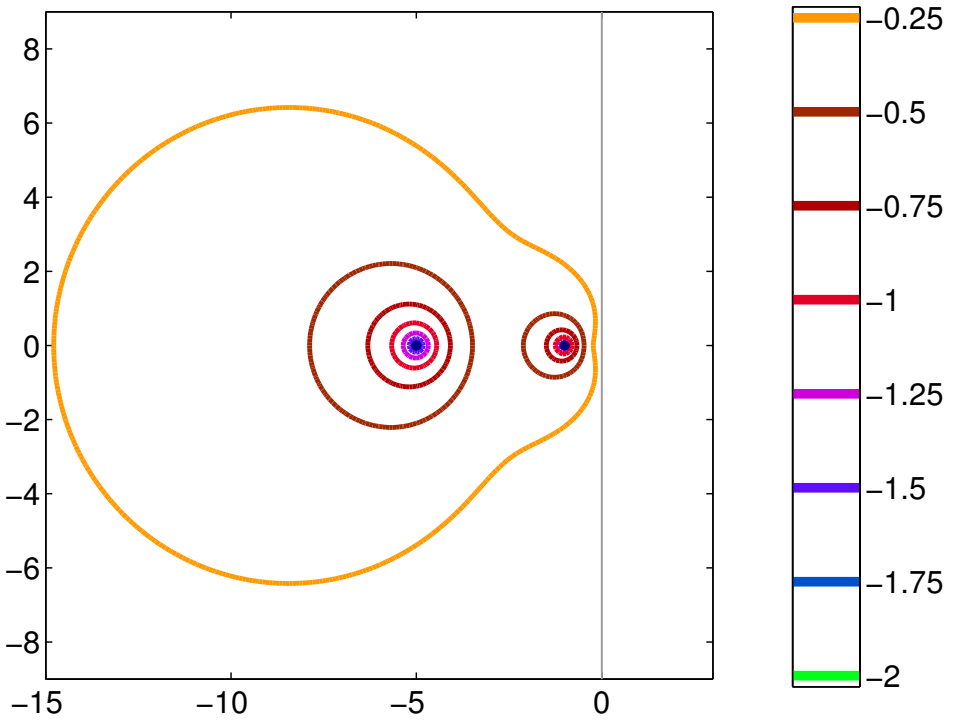


$M = 0.25$

$$\sigma(A, B) = \{-1, -5\}$$

GEPs: Eigenvalue Perturbation Approach

$$\mathbf{A} = \begin{bmatrix} -1 & -5M \\ 0 & -5 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 1 & M \\ 0 & 1 \end{bmatrix}$$

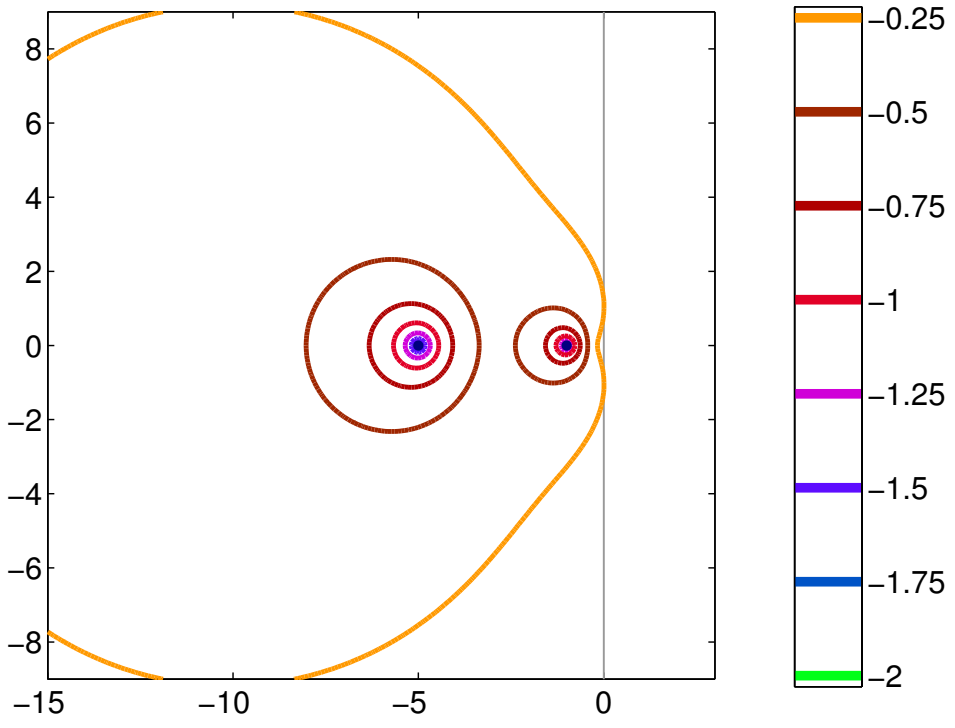


$M = 0.50$

$$\sigma(A, B) = \{-1, -5\}$$

GEPs: Eigenvalue Perturbation Approach

$$\mathbf{A} = \begin{bmatrix} -1 & -5M \\ 0 & -5 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 1 & M \\ 0 & 1 \end{bmatrix}$$

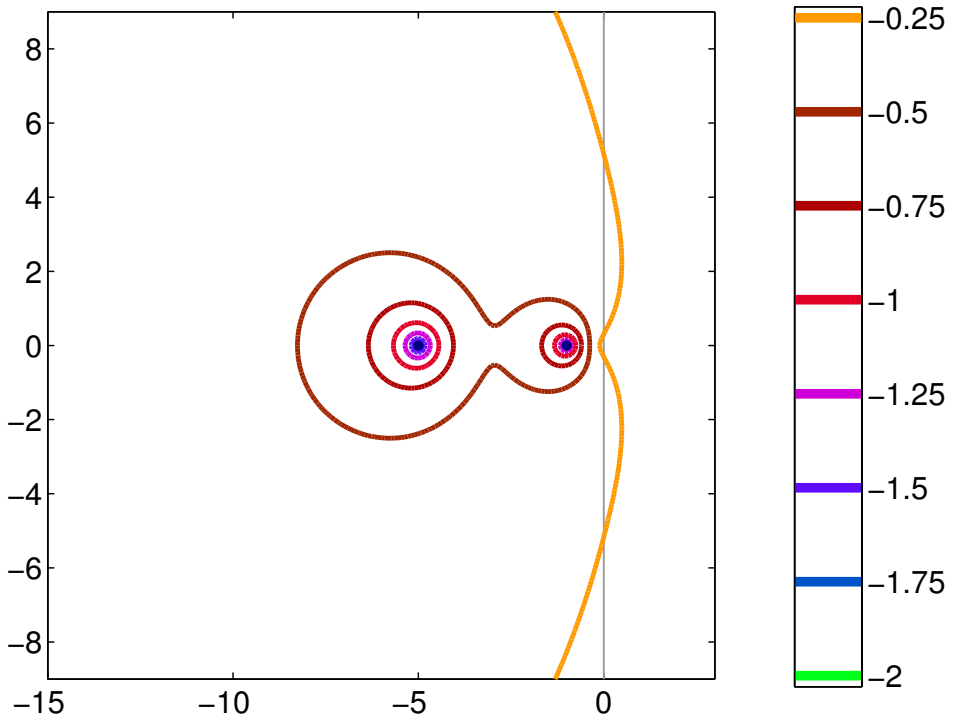


$M = 0.75$

$$\sigma(A, B) = \{-1, -5\}$$

GEPs: Eigenvalue Perturbation Approach

$$\mathbf{A} = \begin{bmatrix} -1 & -5M \\ 0 & -5 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 1 & M \\ 0 & 1 \end{bmatrix}$$

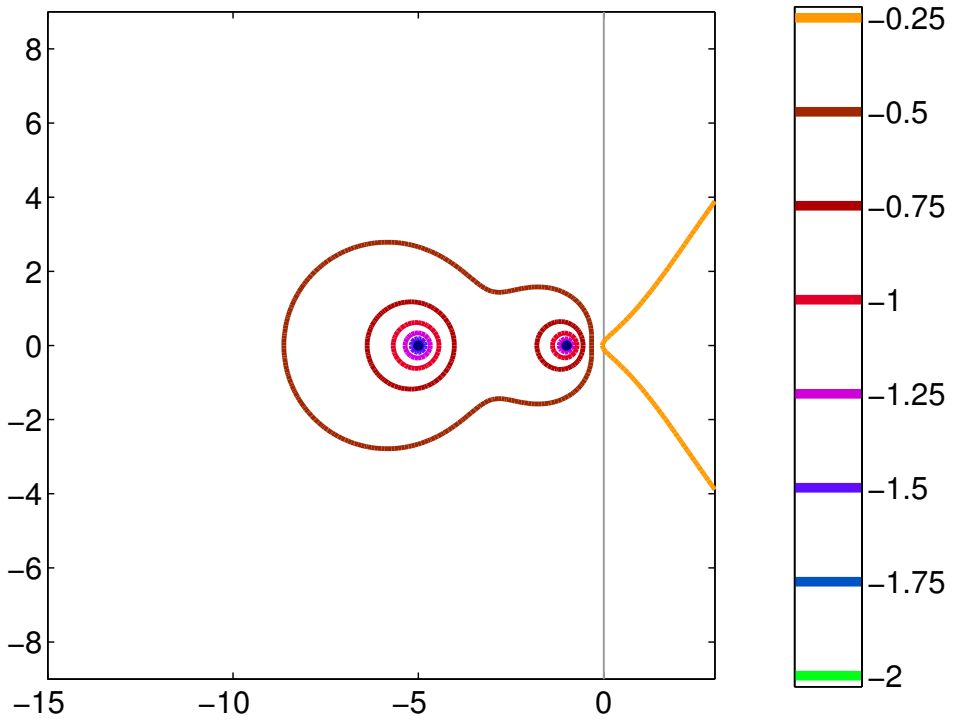


$M = 1.00$

$$\sigma(A, B) = \{-1, -5\}$$

GEPs: Eigenvalue Perturbation Approach

$$\mathbf{A} = \begin{bmatrix} -1 & -5M \\ 0 & -5 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 1 & M \\ 0 & 1 \end{bmatrix}$$

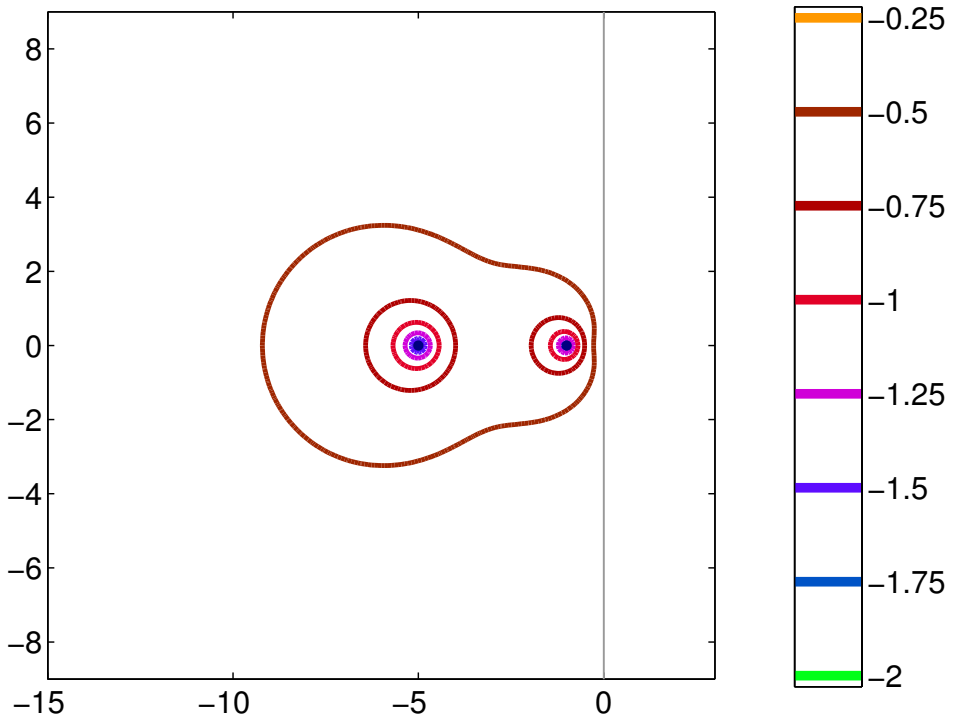


$M = 1.25$

$$\sigma(A, B) = \{-1, -5\}$$

GEPs: Eigenvalue Perturbation Approach

$$\mathbf{A} = \begin{bmatrix} -1 & -5M \\ 0 & -5 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 1 & M \\ 0 & 1 \end{bmatrix}$$

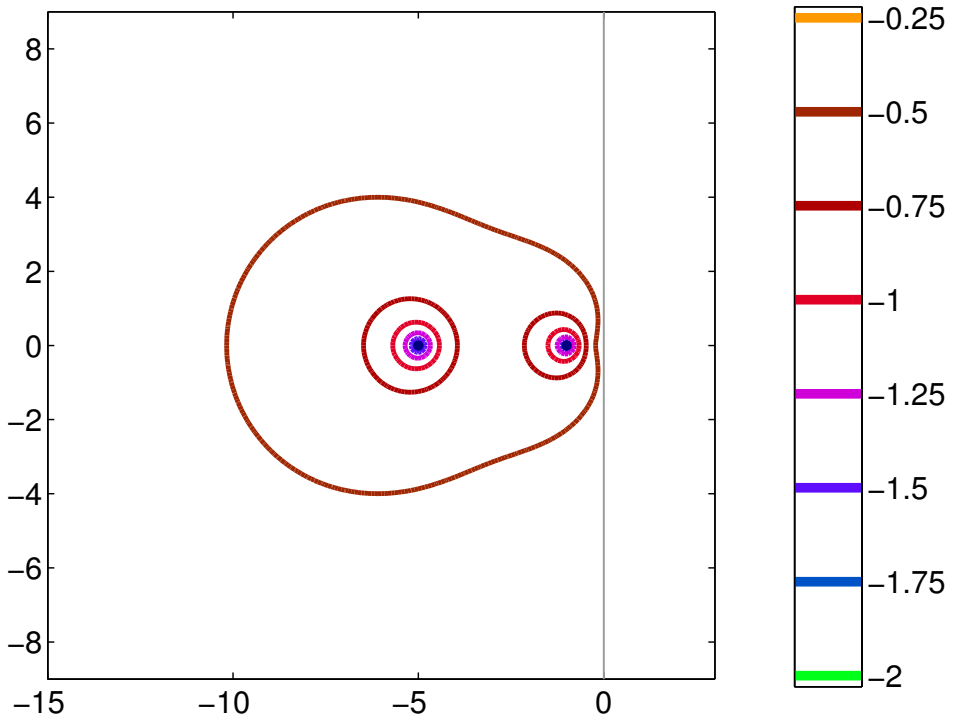


$M = 1.50$

$$\sigma(A, B) = \{-1, -5\}$$

GEPs: Eigenvalue Perturbation Approach

$$\mathbf{A} = \begin{bmatrix} -1 & -5M \\ 0 & -5 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 1 & M \\ 0 & 1 \end{bmatrix}$$

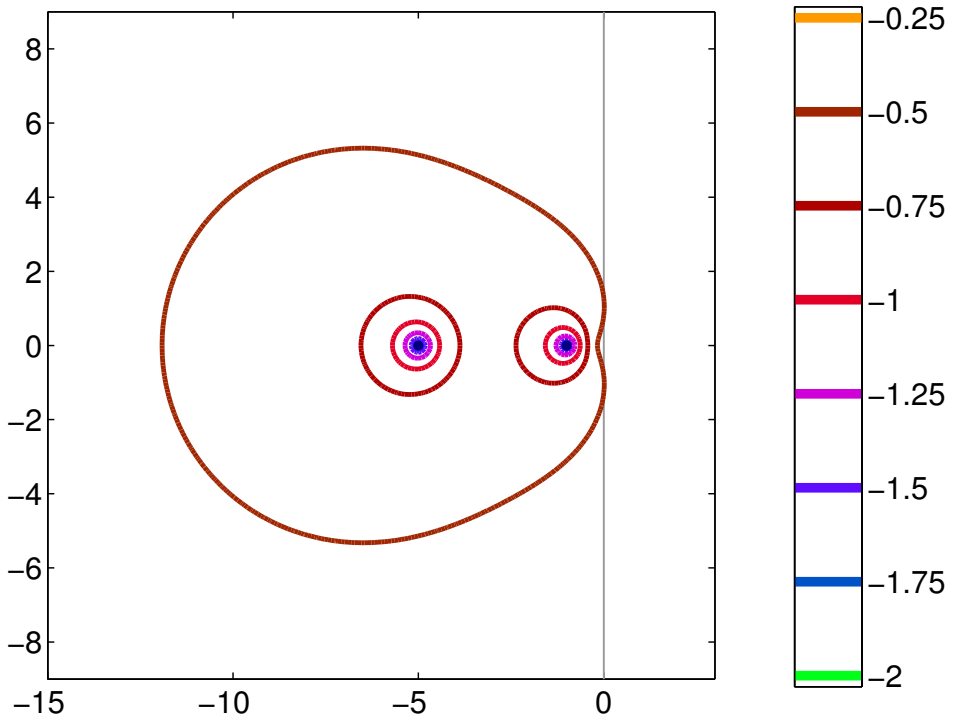


$M = 1.75$

$$\sigma(A, B) = \{-1, -5\}$$

GEPs: Eigenvalue Perturbation Approach

$$\mathbf{A} = \begin{bmatrix} -1 & -5M \\ 0 & -5 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 1 & M \\ 0 & 1 \end{bmatrix}$$



$M = 2.00$

$$\sigma(A, B) = \{-1, -5\}$$

GEPs: Eigenvalue Perturbation Approach

Consider solutions to $\mathbf{B}\mathbf{x}'(t) = \mathbf{A}\mathbf{x}(t)$ for the previous example:

$$\mathbf{A} = \begin{bmatrix} -1 & -5M \\ 0 & -5 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 1 & M \\ 0 & 1 \end{bmatrix}$$

Note that

$$\mathbf{B}^{-1}\mathbf{A} = \begin{bmatrix} -1 & 0 \\ 0 & -5 \end{bmatrix}.$$

- Since $\mathbf{B}^{-1}\mathbf{A}$ is normal and stable, solutions to $\mathbf{B}\mathbf{x}'(t) = \mathbf{A}\mathbf{x}(t)$ cannot exhibit growth.

GEPs: Eigenvalue Perturbation Approach

Consider solutions to $\mathbf{B}\mathbf{x}'(t) = \mathbf{A}\mathbf{x}(t)$ for the previous example:

$$\mathbf{A} = \begin{bmatrix} -1 & -5M \\ 0 & -5 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 1 & M \\ 0 & 1 \end{bmatrix}$$

Note that

$$\mathbf{B}^{-1}\mathbf{A} = \begin{bmatrix} -1 & 0 \\ 0 & -5 \end{bmatrix}.$$

- Since $\mathbf{B}^{-1}\mathbf{A}$ is normal and stable, solutions to $\mathbf{B}\mathbf{x}'(t) = \mathbf{A}\mathbf{x}(t)$ cannot exhibit growth.
- The parameter M affects the stability of eigenvalues of the pencil, but has no influence on the solution of $\mathbf{B}\mathbf{x}'(t) = \mathbf{A}\mathbf{x}(t)$.

GEPs: Matrix Behavior Approach

- More generally, premultiplying

$$\mathbf{B}\mathbf{x}'(t) = \mathbf{A}\mathbf{x}(t)$$

by some invertible matrix $\mathbf{S}$

$$\mathbf{S}\mathbf{B}\mathbf{x}'(t) = \mathbf{S}\mathbf{A}\mathbf{x}(t)$$

affects the perturbation theory of the pencil $(\mathbf{SA}, \mathbf{SB})$, but not the system driven by $(\mathbf{SB})^{-1}(\mathbf{SA}) = \mathbf{B}^{-1}\mathbf{A}$.

This fact suggests an alternative definition.

GEPs: Matrix Behavior Approach

- More generally, premultiplying

$$\mathbf{B}\mathbf{x}'(t) = \mathbf{A}\mathbf{x}(t)$$

by some invertible matrix $\mathbf{S}$

$$\mathbf{S}\mathbf{B}\mathbf{x}'(t) = \mathbf{S}\mathbf{A}\mathbf{x}(t)$$

affects the perturbation theory of the pencil $(\mathbf{SA}, \mathbf{SB})$, but not the system driven by $(\mathbf{SB})^{-1}(\mathbf{SA}) = \mathbf{B}^{-1}\mathbf{A}$.

This fact suggests an alternative definition.

Approach 2: matrix behavior

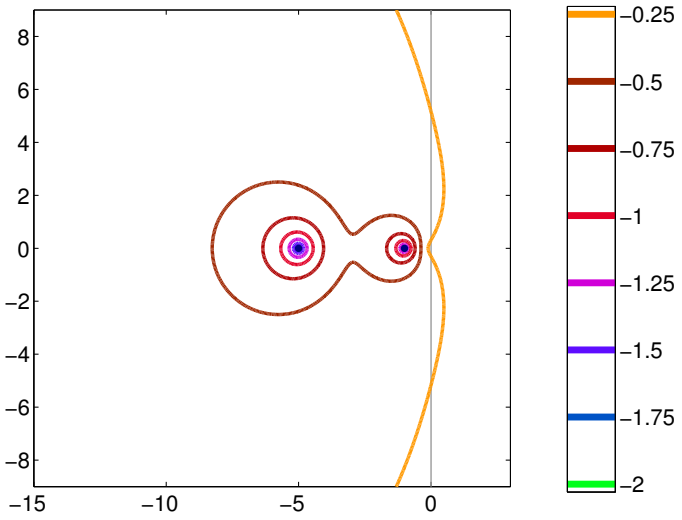
- [Ruhe, 1995], [Riedel, 1994] proposed:
For $\mathbf{A} \in \mathbf{C}^{n\times n}$ and invertible $\mathbf{B} \in \mathbf{C}^{n\times n}$,

$$\sigma_\varepsilon(\mathbf{A}, \mathbf{B}) = \sigma_\varepsilon(\mathbf{B}^{-1}\mathbf{A}).$$

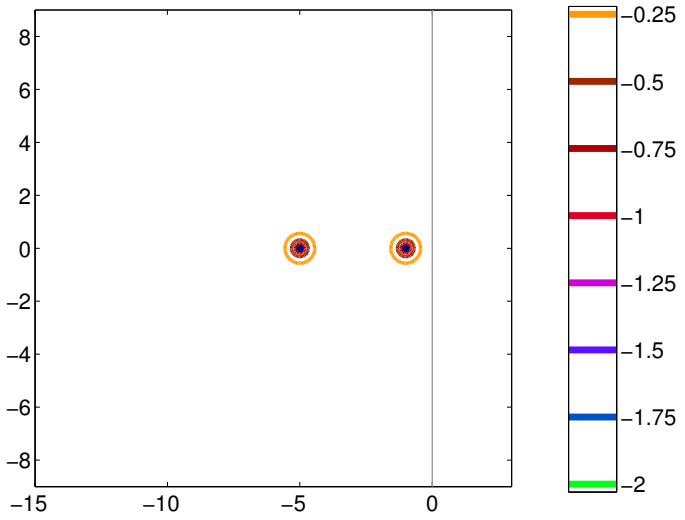
Comparison of GEP Pseudospectra

$$\mathbf{A} = \begin{bmatrix} -1 & -5M \\ 0 & -5 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 1 & M \\ 0 & 1 \end{bmatrix}$$

$$\mathbf{B}^{-1}\mathbf{A} = \begin{bmatrix} -1 & 0 \\ 0 & -5 \end{bmatrix}.$$



FGNT definition, $M = 1$



Ruhe definition

GEP: Behavior in Different Norms

Ruhe's definition is closely related to Riedel's [1994]:

- If $\mathbf{B}$ is Hermitian positive definite with Cholesky factorization $\mathbf{B} = \mathbf{L}\mathbf{L}^*$, then the ε-pseudospectrum of the matrix pencil $(\mathbf{A}, \mathbf{B})$ is the set

$$\sigma_\varepsilon(\mathbf{A}, \mathbf{B}) = \sigma_\varepsilon(\mathbf{L}^{-1}\mathbf{A}\mathbf{L}^{-*}).$$

This definition is the same as Ruhe's definition, but in a different norm. Let

$$\langle \mathbf{x}, \mathbf{y} \rangle_\mathbf{B} = \mathbf{y}^*\mathbf{B}\mathbf{x}, \qquad \|\mathbf{x}\|_\mathbf{B}^2 = \mathbf{x}^*\mathbf{B}\mathbf{x}.$$

Then

$$\|(z - \mathbf{L}^{-1}\mathbf{A}\mathbf{L}^{-*})^{-1}\|_2 = \|(z - \mathbf{B}^{-1}\mathbf{A})^{-1}\|_\mathbf{B}.$$

Pseudospectra for DAEs

Suppose $\mathbf{B}$ is singular, but $(\mathbf{A}, \mathbf{B})$ is regular ($\det(z\mathbf{B} - \mathbf{A}) \neq 0$ for some $z \in \mathbf{C}$).

$\mathbf{B}\mathbf{x}'(t) = \mathbf{A}\mathbf{x}(t)$ is a *differential-algebraic equation* (DAE).

Simple example:

$$\begin{aligned} x_1'(t) &= -x_1(t) \\ x_1(t) &= x_2(t) \end{aligned}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1'(t) \\ x_2'(t) \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

- Campbell & Meyer [1979], Campbell [1980]
- Kunkel & Mehrmann [2006]
- Descriptor systems: Benner, Byers, Mehrmann, Stykel, ...

DAEs, Simplest Case: A Invertible

Suppose that $\mathbf{A}$ is invertible, so that we can write

$$\mathbf{B}\mathbf{x}'(t) = \mathbf{A}\mathbf{x}(t)$$

in the form

$$\mathbf{A}^{-1}\mathbf{B}\mathbf{x}'(t) = \mathbf{x}(t).$$

First take a (generalized) Schur decomposition,

$$\mathbf{A}^{-1}\mathbf{B} = \mathbf{Q}\mathbf{T}\mathbf{Q}^* = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{Q}_2 \end{bmatrix} \begin{bmatrix} \mathbf{G} & \mathbf{D} \\ \mathbf{0} & \mathbf{N} \end{bmatrix} \begin{bmatrix} \mathbf{Q}_1^* \\ \mathbf{Q}_2^* \end{bmatrix},$$

where $\mathbf{Q}$ is unitary, $\mathbf{G}$ is invertible, and $\mathbf{N}$ is nilpotent.
(The degree of nilpotency corresponds to the *index* of the DAE.)

This decomposition reveals the algebraic structure of the problem:

$\mathbf{B}\mathbf{x}'(t) = \mathbf{A}\mathbf{x}(t)$, $\mathbf{x}(0) = \mathbf{x}_0$ has a solution if and only if $\mathbf{x}_0 \in \operatorname{Range} \mathbf{Q}_1$.

DAEs, Simplest Case: A Invertible

$$\mathbf{A}^{-1}\mathbf{B} = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{Q}_2 \end{bmatrix} \begin{bmatrix} \mathbf{G} & \mathbf{D} \\ \mathbf{0} & \mathbf{N} \end{bmatrix} \begin{bmatrix} \mathbf{Q}_1^* \\ \mathbf{Q}_2^* \end{bmatrix}$$

We wish to write the solution to the DAE as

$$\mathbf{x}(t) = \mathbf{Q}_1\mathbf{y}(t) + \mathbf{Q}_2\mathbf{z}(t).$$

One can show that $\mathbf{z}(t) = \mathbf{0}$ for all t, so we seek:

$$\mathbf{G}\mathbf{y}'(t) = \mathbf{y}(t), \qquad \mathbf{z}(t) = \mathbf{0}.$$

Hence write $\mathbf{x}(t) = \mathbf{Q}_1 e^{t\mathbf{G}^{-1}}\mathbf{y}(0)$, i.e.,

$$\boxed{\mathbf{x}(t) = \mathbf{Q}_1 e^{t\mathbf{G}^{-1}}\mathbf{Q}_1^*\mathbf{x}_0.}$$

Special case: $\mathbf{B}$ invertible $\Longrightarrow \mathbf{Q}_1 = \mathbf{I}$ and $\mathbf{G} = \mathbf{A}^{-1}\mathbf{B}$, so

$$\mathbf{x}(t) = e^{t\,\mathbf{G}^{-1}}\mathbf{x}_0 = e^{t\,\mathbf{B}^{-1}\mathbf{A}}\mathbf{x}_0.$$

Pseudospectra of $(\mathbf{A}, \mathbf{B})$ for Transient Analysis of DAEs

Suppose $\mathbf{x}_0 \in \text{Range}\,\mathbf{Q}_1$, with the columns of $\mathbf{Q}_1$ forming an orthonormal basis for the invariant subspace of the pencil associated with finite eigenvalues. Then

$$\mathbf{x}(t) = \mathbf{Q}_1 \mathrm{e}^{t\mathbf{G}^{-1}} \mathbf{Q}_1^* \mathbf{x}_0,$$

where $\mathbf{G} = \mathbf{Q}_1^* \mathbf{A}^{-1} \mathbf{B} \mathbf{Q}_1 \in \mathbf{C}^{m \times m}$ ($m = \#$ of finite eigenvalues).

We can bound the norm of the solution by

$$\|\mathbf{x}(t)\| \leq \|\mathrm{e}^{t\mathbf{G}^{-1}}\| \|\mathbf{x}_0\|.$$

Pseudospectra of $(\mathbf{A}, \mathbf{B})$ for Transient Analysis of DAEs

Suppose $\mathbf{x}_0 \in \text{Range}\,\mathbf{Q}_1$, with the columns of $\mathbf{Q}_1$ forming an orthonormal basis for the invariant subspace of the pencil associated with finite eigenvalues. Then

$$\mathbf{x}(t) = \mathbf{Q}_1 e^{t\mathbf{G}^{-1}} \mathbf{Q}_1^* \mathbf{x}_0,$$

where $\mathbf{G} = \mathbf{Q}_1^* \mathbf{A}^{-1} \mathbf{B} \mathbf{Q}_1 \in \mathbf{C}^{m \times m}$ ($m = \#$ of finite eigenvalues).

We can bound the norm of the solution by

$$\|\mathbf{x}(t)\| \leq \|e^{t\mathbf{G}^{-1}}\| \|\mathbf{x}_0\|.$$

Definition (Pseudospectra of a Regular Pencil, A invertible)

Consider the matrix pencil $\mathbf{A} - \lambda \mathbf{B}$ with $\mathbf{A}$ invertible, and Schur factorization

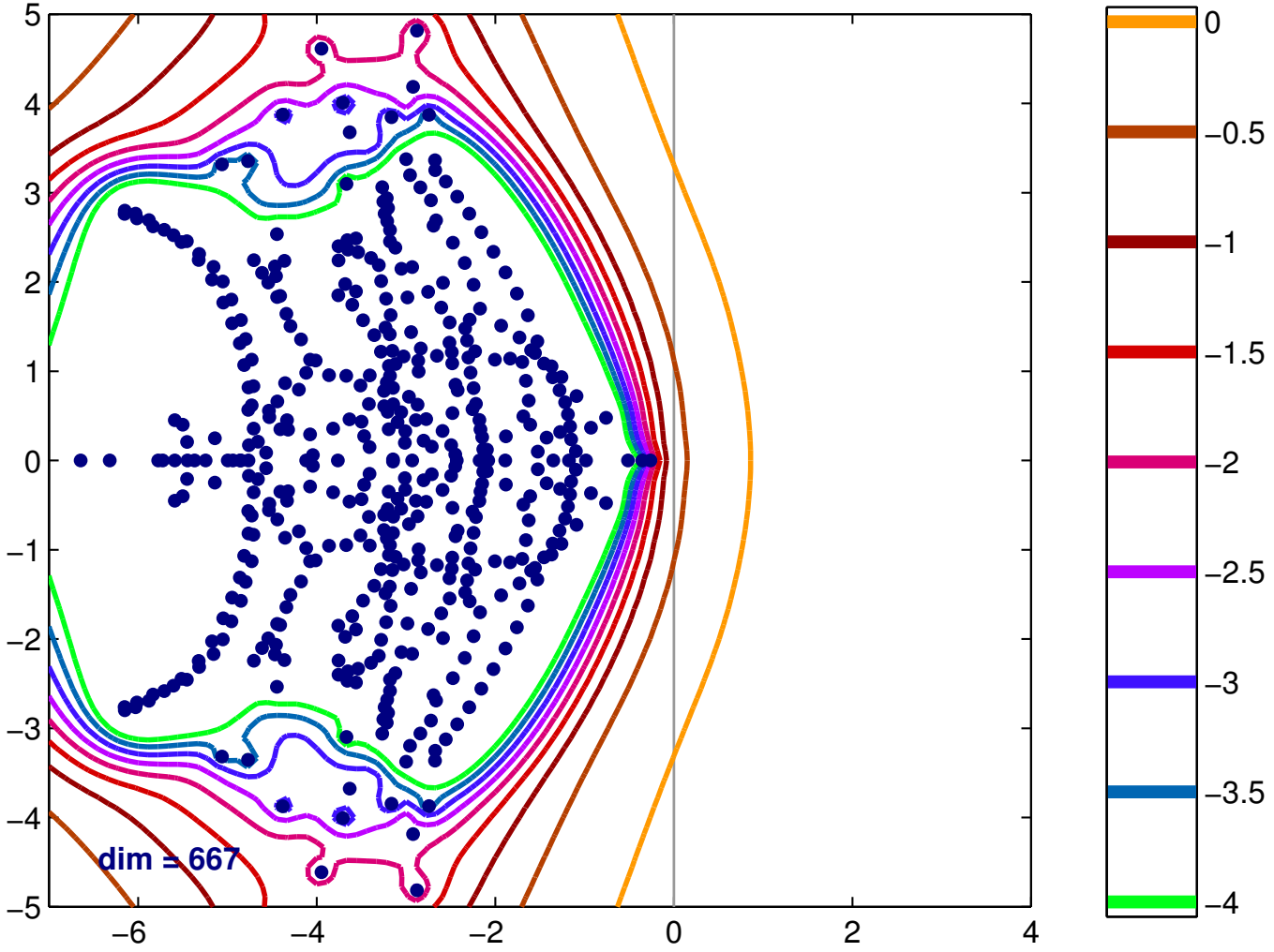
$$\mathbf{A}^{-1}\mathbf{B} = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{Q}_2 \end{bmatrix} \begin{bmatrix} \mathbf{G} & \mathbf{D} \\ \mathbf{0} & \mathbf{N} \end{bmatrix} \begin{bmatrix} \mathbf{Q}_1^* \\ \mathbf{Q}_2^* \end{bmatrix}$$

for $\mathbf{N}$ nilpotent and $0 \notin \sigma(\mathbf{G})$.
The *ε-pseudospectrum of the matrix pencil* $(\mathbf{A}, \mathbf{B})$ is

$$\sigma_\varepsilon(\mathbf{A}, \mathbf{B}) := \sigma_\varepsilon((\mathbf{Q}_1^* \mathbf{A}^{-1} \mathbf{B} \mathbf{Q}_1)^{-1}).$$

Examples: Stability Analysis for Imcompressible Flow

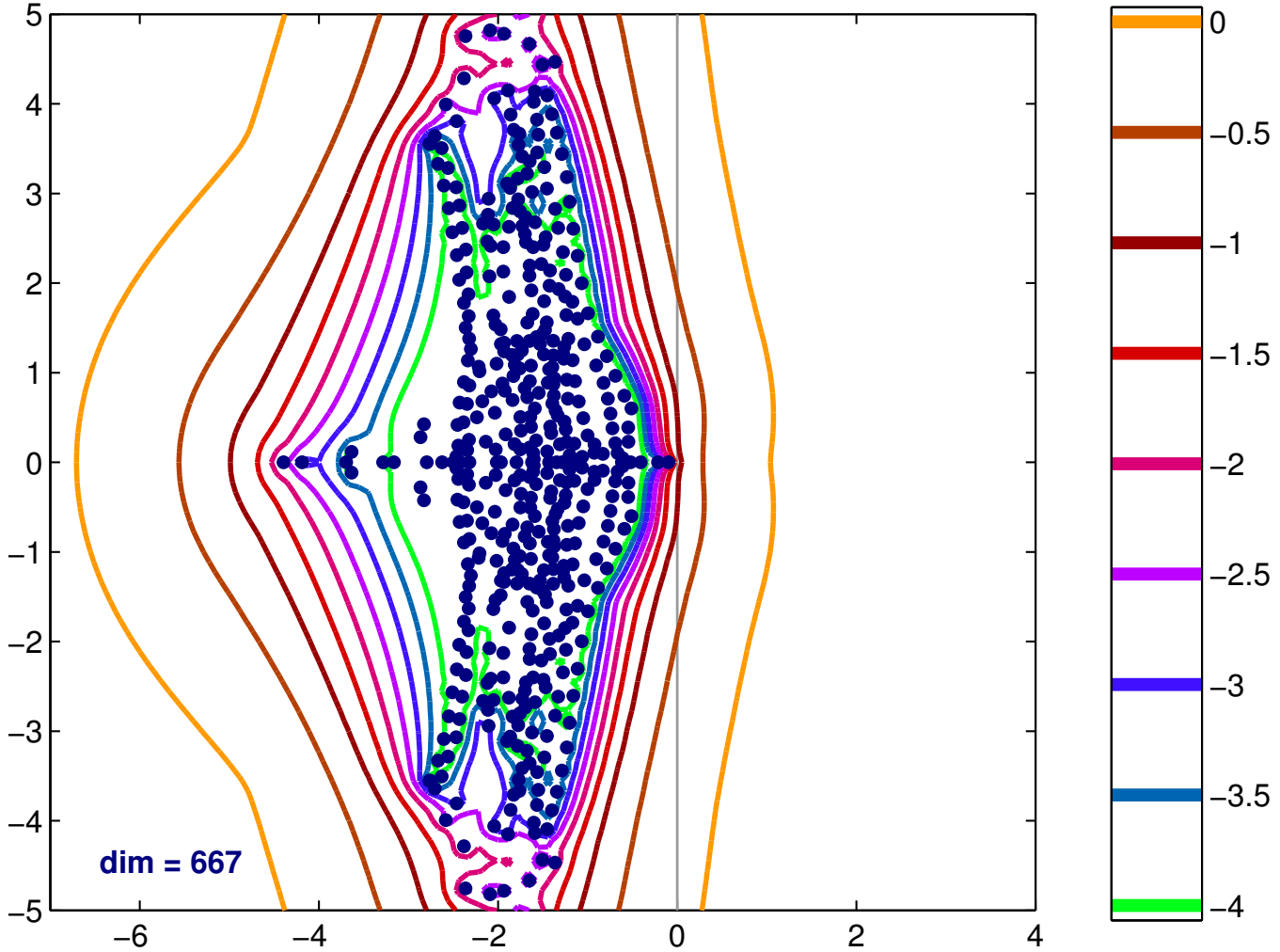
Pseudospectra for a matrix pencil derived from incompressible flow over a backward facing step, discretized via IFISS.



$1/\nu = 50$

Examples: Stability Analysis for Imcompressible Flow

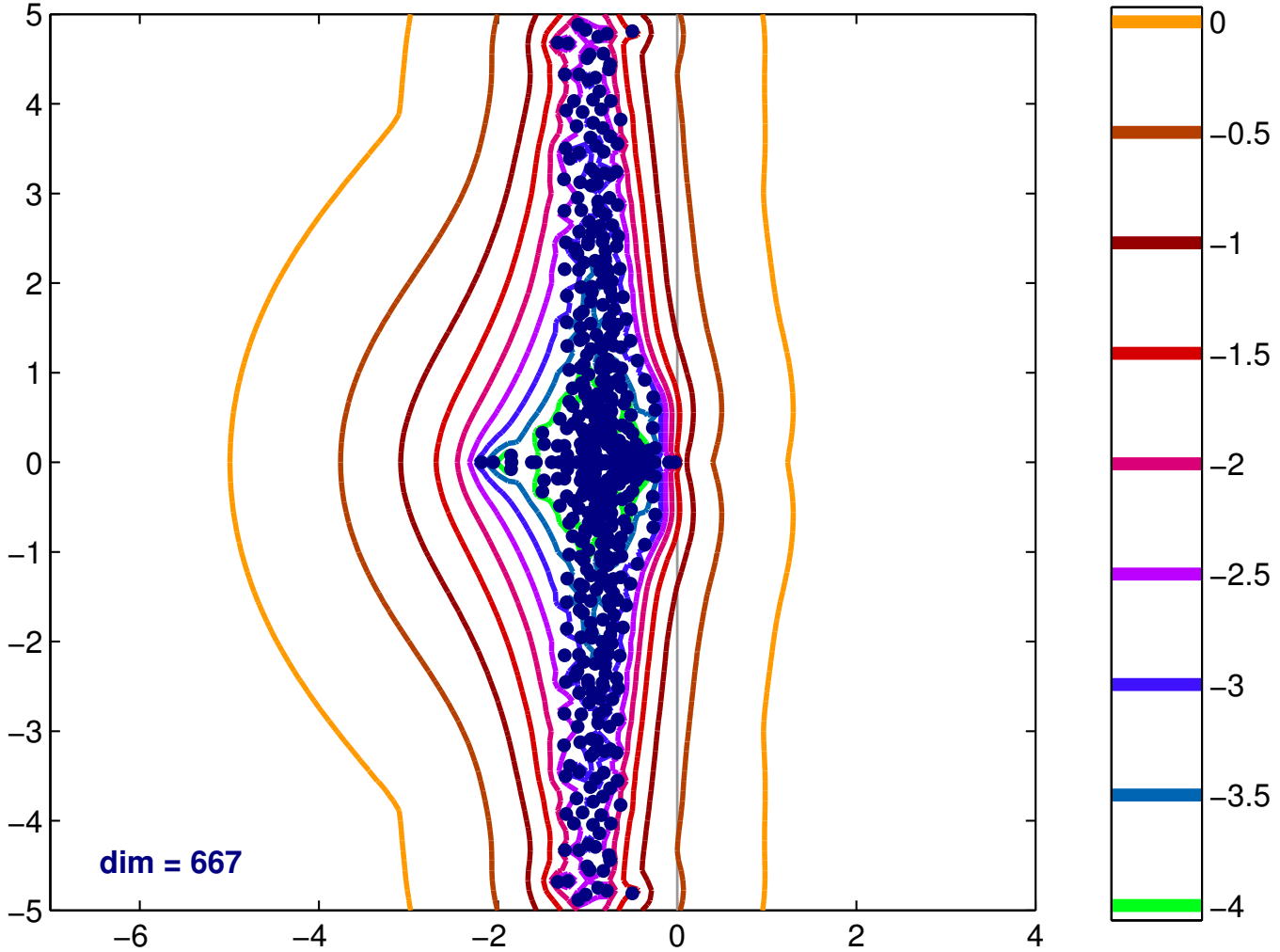
Pseudospectra for a matrix pencil derived from incompressible flow over a backward facing step, discretized via IFISS.



$1/\nu = 100$

Examples: Stability Analysis for Imcompressible Flow

Pseudospectra for a matrix pencil derived from incompressible flow over a backward facing step, discretized via IFISS.



$1/\nu = 200$

Summary

Overview of Lectures

These lectures address modern tools
for the spectral analysis of dynamical systems.
We shall cover a mix of theory, computation, and applications.

Goal: By the end, you should be equipped to understand phenomena that many people find quite mysterious when encountered in the wild.

Lecture 1:

- Normality and Nonnormality
- Pseudospectra
- Bounding Functions of Matrices

Lecture 2:

- Balanced Truncation and Lyapunov Equations
- Moment Matching Model Reduction
- Differential Algebraic Equations

Exercises

Exercises for the Afternoon

- Download EigTool (`http://www.cs.ox.ac.uk/pseudospectra/eigtool`)
  Run `eigtool` from a MATLAB prompt.
  Experiment with menu option: `Demos/Dense/...`
  For dense matrices, experiment with menu option: `Transients`
  For dense matrices, experiment with menu option: `Numbers`
  For dense matrices, experiment with `Pmode+epsilon` button
  Experiment with menu option: `Demos/Sparse/...`

- Computation of pseudospectra of a vibrating string in energy norm.

- Population dynamics: design a transiently growing population.

- Analysis of a nonlinear dynamical system.

- Estimate pseudospectral bounds on $\|\mathbf{A}^k\|$ and $\|e^{t\mathbf{A}}\|$.