Pseudospectra and Nonnormal Dynamical Systems

Mark Embree and Russell Carden
Computational and Applied Mathematics
Rice University
Houston, Texas

ELGERSBURG

MARCH 2012

Overview of the Course

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

Lecture 1: Introduction to Nonnormality and Pseudospectra

Lecture 2: Functions of Matrices

Lecture 3: Toeplitz Matrices and Model Reduction

Lecture 4: Model Reduction, Numerical Algorithms, Differential Operators

Lecture 5: Discretization, Extensions, Applications

Outline for Today

Lecture 5: Discretization, Extensions, Applications

- Discretization of Differential Operators
- Structured Pseudospectra
- Pseudospectra for Generalized Eigenvalue Problems / DAEs
- Pseudospectra for Polynomial Eigenvalue Problems
- Applications

5(a) Discretization of Differential Operators

Discretization of Differential Operators

- Many numerical methods are available for discretizing differential operators: finite differences, finite elements, spectral/collocation methods – many variants.
- For simple operators, it is difficult to beat spectral/collocation methods, which give (1) high order accuracy; (2) small (but dense) discretization matrices.
- Trefethen's *Spectral Methods in MATLAB* (2000) describes on effective approach that we have used throughout our examples here.

Discretization of Differential Operators

"Chebyshev Pseudospectral Collocation Methods"

- Key idea: represent a function u by the degree-n polynomial p_n that interpolates it at "Chebyshev points" (on $[-1, 1]$, say):

$$x_j = \cos\left(\frac{j\pi}{n}\right), \quad j = 0, \ldots, n,$$

so that

$$p_n(x_j) = u(x_j), \qquad j = 0, \ldots, n.$$

Discretization of Differential Operators

"Chebyshev Pseudospectral Collocation Methods"

- Key idea: represent a function u by the degree-n polynomial p_n that interpolates it at "Chebyshev points" (on $[-1, 1]$, say):

  $$x_j = \cos\left(\frac{j\pi}{n}\right), \quad j = 0, \ldots, n,$$

  so that

  $$p_n(x_j) = u(x_j), \qquad j = 0, \ldots, n.$$

- Then approximate $u'(x)$ by $p_n'(x)$, etc.

Chebyshev Pseudospectral Collocation Methods

- These operations can be encoded in matrices.
  Represent u by its values on the Chebyshev grid:

$$\mathbf{u} = [u(x_0)\ u(x_1)\ \ldots\ u(x_n)]^{\mathsf{T}}.$$

- Differentiation is them represented as a matrix-vector product:

$$\begin{bmatrix} u'(x_0) \\ u'(x_1) \\ \vdots \\ u'(x_n) \end{bmatrix} \approx \begin{bmatrix} p_n'(x_0) \\ p_n'(x_1) \\ \vdots \\ p_n'(x_n) \end{bmatrix} = \mathbf{D}\mathbf{u}.$$

- To apply two derivatives, square the matrix:

$$\begin{bmatrix} u''(x_0) \\ u''(x_1) \\ \vdots \\ u''(x_n) \end{bmatrix} \approx \begin{bmatrix} p_n''(x_0) \\ p_n''(x_1) \\ \vdots \\ p_n''(x_n) \end{bmatrix} = \mathbf{D}^2\mathbf{u}.$$

- To impose Dirichlet boundary conditions on both ends, set $u(x_0) = u(x_n) = 0$, and truncate the discretization to the interior $(n-1) \times (n-1)$ submatrix.

Chebyshev Pseudospectral Collocation Methods

- To approximate, e.g., the L^2 norm, we use *Clenshaw–Curtis quadrature*, i.e., the exact integral of the Chebyshev interpolant.

$$\|u\|_{L^2} \approx \sum_{j=0}^{n} w_j |u(x_j)|^2.$$

Chebyshev Pseudospectral Collocation Methods

- To approximate, e.g., the L^2 norm, we use *Clenshaw–Curtis quadrature*, i.e., the exact integral of the Chebyshev interpolant.

$$\|u\|_{L^2} \approx \sum_{j=0}^{n} w_j |u(x_j)|^2.$$

- Trefethen's `cheb.m` MATLAB code constructs this differentiation matrix:
  `[D,x] = cheb(n)`
- Trefethen's `clencurt.m` MATLAB code constructs the Clenshaw-Curtis notes and weights:
  `[x,w] = clencurt(n)`

Chebyshev Pseudospectral Collocation Methods

- To approximate, e.g., the L^2 norm, we use *Clenshaw–Curtis quadrature*, i.e., the exact integral of the Chebyshev interpolant.

$$\|u\|_{L^2} \approx \sum_{j=0}^{n} w_j |u(x_j)|^2.$$

- Trefethen's `cheb.m` MATLAB code constructs this differentiation matrix:
  `[D,x] = cheb(n)`
- Trefethen's `clencurt.m` MATLAB code constructs the Clenshaw-Curtis notes and weights:
  `[x,w] = clencurt(n)`

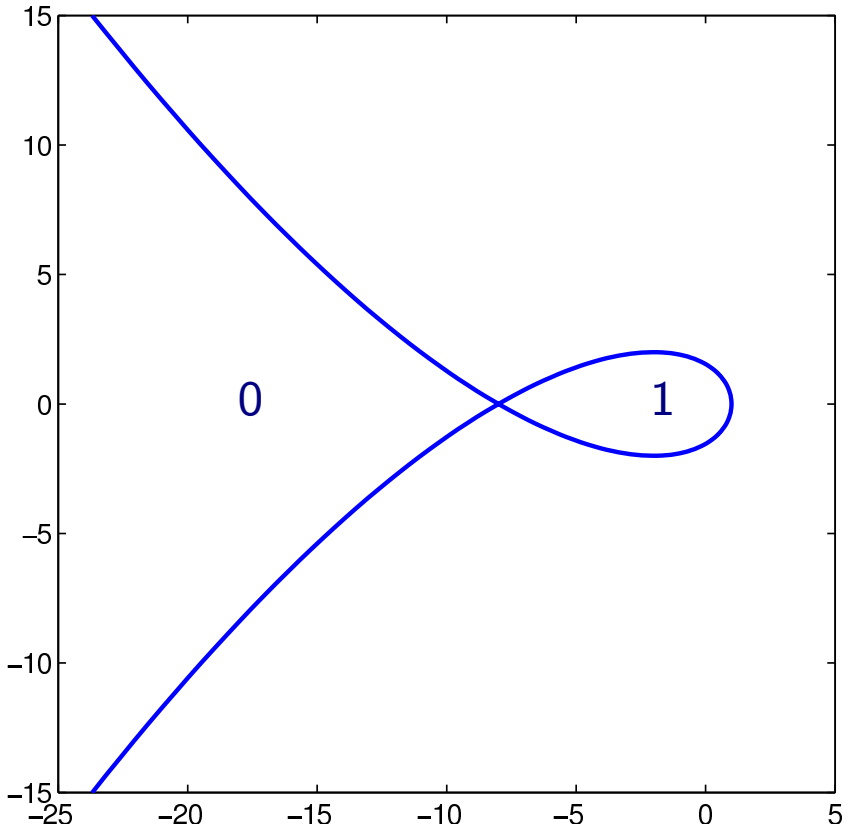
Code to discretize the Laplacian on $[-1,1]$ with homogeneous Dirichlet boundary conditions, and compute its L^2-norm pseudospectra:

```
[D,x] = cheb(n);          % construct differentiation matrix on [-1,1]
L = D*D;                  % compute second derivative matrix
L = L(2:n,2:n);           % impose homogeneous Dirichlet boundary conditions
[x,w] = clencurt(n);      % Clenshaw-Curtis quadrature weights
R = diag(sqrt(w(2:n)));   % matrix that defines the L2 norm, G = R'*R
eigtool(R*L/R)            % compute pseudospectra
```

Pseudospectra of Differential Operators: example 3

Consider the operator

$$\mathbf{A}u = u + 3u' + 3u'' + u'''$$

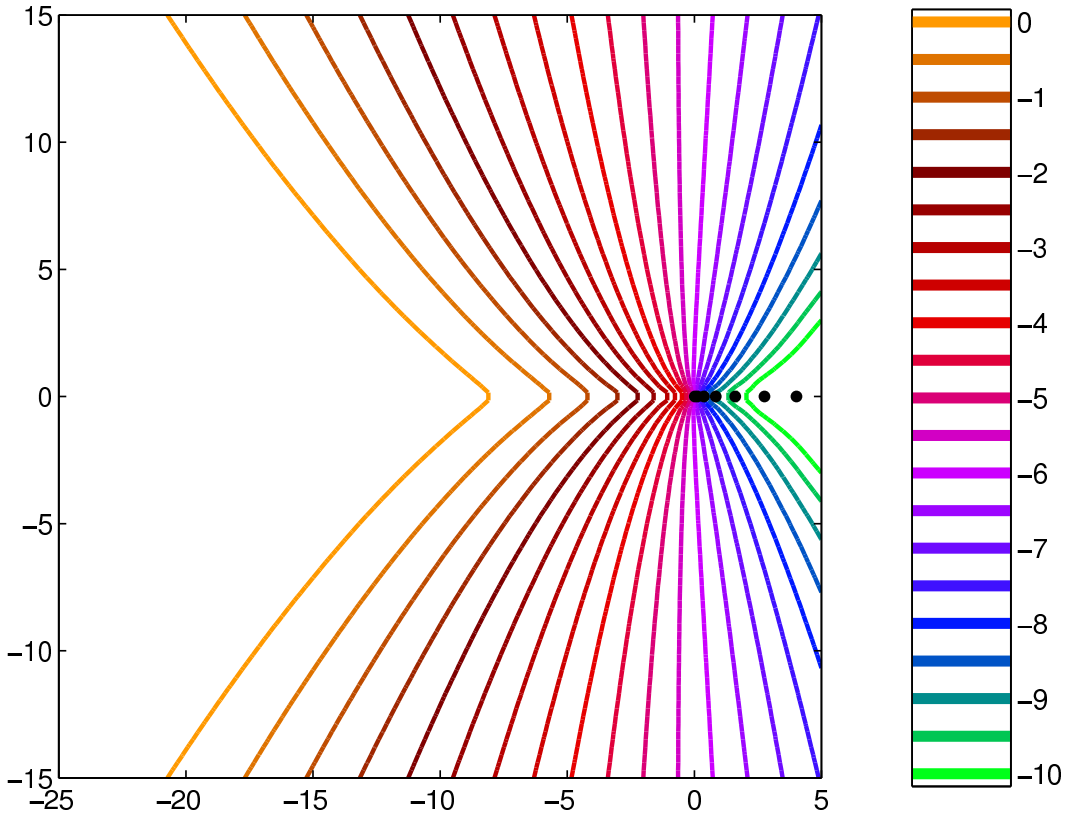


$a(\Gamma_R)$ for $R = 8$

Pseudospectra of Differential Operators: example 3

Consider the operator

$$\mathbf{A}u = u + 3u' + 3u'' + u'''$$

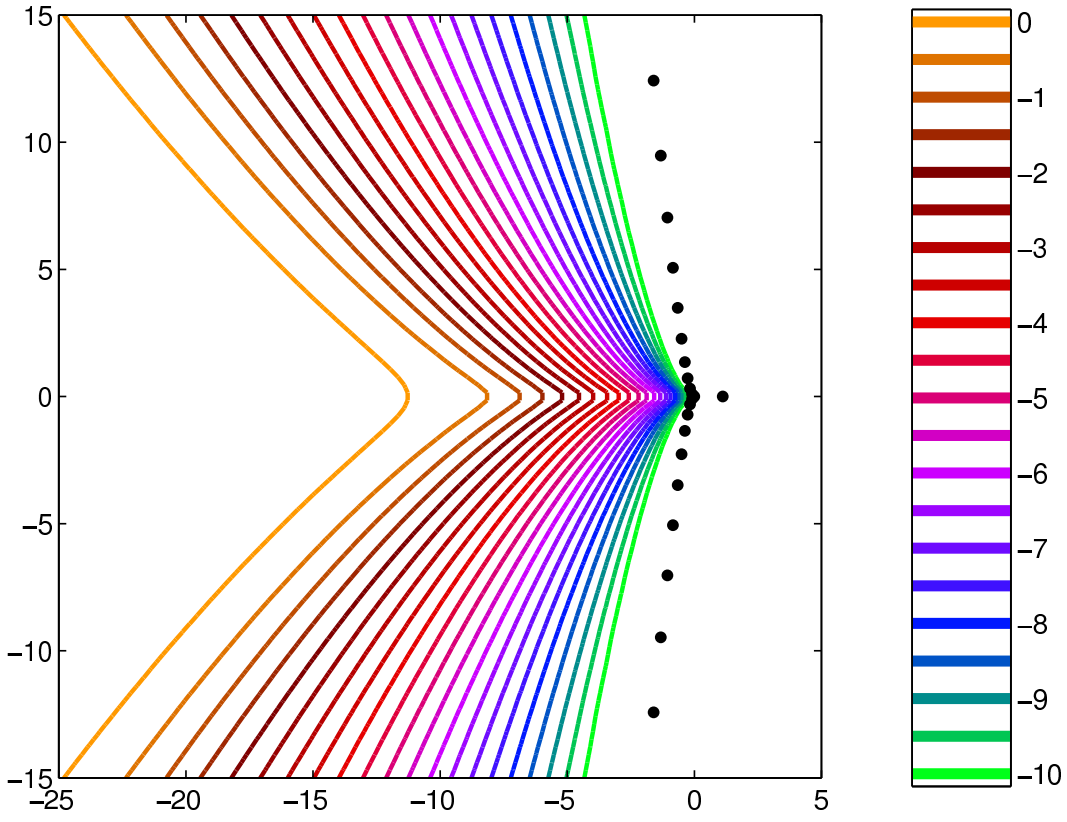


$\sigma_\varepsilon(\mathbf{A})$ for $\ell = 16$, $\beta = 1$, $\gamma = 2$

Pseudospectra of Differential Operators: example 3

Consider the operator

$$\mathbf{A}u = u + 3u' + 3u'' + u'''$$

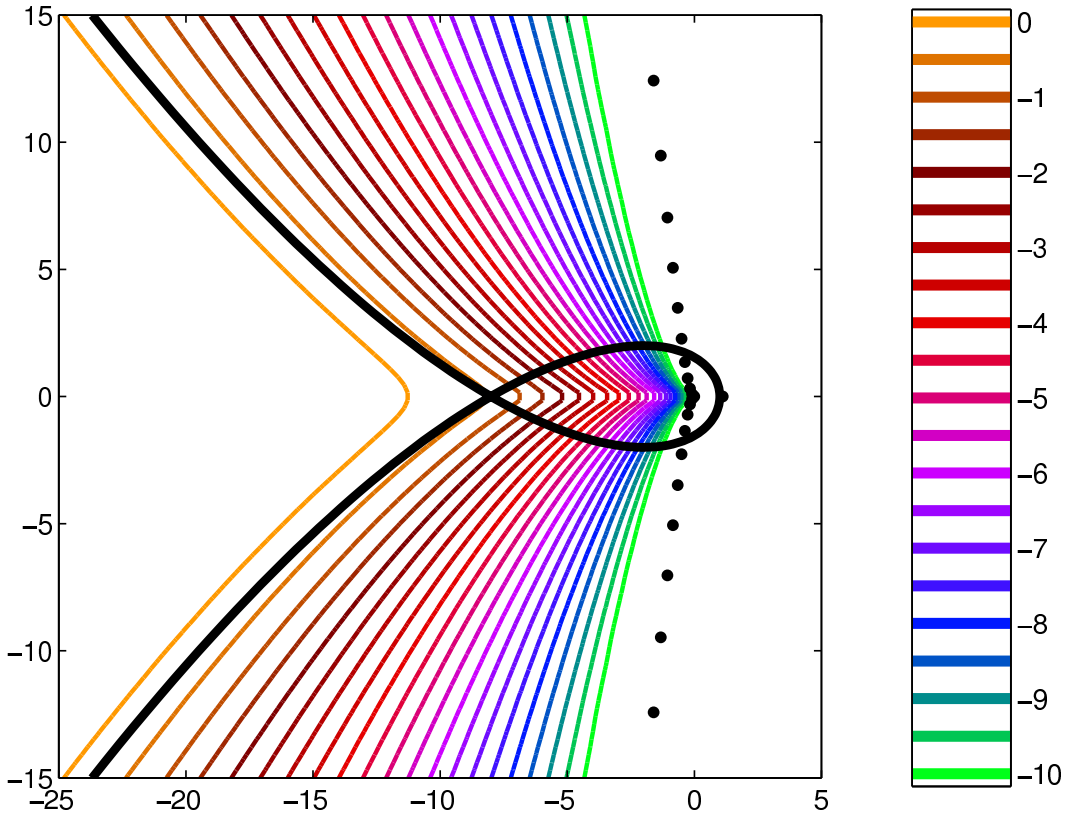


$\sigma_\varepsilon(\mathbf{A})$ for $\ell = 32$, $\beta = 1$, $\gamma = 2$

Pseudospectra of Differential Operators: example 3

Consider the operator

$$\mathbf{A}u = u + 3u' + 3u'' + u'''$$

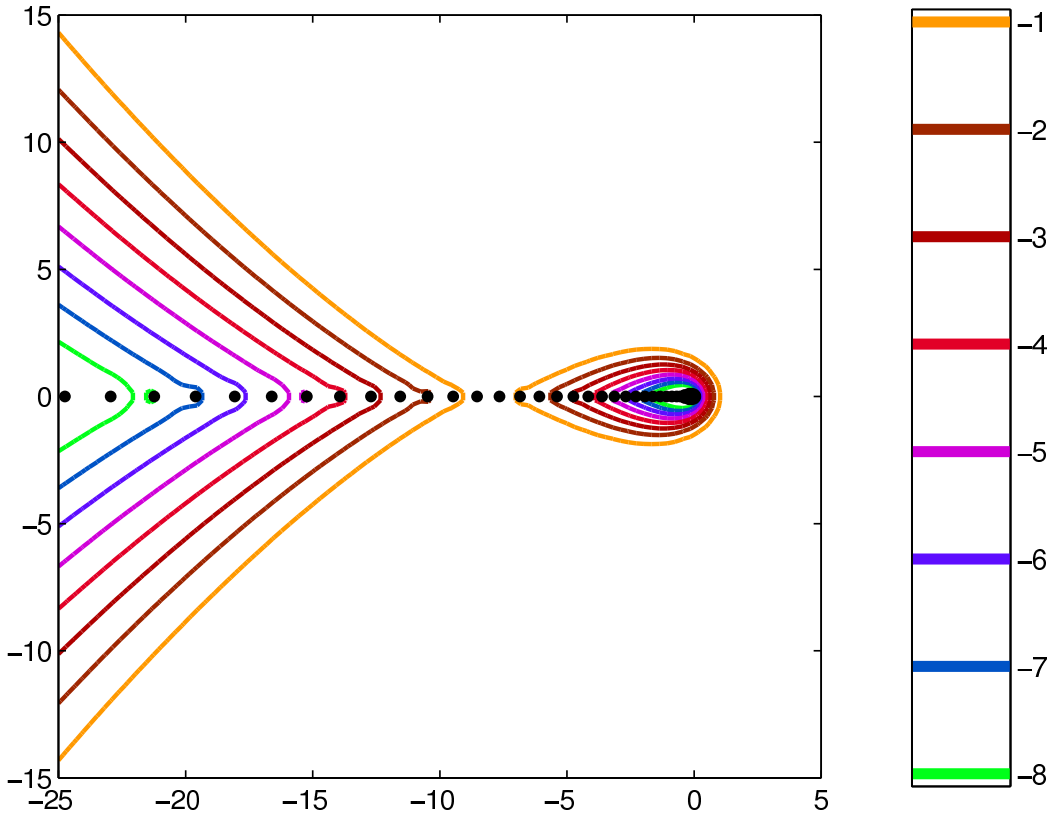


$\sigma_\varepsilon(\mathbf{A})$ for $\ell = 32$, $\beta = 1$, $\gamma = 2$

Pseudospectra of Differential Operators: example 3

Consider the operator

$$\mathbf{A}u = u + 3u' + 3u'' + u'''$$

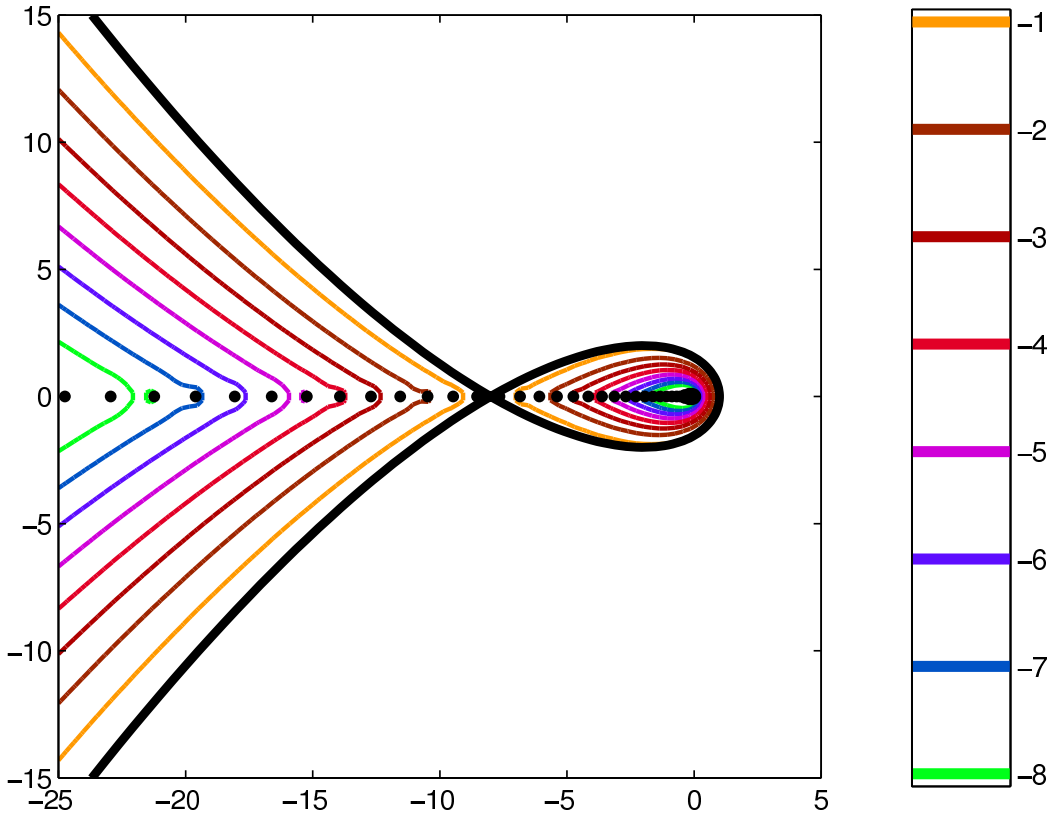


$\sigma_\varepsilon(\mathbf{A})$ for $\ell = 50$, $\beta = 2$, $\gamma = 1$

Pseudospectra of Differential Operators: example 3

Consider the operator

$$\mathbf{A}u = u + 3u' + 3u'' + u'''$$



$\sigma_\varepsilon(\mathbf{A})$ for $\ell = 50$, $\beta = 2$, $\gamma = 1$

5(b) Structured Pseudospectra

Structured Pseudospectra

Definition (ε-pseudospectrum)

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

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

The choice $\mathbf{E} \in \mathbf{C}^{n\times n}$ might seem contrived: if you view pseudospectra as a tool for understanding eigenvalue sensitivity, this perspective is reasonable.

Structured Pseudospectra

Definition (ε-pseudospectrum)

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

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

The choice $\mathbf{E} \in \mathbf{C}^{n\times n}$ might seem contrived: if you view pseudospectra as a tool for understanding eigenvalue sensitivity, this perspective is reasonable.

- Eigenvalue computations
  Stable algorithms give exact eigenvalues of $\mathbf{A} + \mathbf{E}$ for $\|\mathbf{E}\| = \mathcal{O}(\|\mathbf{A}\|\varepsilon_{\text{mach}})$.
  Good methods for real $\mathbf{A}$ give real $\mathbf{E}$.
  Good methods for Hermitian $\mathbf{A}$ give Hermitian $\mathbf{E}$.
  Good methods for Hamiltonian $\mathbf{A}$ give Hamiltonian $\mathbf{E}$.
  Interval matrices have structured entries.

Structured Pseudospectra

Definition (ε-pseudospectrum)

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

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

The choice $\mathbf{E} \in \mathbf{C}^{n\times n}$ might seem contrived: if you view pseudospectra as a tool for understanding eigenvalue sensitivity, this perspective is reasonable.

- Eigenvalue computations
  Stable algorithms give exact eigenvalues of $\mathbf{A} + \mathbf{E}$ for $\|\mathbf{E}\| = \mathcal{O}(\|\mathbf{A}\|\varepsilon_{\text{mach}})$.
  Good methods for real $\mathbf{A}$ give real $\mathbf{E}$.
  Good methods for Hermitian $\mathbf{A}$ give Hermitian $\mathbf{E}$.
  Good methods for Hamiltonian $\mathbf{A}$ give Hamiltonian $\mathbf{E}$.
  Interval matrices have structured entries.
- Uncertain systems
  The physical system often dictates that $\mathbf{A}$ have certain structure (e.g., Toeplitz, nonnegative, stochastic, etc.). Model uncertainty causes matrix entries to change while preserving the structure.

Structured Pseudospectra

Definition (ε-pseudospectrum)

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

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

Concerns about structured perturbations motivate a new definition.

Definition (structured ε-pseudospectrum)

Let $\mathbb{S}$ denote some subset of $\mathbf{C}^{n\times n}$. For any $\varepsilon > 0$, the *$\mathbb{S}$-structured ε-pseudospectrum of* $\mathbf{A}$, denoted $\sigma_\varepsilon^{\mathbb{S}}(\mathbf{A})$, is the set

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

- $\sigma_\varepsilon^{\mathbb{S}}(\mathbf{A}) \subseteq \sigma_\varepsilon(\mathbf{A})$ for any choice of $\mathbb{S}$.
- We cannot equivalently define $\sigma_\varepsilon^{\mathbb{S}}(\mathbf{A})$ in terms of the resolvent norm or pseudoeigenvectors.

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

Structured Pseudospectra: Stochastic Example

Stochastic matrices have entries in $[0, 1]$ with rows that sum to one.

$$\begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/8 & 3/4 & 1/8 \\ 0 & 1/2 & 1/2 \end{bmatrix} \qquad \begin{bmatrix} 1/4 & 3/4 & 0 \\ 1/4 & 1/4 & 1/2 \\ 1/2 & 1/4 & 1/4 \end{bmatrix}$$

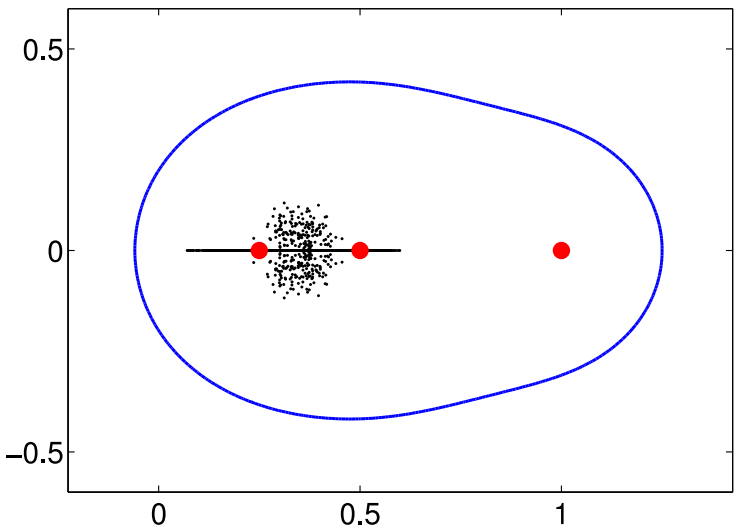
We shall make structured perturbations of size bounded by $\varepsilon = 1/4$ that preserve the stochastic structure of $\mathbf{A}$.

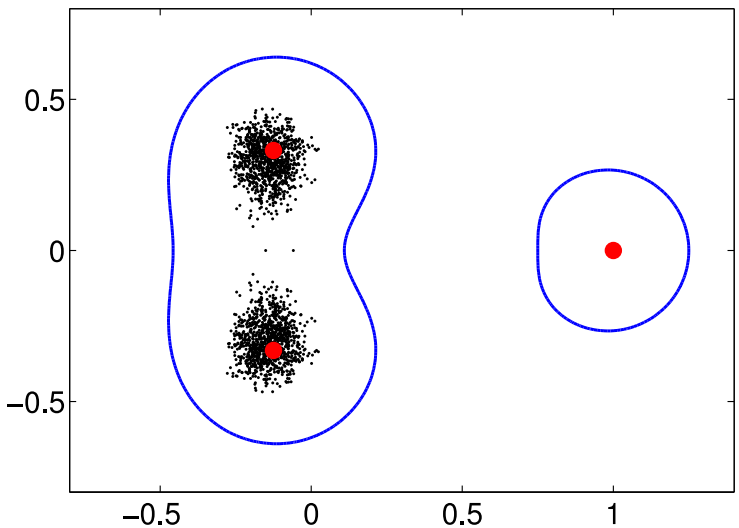
Structured Pseudospectra: Stochastic Example

Stochastic matrices have entries in $[0, 1]$ with rows that sum to one.

$$\begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/8 & 3/4 & 1/8 \\ 0 & 1/2 & 1/2 \end{bmatrix} \qquad \begin{bmatrix} 1/4 & 3/4 & 0 \\ 1/4 & 1/4 & 1/2 \\ 1/2 & 1/4 & 1/4 \end{bmatrix}$$

We shall make structured perturbations of size bounded by $\varepsilon = 1/4$ that preserve the stochastic structure of $\mathbf{A}$.





Blue curves shows the boundary of $\sigma_\varepsilon(\mathbf{A})$ for $\|\cdot\| = \|\cdot\|_\infty$.
Red dots show the eigenvalues of $\mathbf{A}$. Black dots show eigenvalues for 1000 $\mathbf{E}$.

Structured Pseudospectra: Toeplitz Example

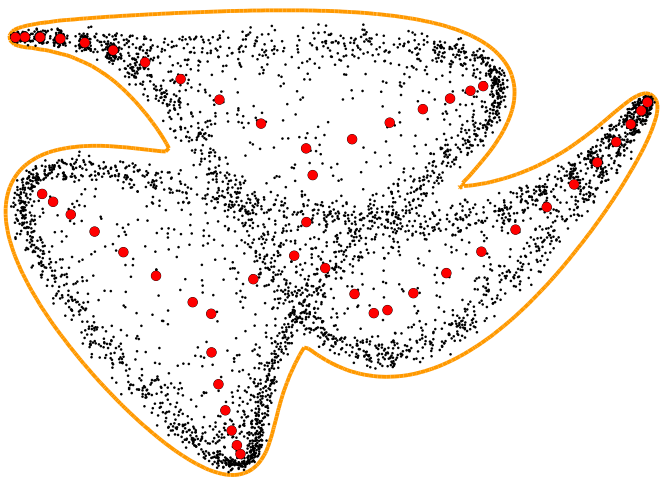
Consider a Toeplitz matrix with $N = 50$, a symbol used in Lecture 3:

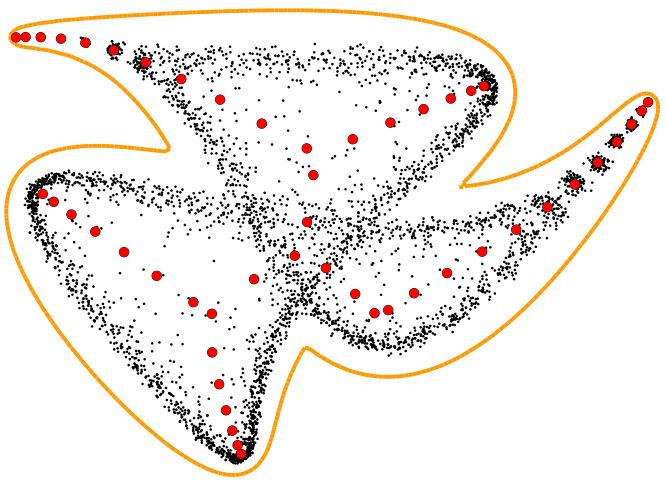
$$a(z) = iz^{-4} + z^{-2} + 2z^{-1} + 5z^2 + iz^5.$$

The orange curves show the boundary of $\sigma_\varepsilon(\mathbf{A})$ for $\varepsilon = 10^{-1}$ (unstructured).

- Compute 200 random *unstructured* perturbations and display eigenvalues.
- Compute 200 random *Toeplitz* perturbations and display eigenvalues.

Which plot is which?





Structured Pseudospectra: Toeplitz Example

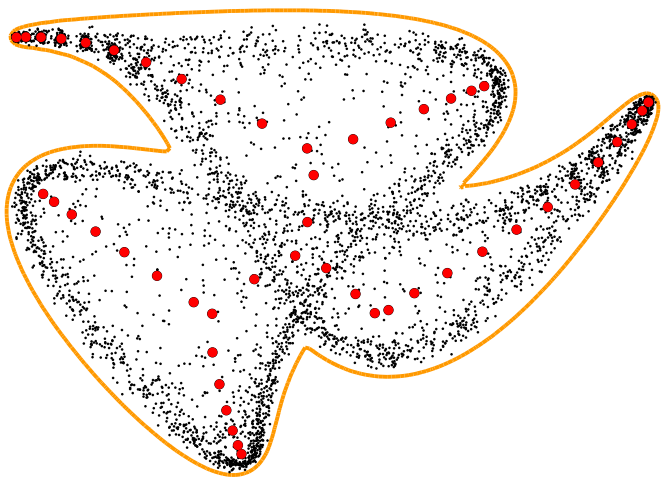
Consider a Toeplitz matrix with $N = 50$, a symbol used in Lecture 3:

$$a(z) = iz^{-4} + z^{-2} + 2z^{-1} + 5z^2 + iz^5.$$

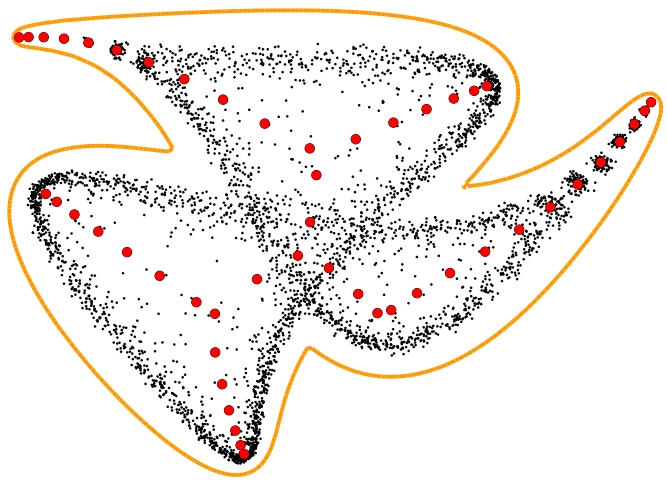
The orange curves show the boundary of $\sigma_\varepsilon(\mathbf{A})$ for $\varepsilon = 10^{-1}$ (unstructured).

- Compute 200 random *unstructured* perturbations and display eigenvalues.
- Compute 200 random *Toeplitz* perturbations and display eigenvalues.

Which plot is which?



TOEPLITZ



UNSTRUCTURED

Some Special Structures

On the last slide, we saw close apparent resemblance of the structured Toepltiz pseudospectrum and unstructured pseudospectrum. In fact, this is a general phenomenon for a number of important structures.

Theorem (Rump, 2006)

Suppose $\mathbf{A} \in \mathbb{S}$, *where* $\mathbb{S}$ *comprises one of: complex symmetric, persymmetric, circulant, Toeplitz, symmetric Toeplitz, Hankel or persymmetric Hankel matrices. Then for all* $\varepsilon > 0$,

$$\sigma_\varepsilon^{\mathbb{S}}(\mathbf{A}) = \sigma_\varepsilon(\mathbf{A}).$$

If $\mathbb{S}$ *comprises Hermitian matrices and* $\mathbf{A} \in \mathbb{S}$, *then*

$$\sigma_\varepsilon^{\mathbb{S}}(\mathbf{A}) = \sigma_\varepsilon(\mathbf{A}) \cap \boldsymbol{R}.$$

If $\mathbb{S}$ *comprises skew-Hermitian matrices and* $\mathbf{A} \in \mathbb{S}$, *then*

$$\sigma_\varepsilon^{\mathbb{S}}(\mathbf{A}) = \sigma_\varepsilon(\mathbf{A}) \cap i\boldsymbol{R}.$$

Real Structured Perturbations

For most structures not covered by Rump's Theorem, $\sigma_\varepsilon^{\mathbb{S}}(\mathbf{A})$ is very hard to compute (e.g., when $\mathbb{S}$ contains matrices with the same sparsity pattern as $\mathbf{A}$).

But several special cases are computable, e.g., with $\sigma_\varepsilon^{\mathbb{S}}(\mathbf{A})$ characterized for small ε; see Karow et al. [2003 – 2011...].

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

Definition (real structured ε-pseudospectrum)

For any $\varepsilon > 0$, the *real structured ε-pseudospectrum of* $\mathbf{A}$ is the set

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

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

Real Structured Perturbations

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

$$d_{\mathbf{C}}(\mathbf{A}, z) = \left(s_1\left((z - \mathbf{A})^{-1} \right) \right)^{-1}.$$

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

$$d_{\mathbf{R}}(\mathbf{A}, z) = \left(\inf_{\gamma \in (0,1]} s_2 \left(\begin{bmatrix} \mathrm{Re}(z - \mathbf{A})^{-1} & -\gamma \mathrm{Im}(z - \mathbf{A})^{-1} \\ \gamma^{-1} \mathrm{Im}(z - \mathbf{A})^{-1} & \mathrm{Re}(z - \mathbf{A})^{-1} \end{bmatrix} \right) \right)^{-1}$$

[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]:

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

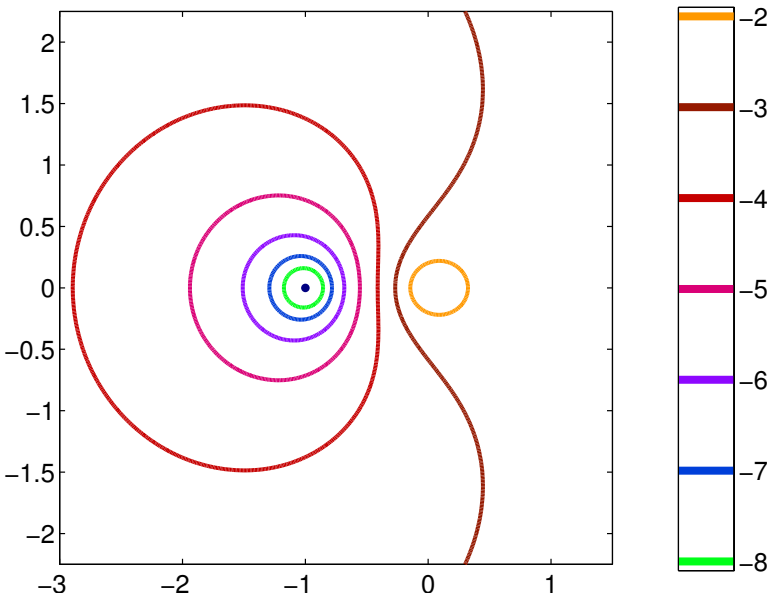
with $M = 10$.

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

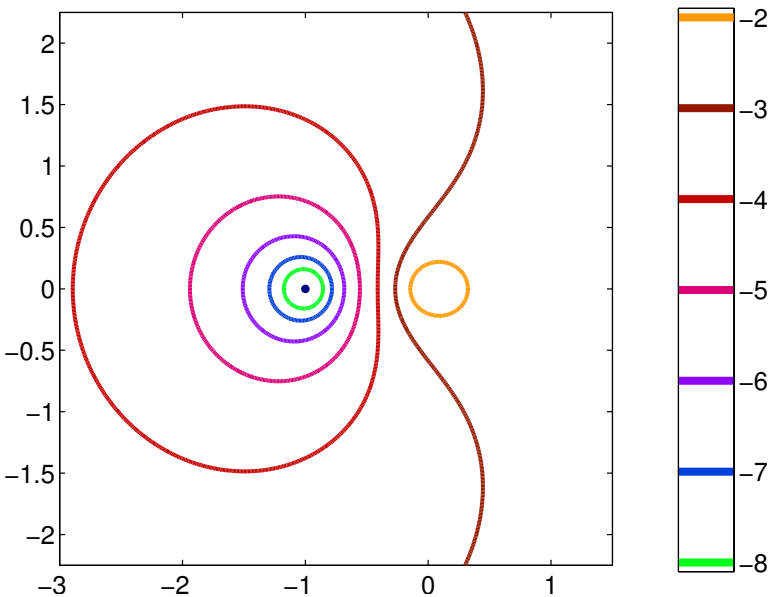
Complex and real pseudospectra for 5×5 matrix, $\varepsilon = 10^{-2}, 10^{-3}, \ldots, 10^{-8}$



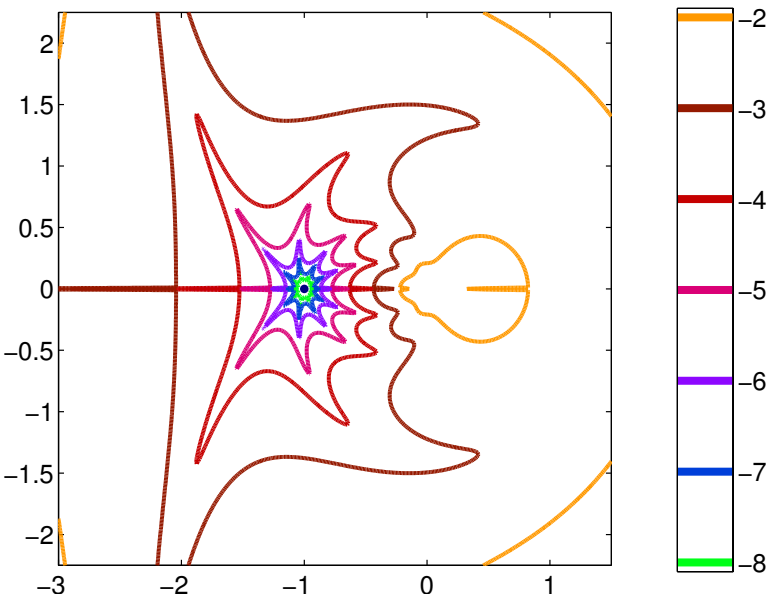
complex

Complex versus Real Perturbations: Example 1

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



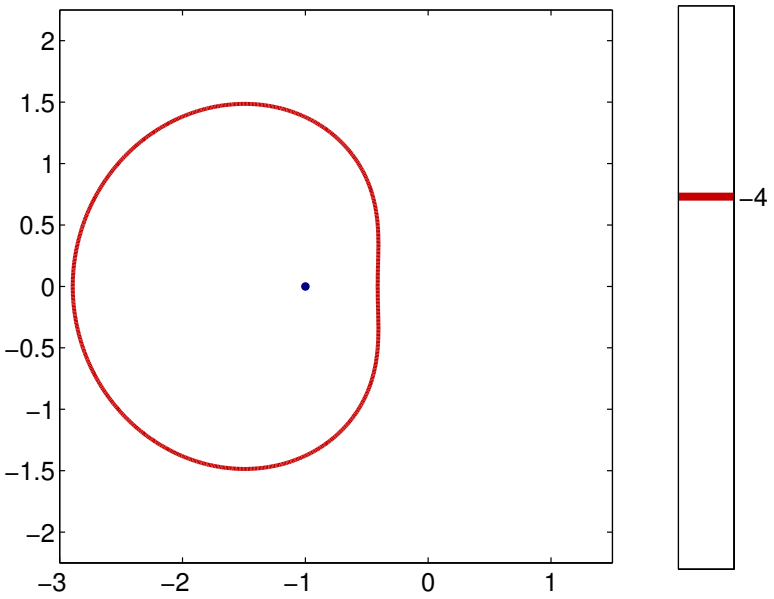
complex



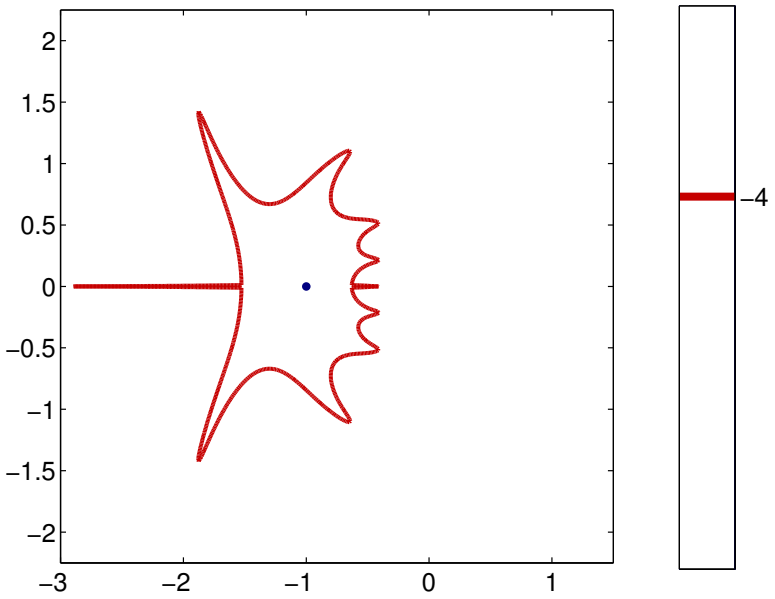
real

Complex versus Real Perturbations: Example 1

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



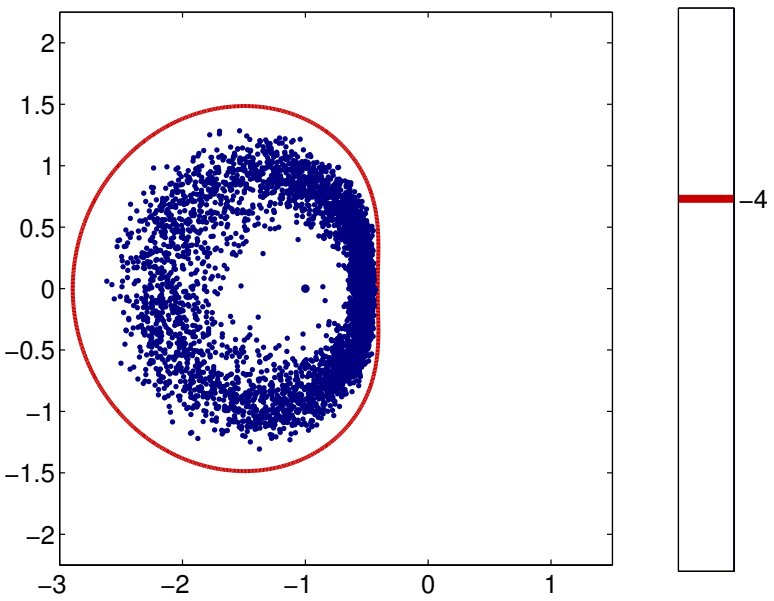
complex



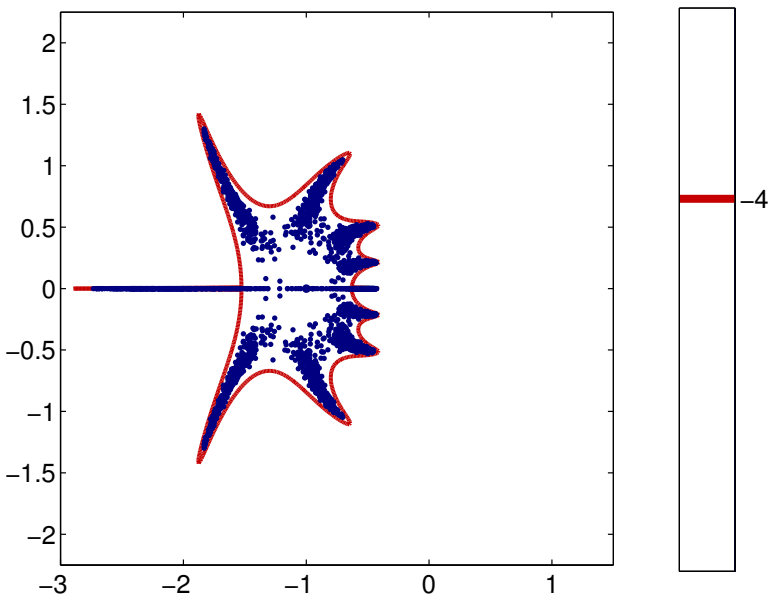
real

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

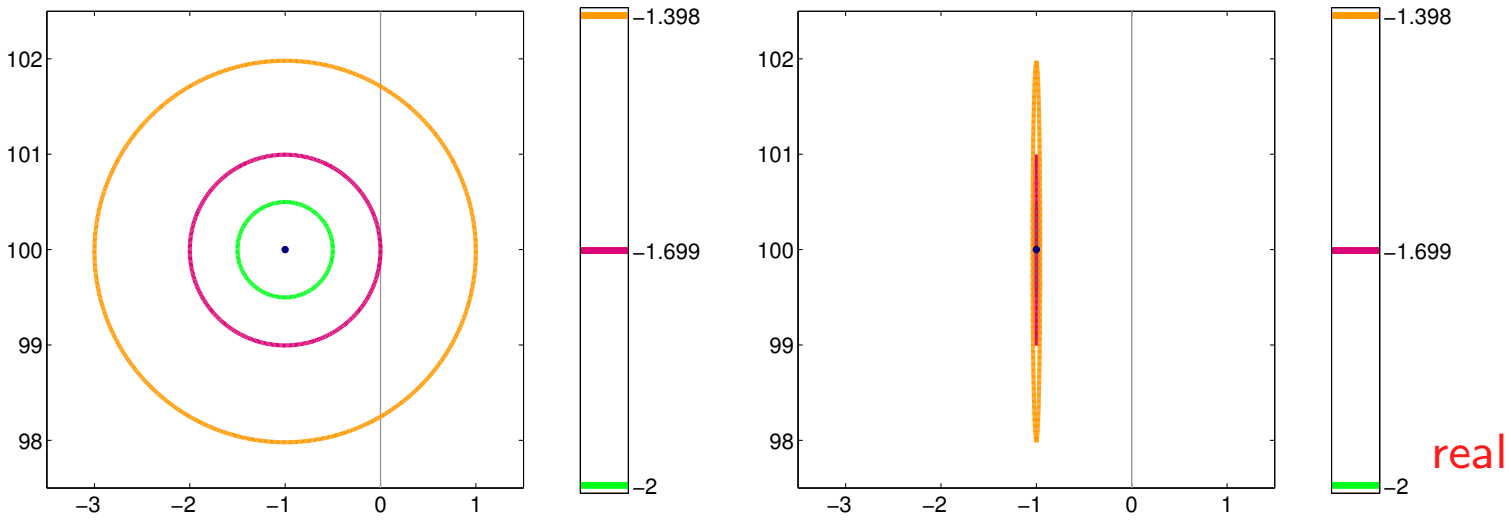
Real perturbations need not describe *transient* behavior of dynamical systems.

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

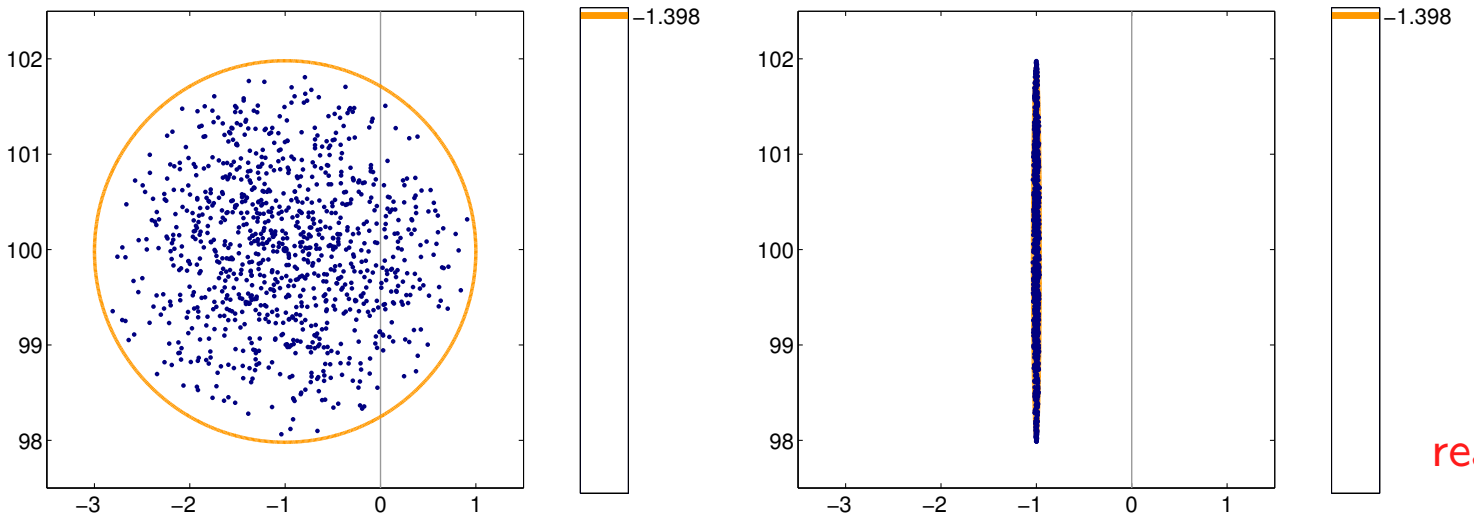
Real perturbations need not describe *transient* behavior of dynamical systems.

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

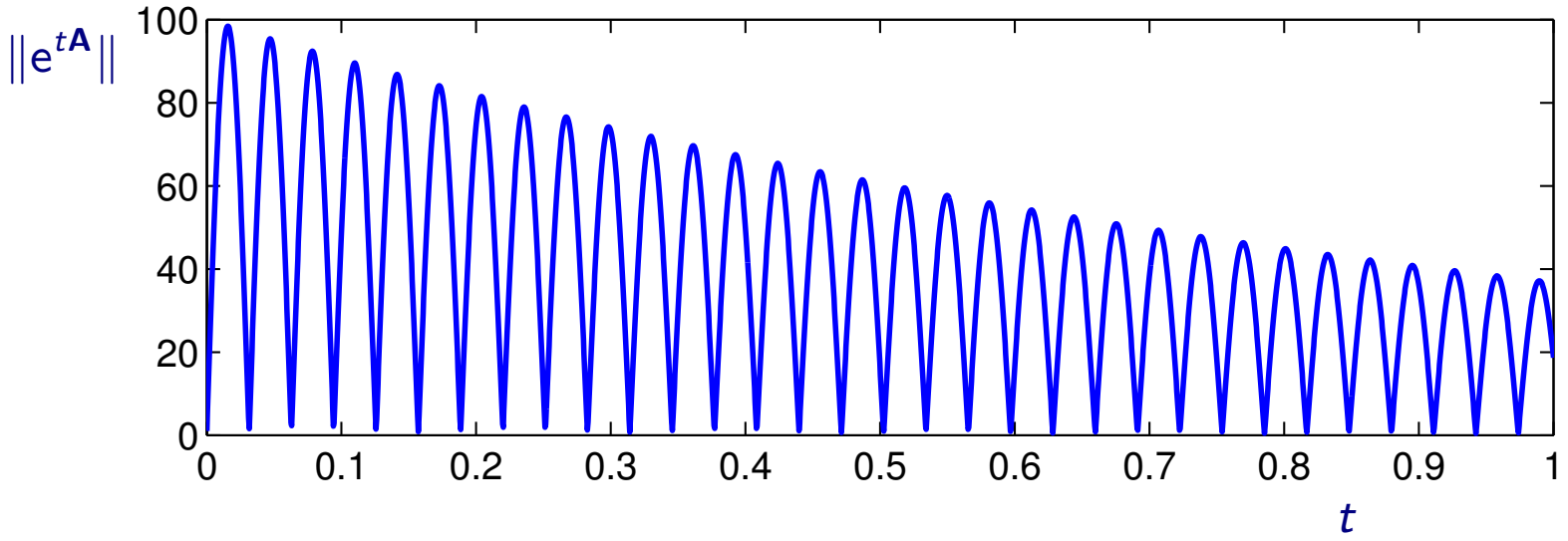
eigenvalues of 1000 random perturbations of size 400/10001

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.

5(c) Generalized Eigenvalue Problems / 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 $\alpha_1 = \|\mathbf{B}\|$.

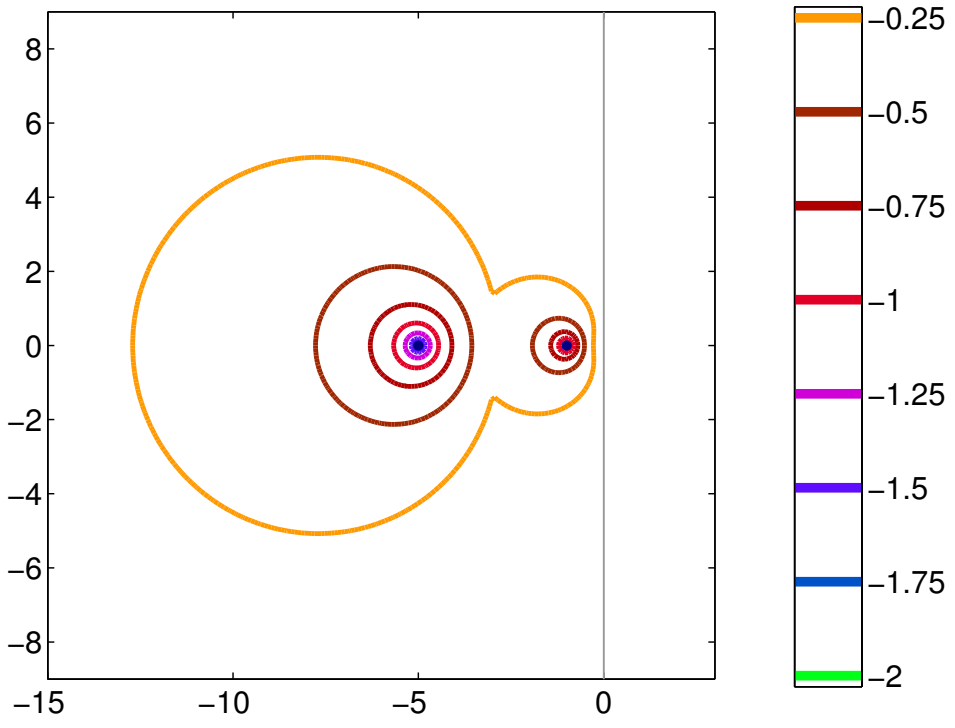
- This has an equivalent resolvent-like formulation:

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

- 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

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

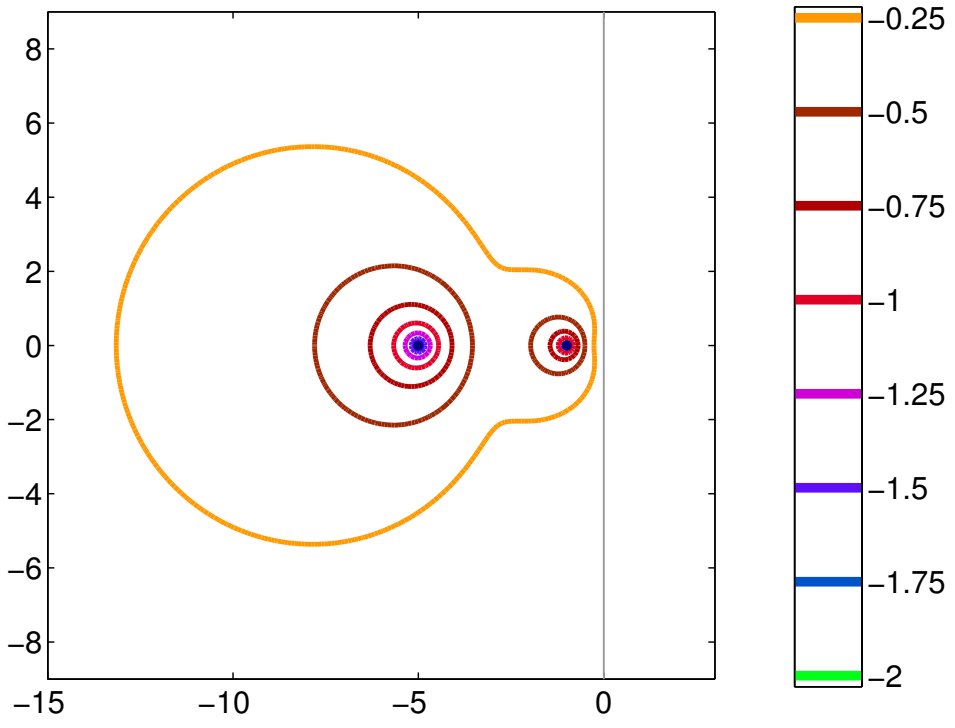


$M = 0$

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

GEPs: Eigenvalue Perturbation Approach

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

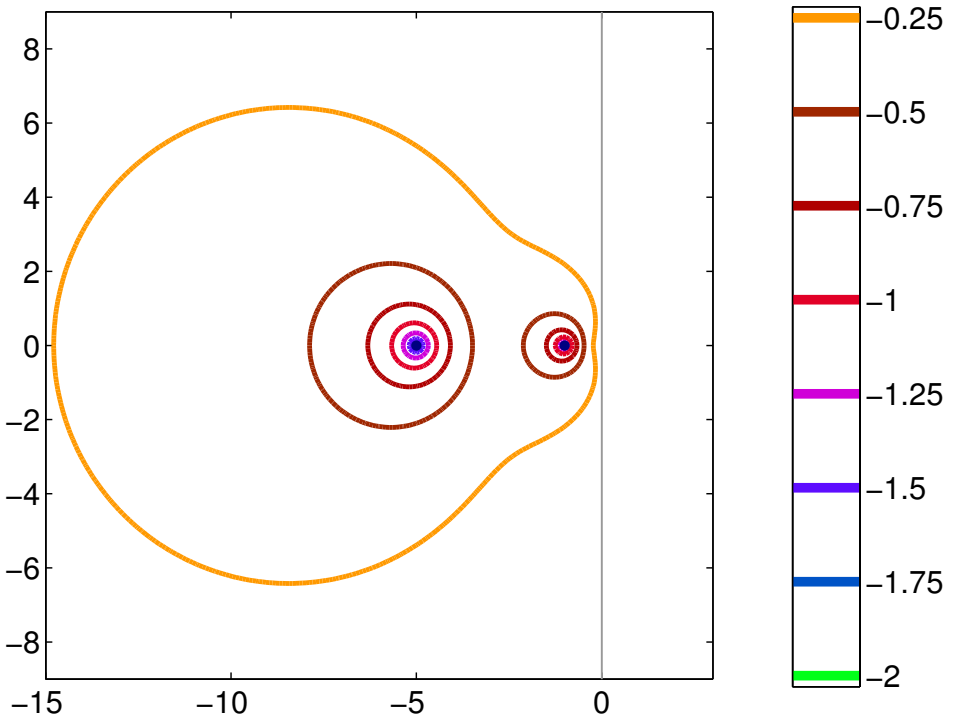


$M = 0.25$

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

GEPs: Eigenvalue Perturbation Approach

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

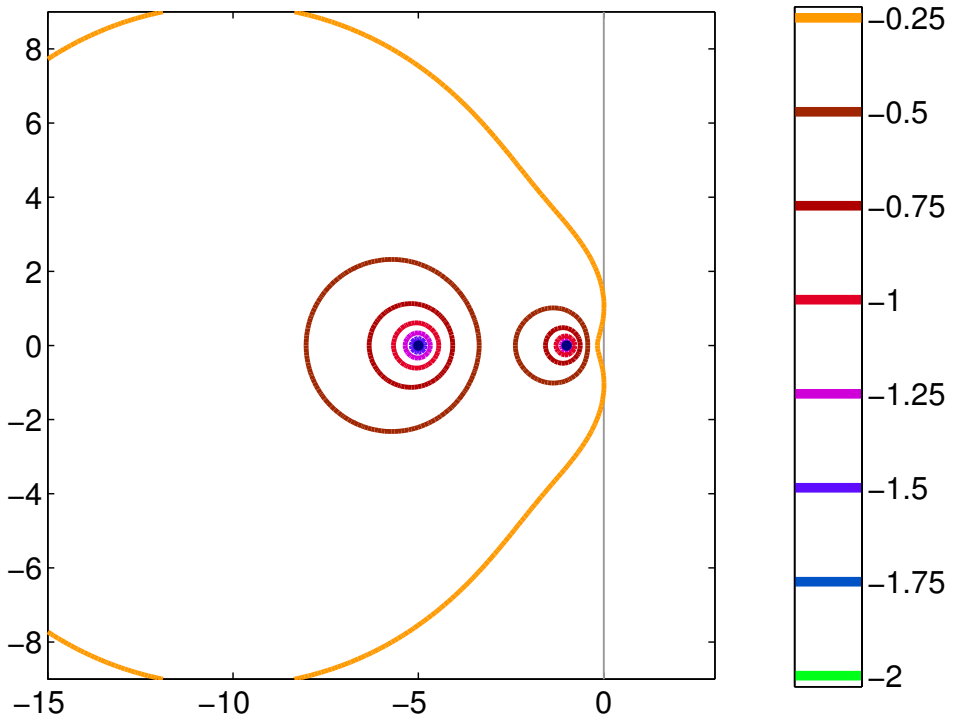


$M = 0.50$

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

GEPs: Eigenvalue Perturbation Approach

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

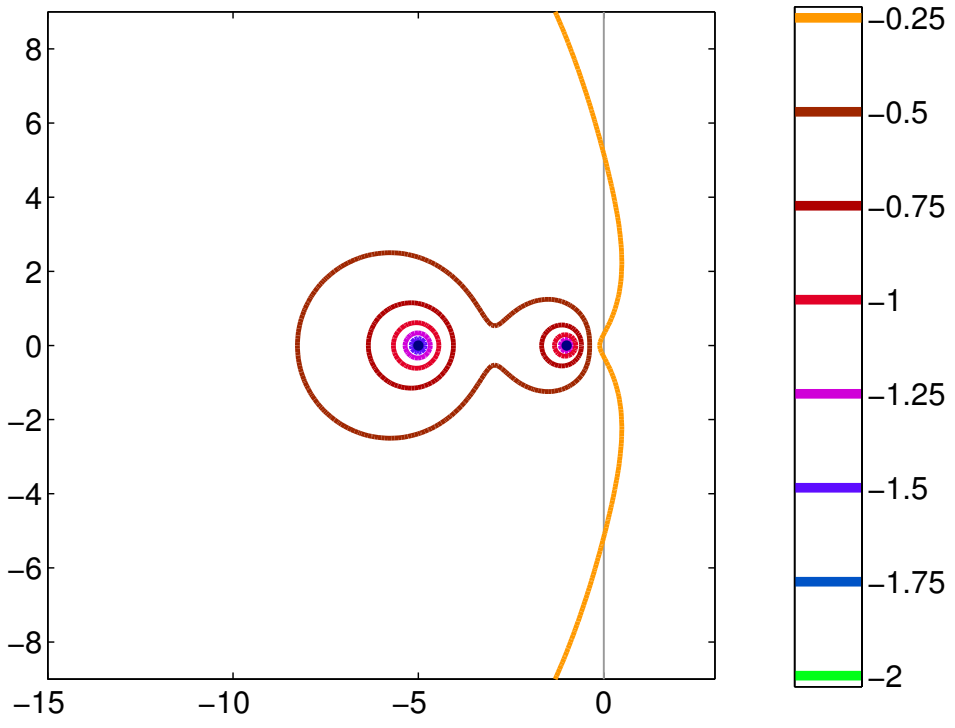


$M = 0.75$

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

GEPs: Eigenvalue Perturbation Approach

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

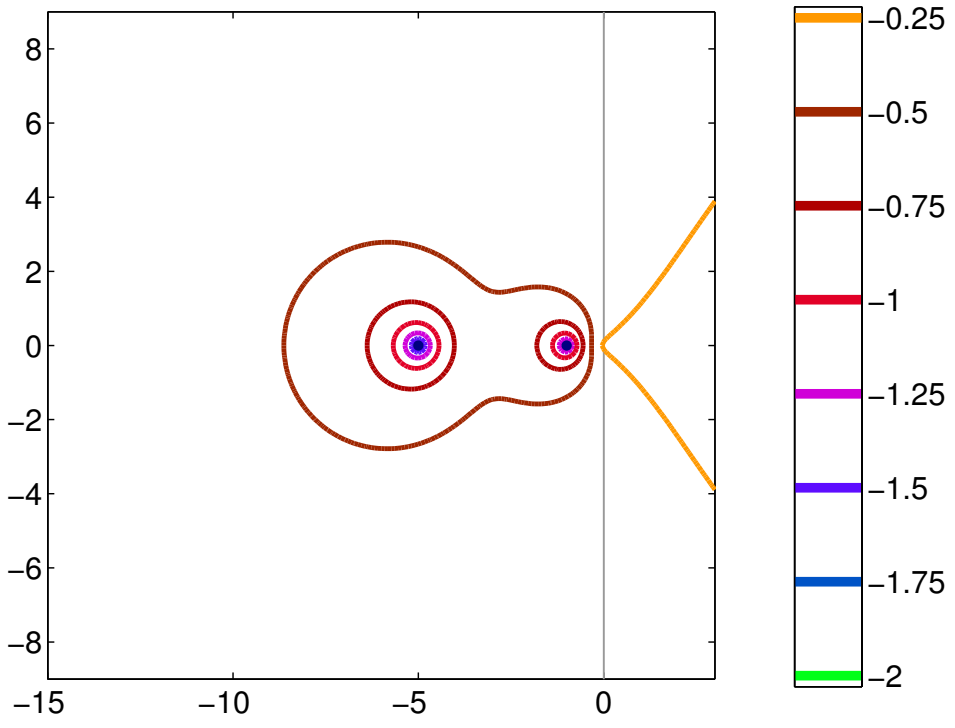


$M = 1.00$

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

GEPs: Eigenvalue Perturbation Approach

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

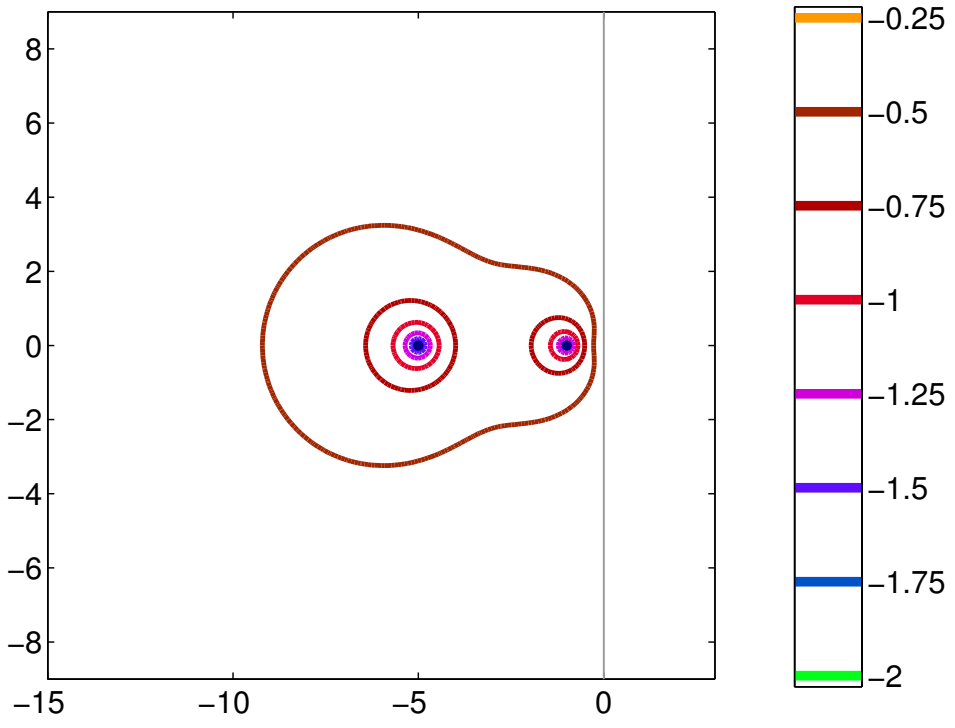


$M = 1.25$

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

GEPs: Eigenvalue Perturbation Approach

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

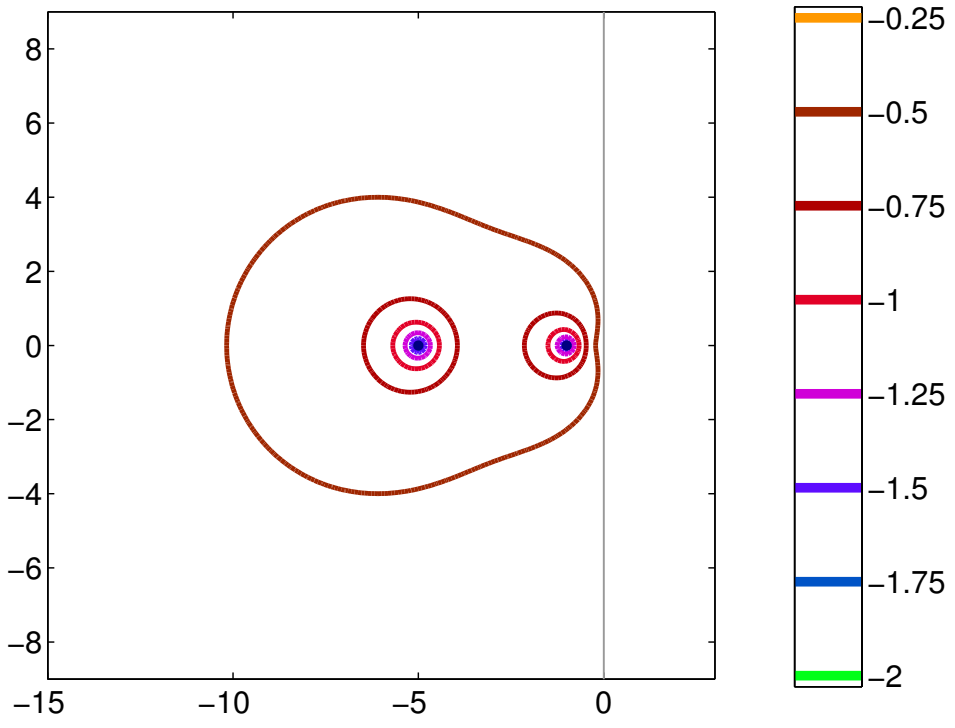


$M = 1.50$

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

GEPs: Eigenvalue Perturbation Approach

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

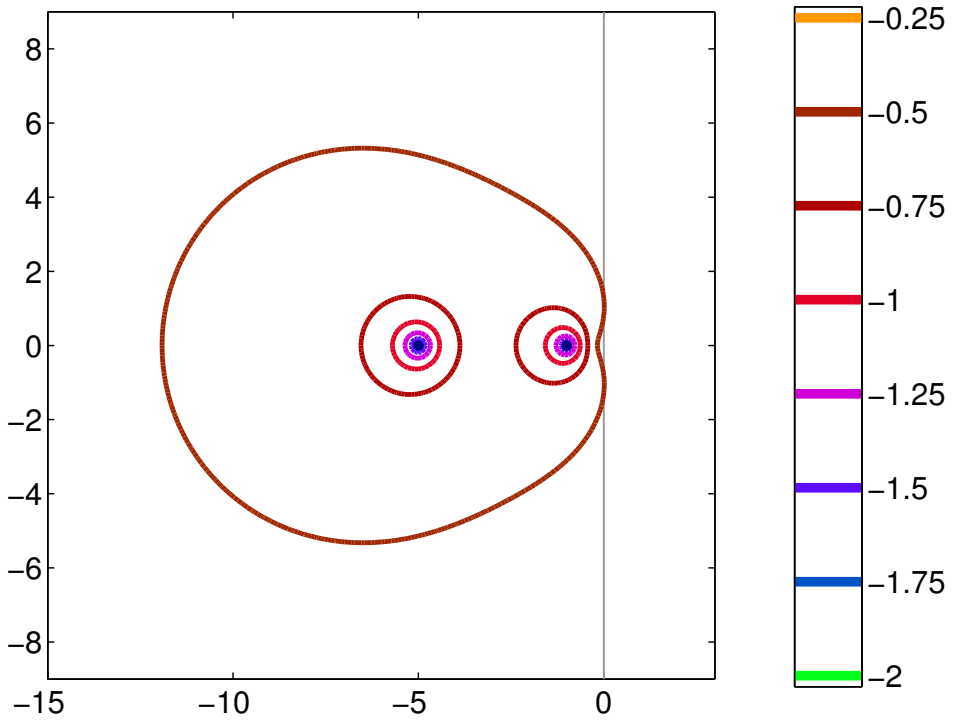


$M = 1.75$

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

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 $\mathbf{B}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$ 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 $\mathbf{B}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$ cannot exhibit growth.

GEPs: Eigenvalue Perturbation Approach

Consider solutions to $\mathbf{B}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$ 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 $\mathbf{B}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$ cannot exhibit growth.
- The parameter M affects the stability of eigenvalues of the pencil, but has no influence on the solution of $\mathbf{B}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$.

GEPs: Matrix Behavior Approach

- More generally, premultiplying

$$\mathbf{B}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$$

by some invertible matrix $\mathbf{S}$

$$\mathbf{S}\mathbf{B}\dot{\mathbf{x}}(t) = \mathbf{S}\mathbf{A}\mathbf{x}(t)$$

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

$$\mathbf{B}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$$

by some invertible matrix $\mathbf{S}$

$$\mathbf{SB}\dot{\mathbf{x}}(t) = \mathbf{SA}\mathbf{x}(t)$$

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

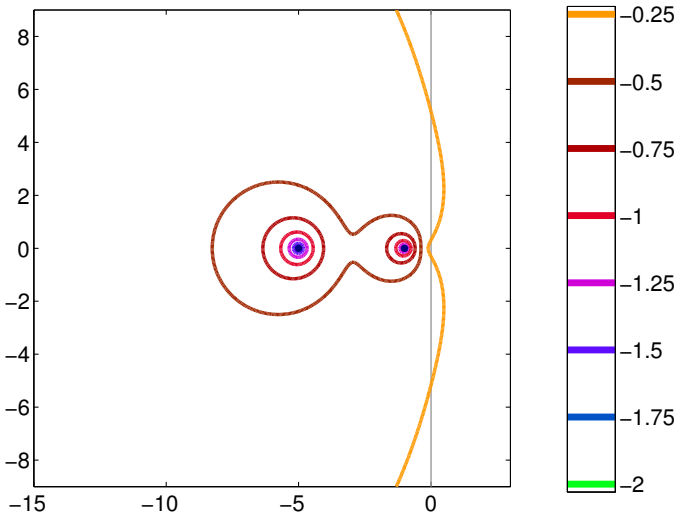
- [Ruhe, 1995], [Riedel, 1994] proposed:
For $\mathbf{A} \in \mathbf{C}^{n\times n}$ and invertible $\mathbf{B} \in \mathbf{C}^{n\times n}$,

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

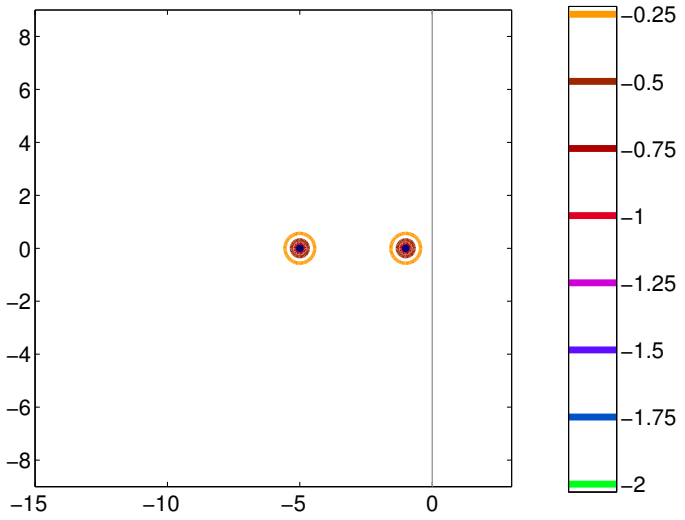
Comparison of GEP Pseudospectra

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

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



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 $z \in \mathbf{C}$).

$\mathbf{B}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$ is a *differential-algebraic equation* (DAE).

Simple example:

$$\begin{aligned} \dot{x}_1(t) &= -x_1(t) \\ x_1(t) &= x_2(t) \end{aligned}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

- 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

$$\mathbf{B}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$$

in the form

$$\mathbf{A}^{-1}\mathbf{B}\dot{\mathbf{x}}(t) = \mathbf{x}(t).$$

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:

$\mathbf{B}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$, $\mathbf{x}(0) = \mathbf{x}_0$ has a solution if and only if $\mathbf{x}_0 \in \text{Range}\,\mathbf{Q}_1$.

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:

$$\mathbf{G}\dot{\mathbf{y}}(t) = \mathbf{y}(t), \qquad \mathbf{z}(t) = \mathbf{0}.$$

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

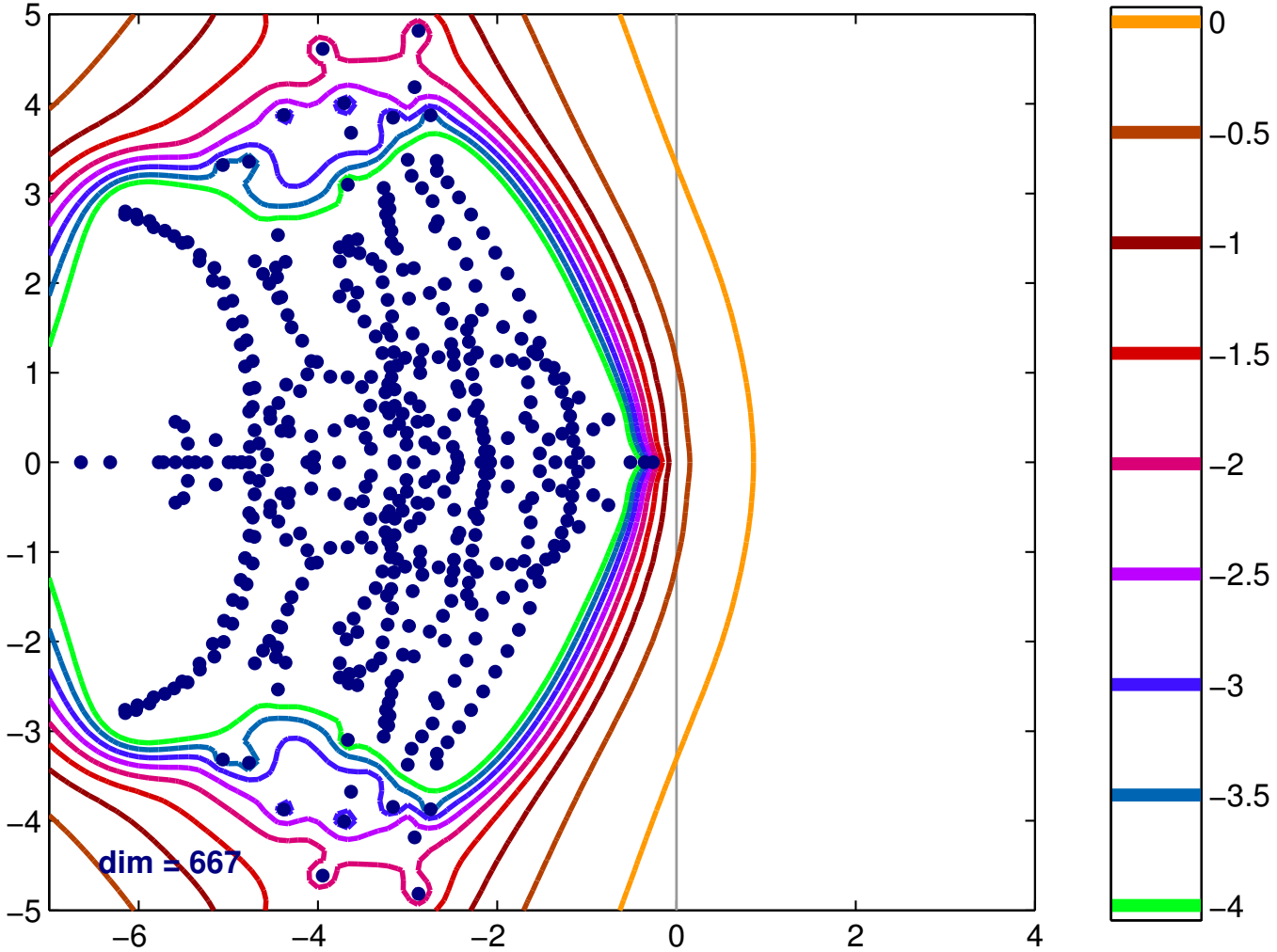
$$\mathbf{A}^{-1}\mathbf{B} = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{Q}_2 \end{bmatrix} \begin{bmatrix} \mathbf{G} & \mathbf{D} \\ \mathbf{0} & \mathbf{N} \end{bmatrix} \begin{bmatrix} \mathbf{Q}_1^* \\ \mathbf{Q}_2^* \end{bmatrix}$$

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 Incompressible Flow

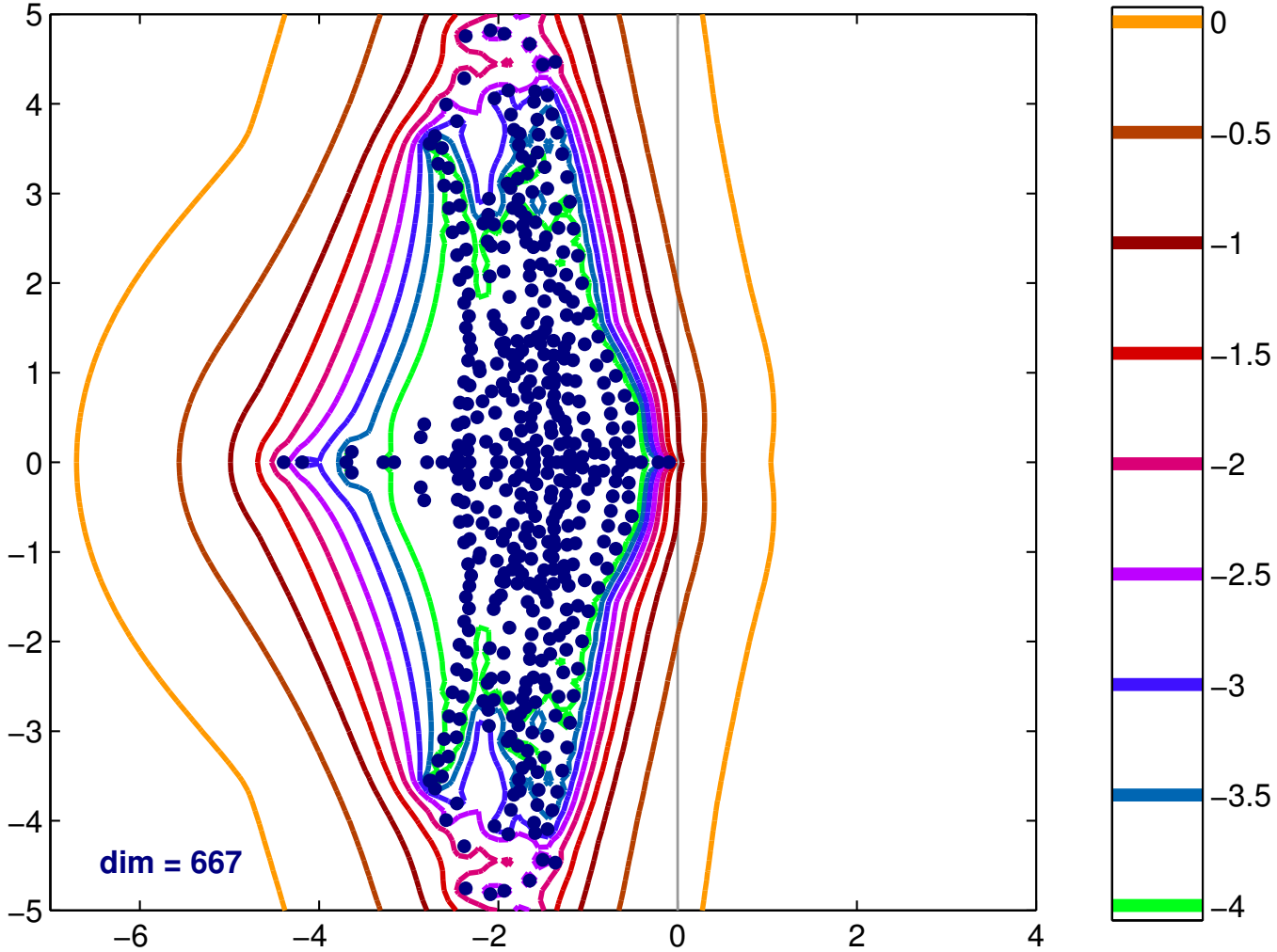
Pseudospectra for a matrix pencil derived from incompressible flow over a backward facing step, discretized via IFISS.



$1/\nu = 50$

Examples: Stability Analysis for Incompressible Flow

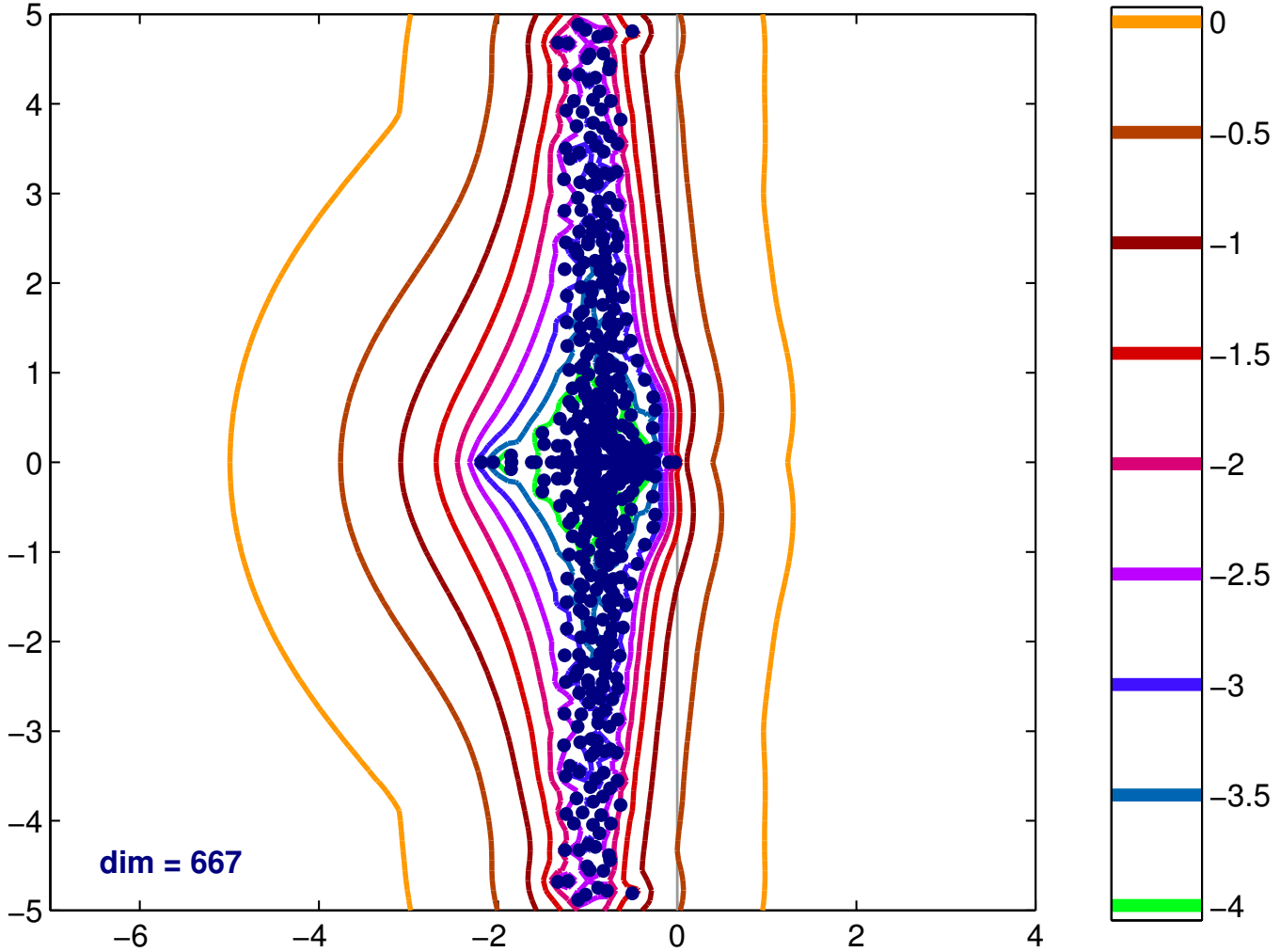
Pseudospectra for a matrix pencil derived from incompressible flow over a backward facing step, discretized via IFISS.



$1/\nu = 100$

Examples: Stability Analysis for Incompressible Flow

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



$1/\nu = 200$

5(d) Polynomial Eigenvalue Problems

Pseudospectra for Polynomial Eigenvalue Problems

Higher order problems give rise to *polynomial eigenvalue problems*.

For example, with the second-order ODE

$$\mathbf{M}\ddot{\mathbf{u}}(t) + \mathbf{G}\dot{\mathbf{u}}(t) + \mathbf{K}\mathbf{u}(t) = \mathbf{0}$$

we associate the *quadratic eigenvalue problem*

$$(\lambda^2\mathbf{M} + \lambda\mathbf{G} + \mathbf{K})\mathbf{u} = \mathbf{0}.$$

A typical numerical approach is to "linearize" this equation to obtain a generalized eigenvalue problem: introduce

$$\mathbf{v} := \lambda\mathbf{u},$$

so that

$$\lambda\mathbf{M}\mathbf{v} + \mathbf{G}\mathbf{v} + \mathbf{K}\mathbf{u} = \mathbf{0}.$$

Pseudospectra for Polynomial Eigenvalue Problems

Higher order problems give rise to *polynomial eigenvalue problems*.

For example, with the second-order ODE

$$\mathbf{M}\ddot{\mathbf{u}}(t) + \mathbf{G}\dot{\mathbf{u}}(t) + \mathbf{K}\mathbf{u}(t) = \mathbf{0}$$

we associate the *quadratic eigenvalue problem*

$$(\lambda^2\mathbf{M} + \lambda\mathbf{G} + \mathbf{K})\mathbf{u} = \mathbf{0}.$$

A typical numerical approach is to "linearize" this equation to obtain a generalized eigenvalue problem: introduce

$$\mathbf{v} := \lambda\mathbf{u},$$

so that

$$\lambda\mathbf{M}\mathbf{v} + \mathbf{G}\mathbf{v} + \mathbf{K}\mathbf{u} = \mathbf{0}.$$

Together, we have $\mathbf{A}\mathbf{x} = \lambda\mathbf{B}\mathbf{x}$ of the form

$$\begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{K} & -\mathbf{G} \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix} = \lambda \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{M} \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix}$$

Pseudospectra for Polynomial Eigenvalue Problems

By transforming the quadratic problem

$$(\lambda^2\mathbf{M} + \lambda\mathbf{G} + \mathbf{K})\mathbf{u} = \mathbf{0}.$$

into the form $\mathbf{Ax} = \lambda\mathbf{Bx}$, i.e.,

$$\begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{K} & -\mathbf{G} \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix} = \lambda \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{M} \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix}$$

we can perform pseudospectral analysis on the generalized eigenvalue problem.

However: *Generic perturbations* $\mathbf{A} + \mathbf{E}_0$ *and* $\mathbf{B} + \mathbf{E}_1$ *will destroy the structure that derives from the second-order form.*

Pseudospectra for Polynomial Eigenvalue Problems

By transforming the quadratic problem

$$(\lambda^2\mathbf{M} + \lambda\mathbf{G} + \mathbf{K})\mathbf{u} = \mathbf{0}.$$

into the form $\mathbf{Ax} = \lambda\mathbf{Bx}$, i.e.,

$$\begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{K} & -\mathbf{G} \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix} = \lambda \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{M} \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix}$$

we can perform pseudospectral analysis on the generalized eigenvalue problem.

However: *Generic perturbations* $\mathbf{A} + \mathbf{E}_0$ *and* $\mathbf{B} + \mathbf{E}_1$ *will destroy the structure that derives from the second-order form.*

Definition (Tisseur & Higham, 2001)

The *ε-pseudospectrum* $\sigma_\varepsilon(\mathbf{P})$ *of the matrix polynomial* $\mathbf{P}(\lambda) = \mathbf{A}_0 + \mathbf{A}_1\lambda + \cdots + \lambda^d\mathbf{A}_d$ is the set

$$\begin{aligned} \sigma_\varepsilon(\mathbf{P}) \quad = \quad & \{z \in \mathbf{C} : z \in \sigma(\mathbf{P} + \mathbf{E}) \text{ for some} \\ & \mathbf{E}(\lambda) = \mathbf{E}_0 + \lambda\mathbf{E}_1 + \cdots + \lambda^d\mathbf{E}_d,\ \|\mathbf{E}_j\| \le \varepsilon\alpha_j, \quad j = 0, \ldots, d\}, \end{aligned}$$

for nonnegative $\alpha_0, \ldots, \alpha_d$.

For most applications, one would either take $\alpha_j = 1$ for all j, or $\alpha_j = \|A_j\|$.

Pseudospectra for Polynomial Eigenvalue Problems

Definition (Tisseur & Higham, 2001)

The *ε-pseudospectrum $\sigma_\varepsilon(\mathbf{P})$ of the matrix polynomial* $\mathbf{P}(\lambda) = \mathbf{A}_0 + \mathbf{A}_1\lambda + \cdots + \lambda^d\mathbf{A}_d$ is the set

$$\begin{aligned}\sigma_\varepsilon(\mathbf{P}) \quad = \quad & \{z \in \mathbf{C} : z \in \sigma(\mathbf{P}+\mathbf{E}) \text{ for some} \\ & \mathbf{E}(\lambda) = \mathbf{E}_0 + \lambda\mathbf{E}_1 + \cdots + \lambda^d\mathbf{E}_d,\ \|\mathbf{E}_j\| \le \varepsilon\alpha_j, \quad j = 0,\ldots,d\},\end{aligned}$$

for nonnegative $\alpha_0, \ldots, \alpha_d$.

For most applications, one would either take $\alpha_j = 1$ for all j, or $\alpha_j = \|A_j\|$.

This definition restricts the perturbed problem to have the same structure as the unperturbed problem.

$$\begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{K}-\mathbf{E}_0 & -\mathbf{G}-\mathbf{E}_1 \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix} = \lambda \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}+\mathbf{E}_2 \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix}$$

The resulting spectrum of each of these perturbations will be the spectrum associated with a second-order system.

Pseudospectra for Polynomial Eigenvalue Problems

Definition (Tisseur & Higham, 2001)

The *ε-pseudospectrum $\sigma_\varepsilon(\mathbf{P})$ of the matrix polynomial* $\mathbf{P}(\lambda) = \mathbf{A}_0 + \mathbf{A}_1\lambda + \cdots + \lambda^d\mathbf{A}_d$ is the set

$$\begin{aligned}\sigma_\varepsilon(\mathbf{P}) \quad = \quad & \{z \in \mathbf{C} : z \in \sigma(\mathbf{P}+\mathbf{E}) \text{ for some} \\ & \mathbf{E}(\lambda) = \mathbf{E}_0 + \lambda\mathbf{E}_1 + \cdots + \lambda^d\mathbf{E}_d,\ \|\mathbf{E}_j\| \le \varepsilon\alpha_j, \quad j = 0,\ldots,d\},\end{aligned}$$

for nonnegative $\alpha_0, \ldots, \alpha_d$.

There is also an equivalent formulation in terms of resolvent norms.

Theorem (Tisseur & Higham, 2001)

The ε-pseudospectrum $\sigma_\varepsilon(\mathbf{P})$ of the matrix polynomial can equivalently be characterized as

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

where $\phi(z) = \sum_{j=0}^{d} \alpha_k z^k$ for the same values of $\alpha_0, \ldots, \alpha_d$ used earlier.

Pseudospectra for Polynomial Eigenvalue Problems

From the perspective of *eigenvalue computations*, this approach is very helpful.

- Subsequent work by Higham, Tisseur, Mehl, Mehrmann, Mackey, and Mackey investigates the eigenvalue conditioning of alternative linearizations of the polynomial problem.

- This work aims to pick the linearization that gives the most stable eigenvalues, to obtain accurately computed eigenvalues.

Pseudospectra for Polynomial Eigenvalue Problems

From the perspective of *eigenvalue computations*, this approach is very helpful.

- Subsequent work by Higham, Tisseur, Mehl, Mehrmann, Mackey, and Mackey investigates the eigenvalue conditioning of alternative linearizations of the polynomial problem.

- This work aims to pick the linearization that gives the most stable eigenvalues, to obtain accurately computed eigenvalues.

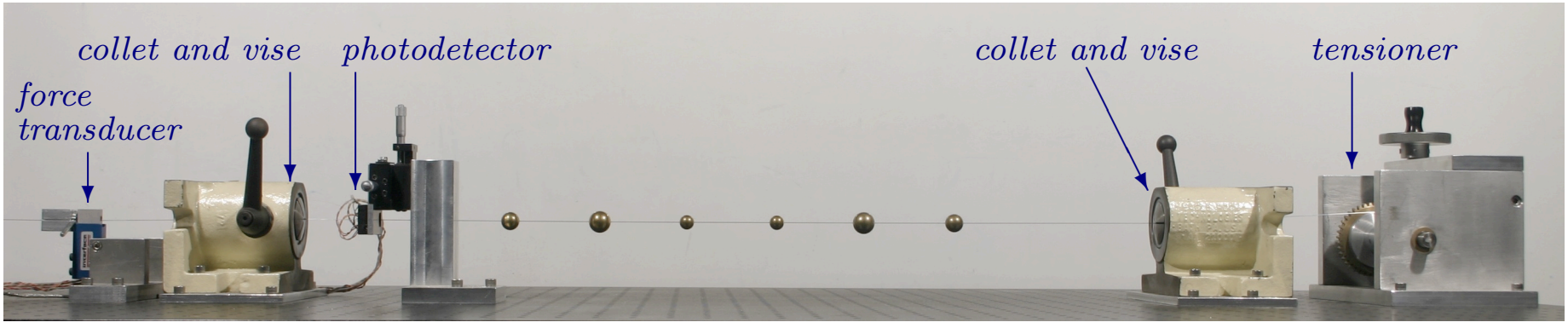
- A different perspective: If we are concerned with understanding the *behavior* of the underlying dynamical system, we *want to see and understand physically-relevant instability* that corresponds to the transient behavior of our system. One can assess such growth through using the conventional pseudospectrum of the linearization, measured in the energy norm for the system.
- We shall investigate this idea using discretizations of a damped wave equation:

$$u_{tt}(x,t) = u_{xx}(x,t) - 2a(x)u_t(x,t),$$

for $x \in [0,1]$ with $u(0) = u(1) = 0$.

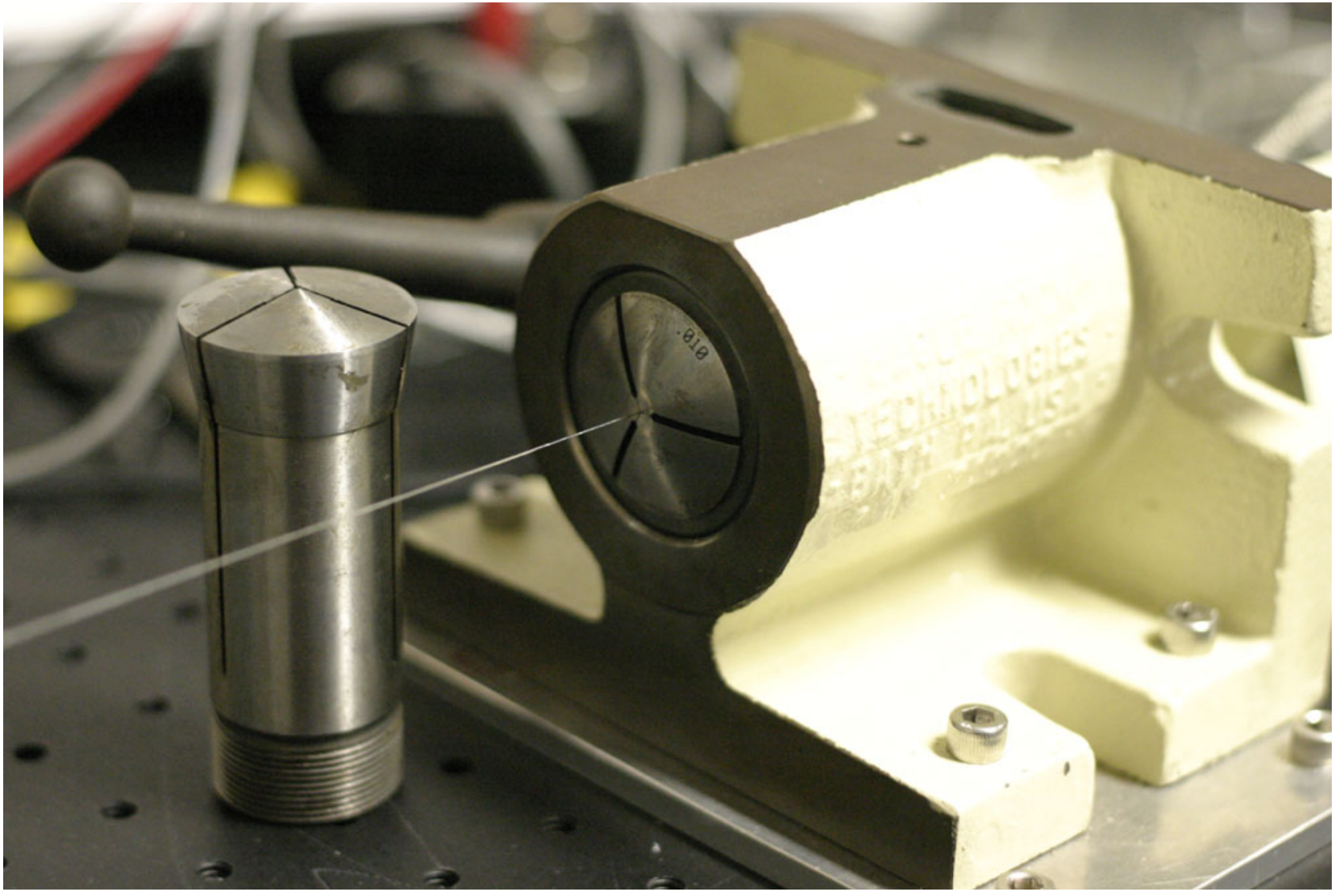
Physical Experiments

We have a small laboratory in which to conduct experiments with these (inverse) eigenvalue problems.

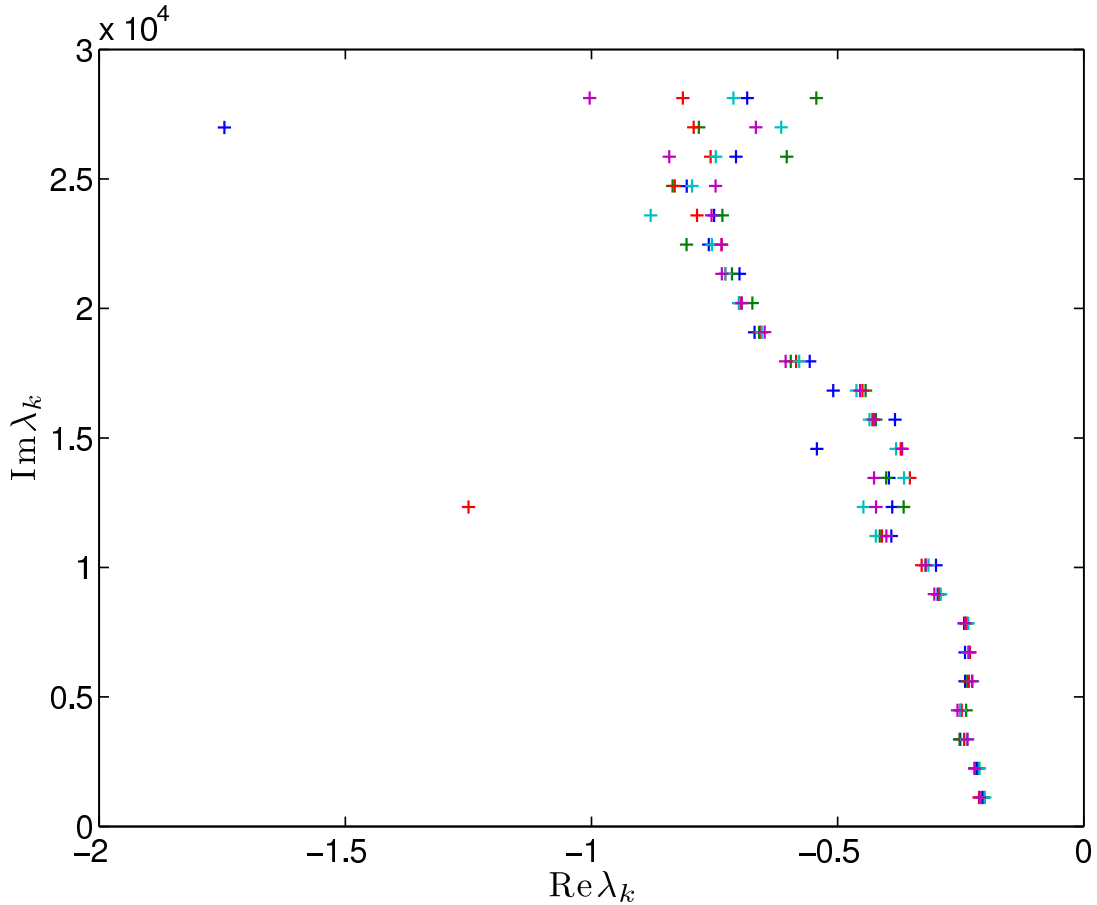


Apparatus inspired by string designer Fan Tao of J. d'Addario;
Designed by Sean Hardesty, Jeffrey Hokanson.

Dirichlet boundary conditions imposed via a collet



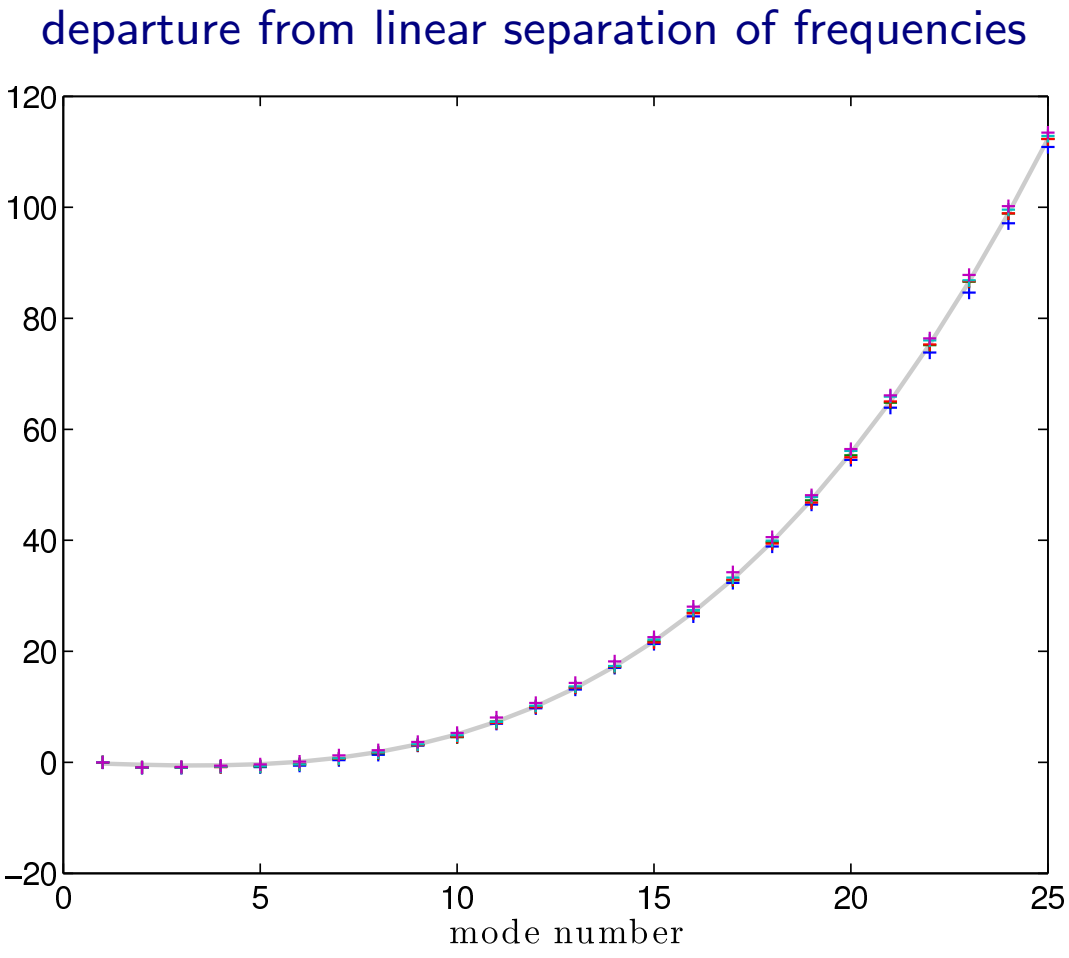
Measured eigenvalues of a real string



Measured eigenvalues of a homogeneous steel wire, driven at frequencies; results from five trials [Hokanson, 2011].

Eigenvalues determined by Hokanson's fast variant of VARPRO.
Frequencies (imaginary parts) are easy to find; real parts are more difficult.

Measured eigenvalues of a real string



Quadratic separation is a characteristic of stiff strings [Hokanson, 2011]; cf. Bensa et al. [2003].

Canonical damping mechanisms

We can consider more exotic classes of damping functions.

viscous damping:	$u_{tt} = u_{xx} - 2a(x)u_t$
Kelvin–Voigt:	$u_{tt} = u_{xx} + (a(x)u_{xt})_x$
magnetic damping:	$u_{tt} = u_{xx} - a(x)\int_0^1 a(s)u_t(s,t)\,\mathrm{d}s$
stiff strings:	$u_{tt} = c^2 u_{xx} - \kappa^2 u_{xxxx} - 2a(x)u_t + 2b(x)u_{xxt}$

Each is posed on $x \in [0,1]$, $t \geq 0$, with Dirichlet boundary conditions $u(0,t) = u(1,t) = 0$ and initial conditions

$$u(x,0) = u_0(x), \qquad u_t(x,0) = v_0(x).$$

Canonical damping mechanisms

We can consider more exotic classes of damping functions.

viscous damping:	$u_{tt} = u_{xx} - 2a(x)u_t$
Kelvin–Voigt:	$u_{tt} = u_{xx} + (a(x)u_{xt})_x$
magnetic damping:	$u_{tt} = u_{xx} - a(x)\int_0^1 a(s)u_t(s,t)\,\mathrm{d}s$
stiff strings:	$u_{tt} = c^2 u_{xx} - \kappa^2 u_{xxxx} - 2a(x)u_t + 2b(x)u_{xxt}$

Each is posed on $x \in [0,1]$, $t \geq 0$, with Dirichlet boundary conditions $u(0,t) = u(1,t) = 0$ and initial conditions

$$u(x,0) = u_0(x), \qquad u_t(x,0) = v_0(x).$$

These mechanisms are more sophisticated that "Rayleigh damping" that is often seen in computational codes, where the eigenvalues of the damped problem are a simple function of the eigenvalues of the undamped problem.

General framework

To analyze, we write these operators as first-order time-evolution problems, (mostly) of the form

$$\begin{bmatrix} u \\ u_t \end{bmatrix}_t = \begin{bmatrix} 0 & I \\ \mathrm{d}^2/\mathrm{d}x^2 & G \end{bmatrix} \begin{bmatrix} u \\ u_t \end{bmatrix},$$

where G denotes the damping operator. We rewrite as

$$U_t = AU,$$

and are interested in the decay of the semigroup A generates. Here (usually)

$$U = \begin{bmatrix} u \\ v \end{bmatrix} \in \mathrm{Dom}(A) = (H_0^1(0,1) \cap H^2(0,1)) \times H_0^1(0,1)$$

with energy inner product

$$\left\langle \begin{bmatrix} f \\ g \end{bmatrix}, \begin{bmatrix} u \\ v \end{bmatrix} \right\rangle_E = \int_0^1 \left(\overline{f'(x)} u'(x) + \overline{g(x)} v(x) \right) \mathrm{d}x.$$

General framework

To analyze, we write these operators as first-order time-evolution problems, (mostly) of the form

$$\begin{bmatrix} u \\ u_t \end{bmatrix}_t = \begin{bmatrix} 0 & I \\ \mathrm{d}^2/\mathrm{d}x^2 & G \end{bmatrix} \begin{bmatrix} u \\ u_t \end{bmatrix},$$

where G denotes the damping operator. We rewrite as

$$U_t = AU,$$

and are interested in the decay of the semigroup A generates. Here (usually)

$$U = \begin{bmatrix} u \\ v \end{bmatrix} \in \mathrm{Dom}(A) = (H_0^1(0,1) \cap H^2(0,1)) \times H_0^1(0,1)$$

with energy inner product

$$\left\langle \begin{bmatrix} f \\ g \end{bmatrix}, \begin{bmatrix} u \\ v \end{bmatrix} \right\rangle_E = \int_0^1 \left(\overline{f'(x)} u'(x) + \overline{g(x)} v(x) \right) \mathrm{d}x.$$

Discretized computations should be faithful to this inner product.

Discretization

With finite elements,

$$\begin{bmatrix} u \\ u_t \end{bmatrix}_t = \begin{bmatrix} 0 & I \\ \mathrm{d}^2/\mathrm{d}x^2 & G \end{bmatrix} \begin{bmatrix} u \\ u_t \end{bmatrix}$$

leads to a system of the form

$$\begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{M} \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix}_t = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{K} & -\mathbf{G} \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix}.$$

Discretization

With finite elements,

$$\begin{bmatrix} u \\ u_t \end{bmatrix}_t = \begin{bmatrix} 0 & I \\ \mathrm{d}^2/\mathrm{d}x^2 & G \end{bmatrix} \begin{bmatrix} u \\ u_t \end{bmatrix}$$

leads to a system of the form

$$\begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{M} \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix}_t = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ \underbrace{-\mathbf{K}} & -\mathbf{G} \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix}.$$

SPD SPD structured

Discretization

With finite elements,

$$\begin{bmatrix} u \\ u_t \end{bmatrix}_t = \begin{bmatrix} 0 & I \\ \mathrm{d}^2/\mathrm{d}x^2 & G \end{bmatrix} \begin{bmatrix} u \\ u_t \end{bmatrix}$$

leads to a system of the form

$$\begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{M} \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix}_t = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{K} & -\mathbf{G} \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix}.$$

SPD SPD structured

When G is well-behaved, we prefer a Chebyshev spectral discretization, where $\mathbf{M} = \mathbf{I}$ and $\mathbf{K}$ is only "spiritually SPD".

Undamped problem

When $G = 0$, we have the undamped wave operator

$$A = \begin{bmatrix} 0 & I \\ \mathrm{d}^2/\mathrm{d}x^2 & 0 \end{bmatrix},$$

which is skew-adjoint in the energy inner product:

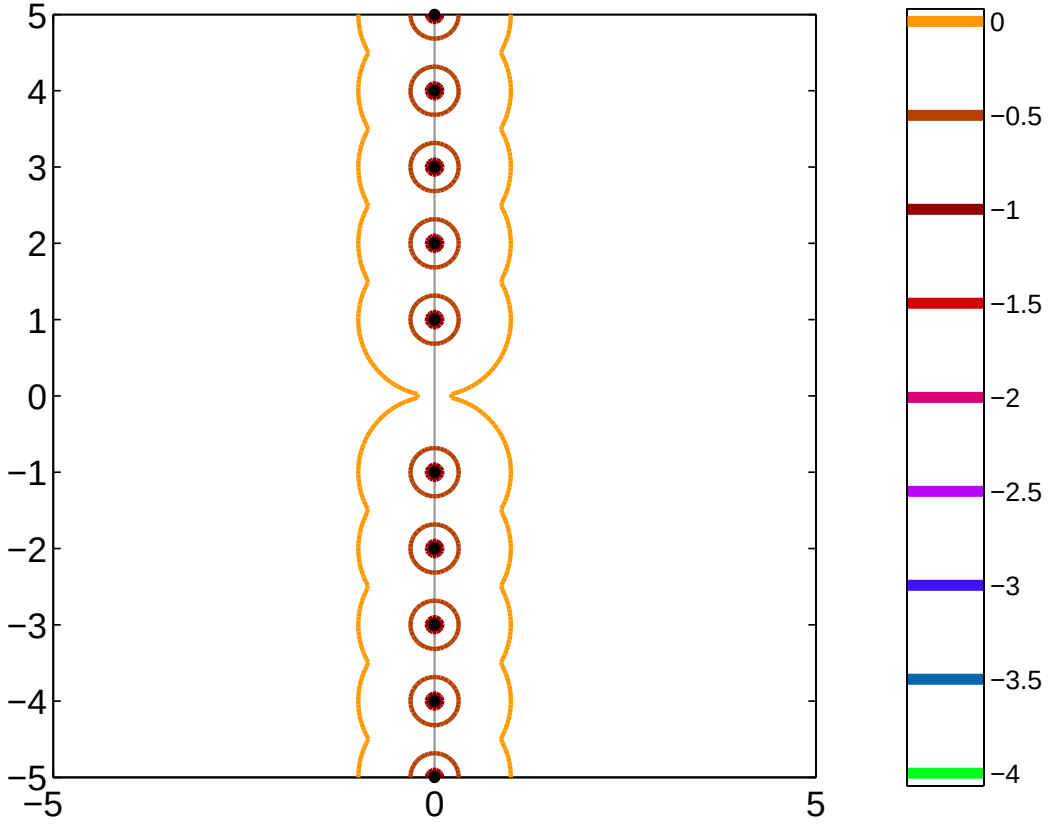
$$A^* = \begin{bmatrix} 0 & -I \\ -\mathrm{d}^2/\mathrm{d}x^2 & 0 \end{bmatrix}.$$

Consequences:

- eigenvalues are purely imaginary: $\sigma(A) = \{\pm k\pi \mathrm{i} : k = 1, 2, \ldots\}$;
- eigenvectors are $U_{\pm k} = \begin{bmatrix} \sin(k\pi x) \\ \pm k\pi \mathrm{i} \sin(k\pi x) \end{bmatrix}$;
- operator is normal;
- $\sigma_\varepsilon(A)$ comprises the union of open ε-balls about each of the eigenvalues.
- energy of solutions to $U_t = AU$ is conserved: $\|U(t)\| = \|U(0)\|$.

Undamped problem

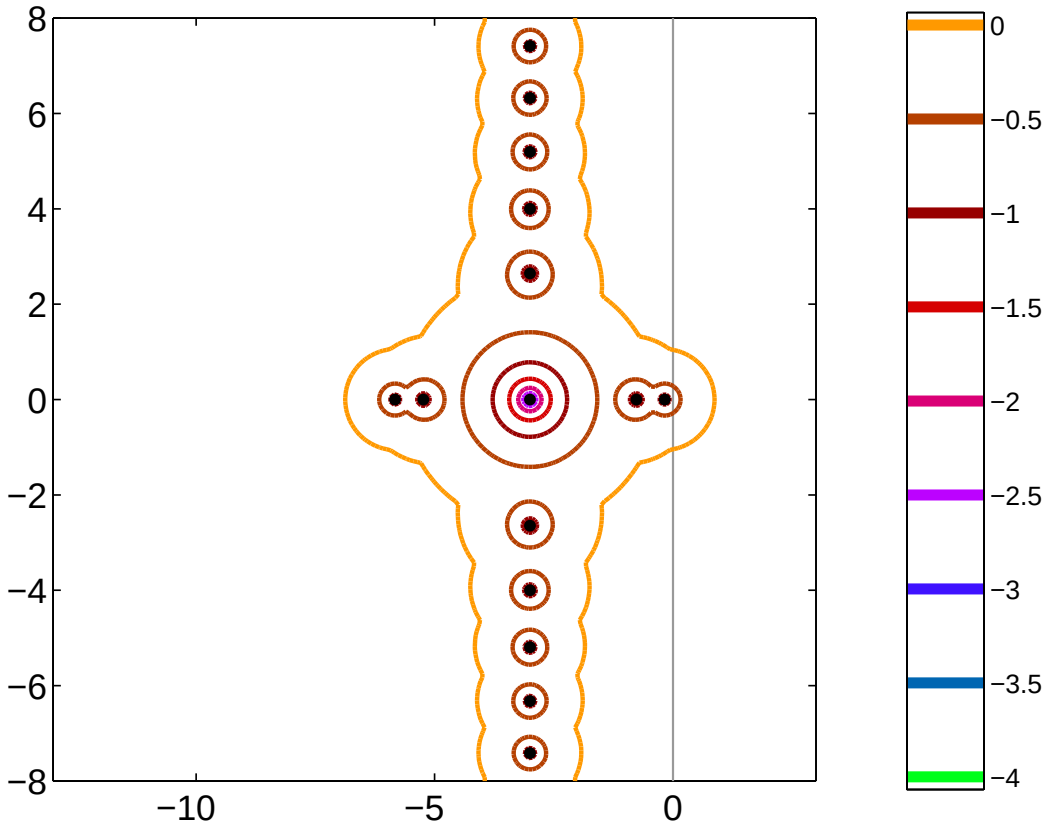
Pseudospectra (posed on $[0, \pi]$, so eigenvalues are $\lambda_{\pm k} = \pm k\,i$)



Since A is normal, pseudospectra are the union of ε-balls about the eigenvalues.

Viscous damping: $G(x)v(x) = -2a(x)v(x)$

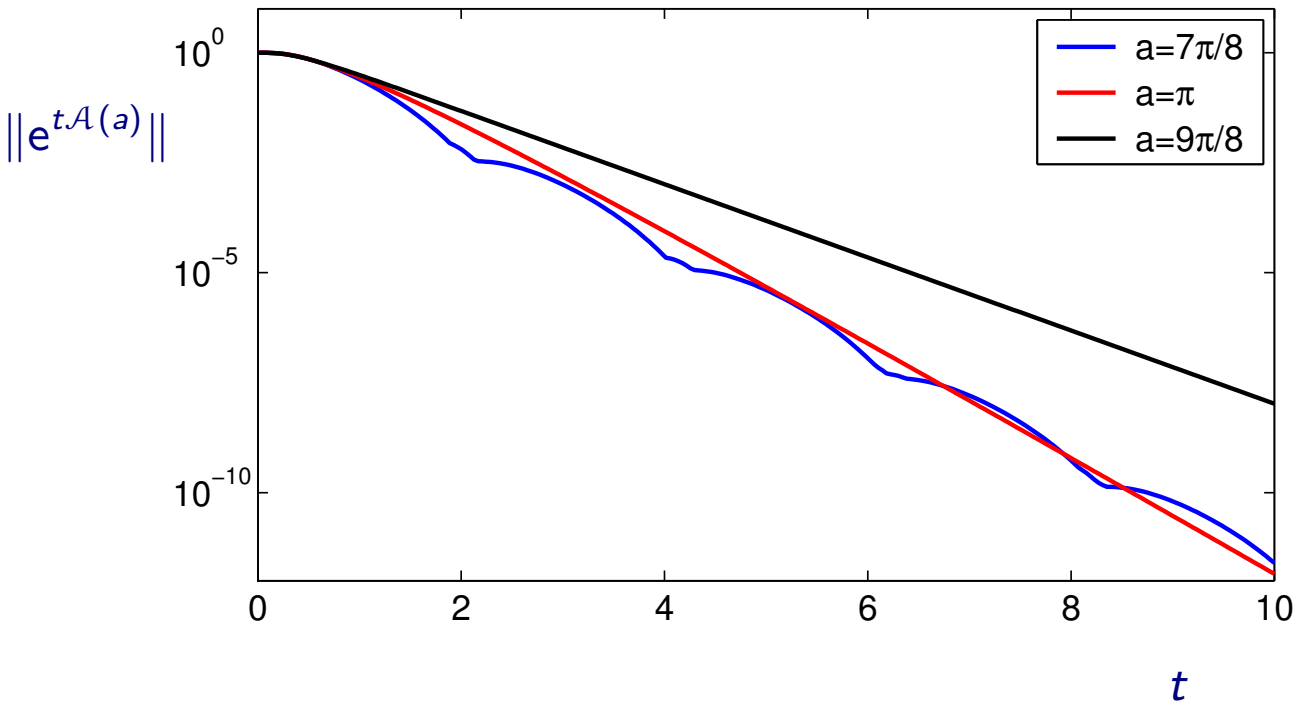
Pseudospectra for $a(x) \equiv 3$ (posed on $[0, \pi]$):



When a is constant, $\sigma(A) = \{-a \pm \sqrt{a^2 - k^2} : k = 1, 2, \ldots\}$

Wave Equation with Constant Damping: Transient Behavior

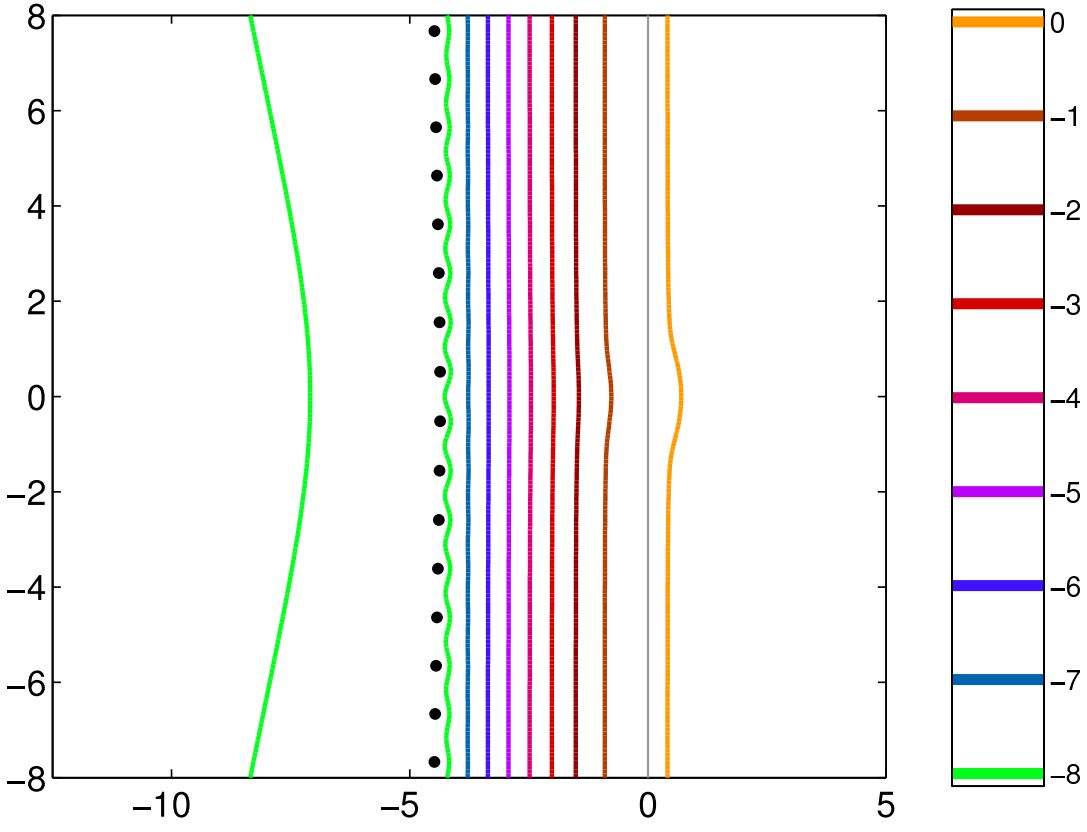
Though $a = \pi$ gives optimal *asymptotic decay*, it is not best at all time scales.



Viscous damping: $G(x)v(x) = -2a(x)v(x)$

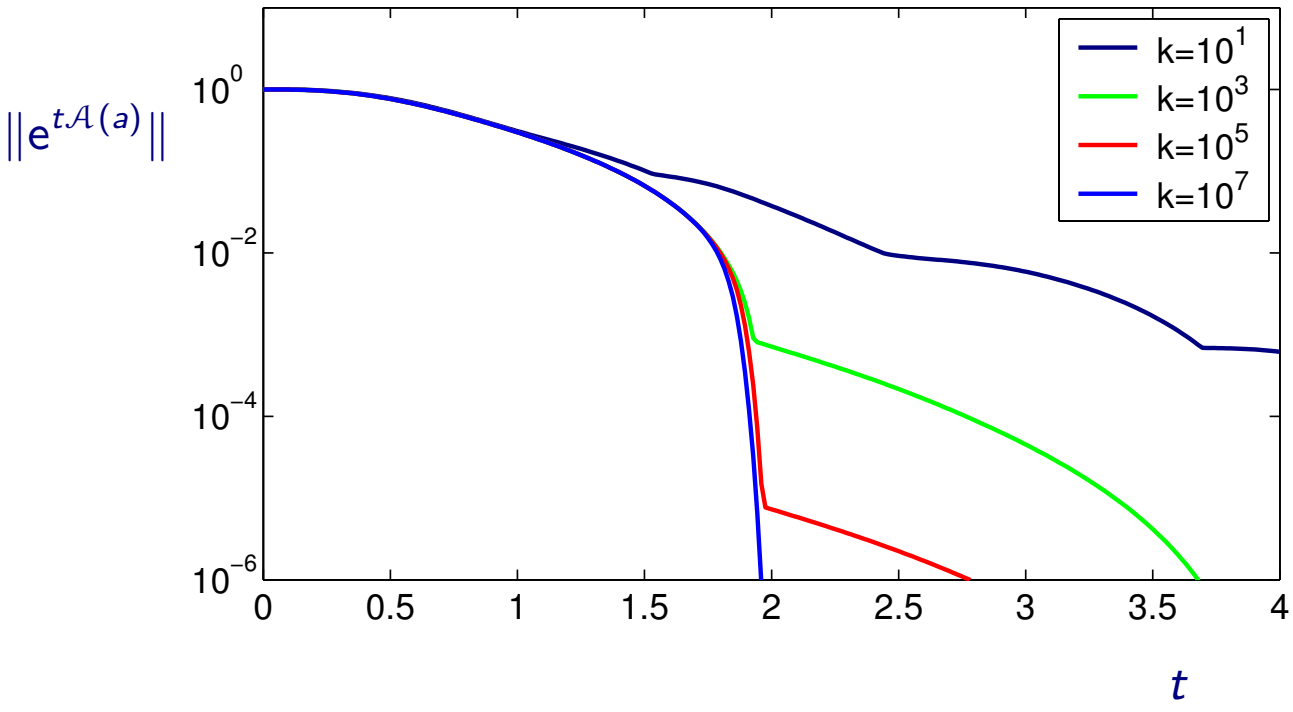
Cox and Overton [1996] asked, "Is constant damping optimal?"
Freitas [1998], Cox and Castro [2001] show not.

Pseudospectra for $a(x) = 1/(x + 10^{-10})$:



. . . but asymptotic optimality comes at the expense of sensitive eigenvalues.

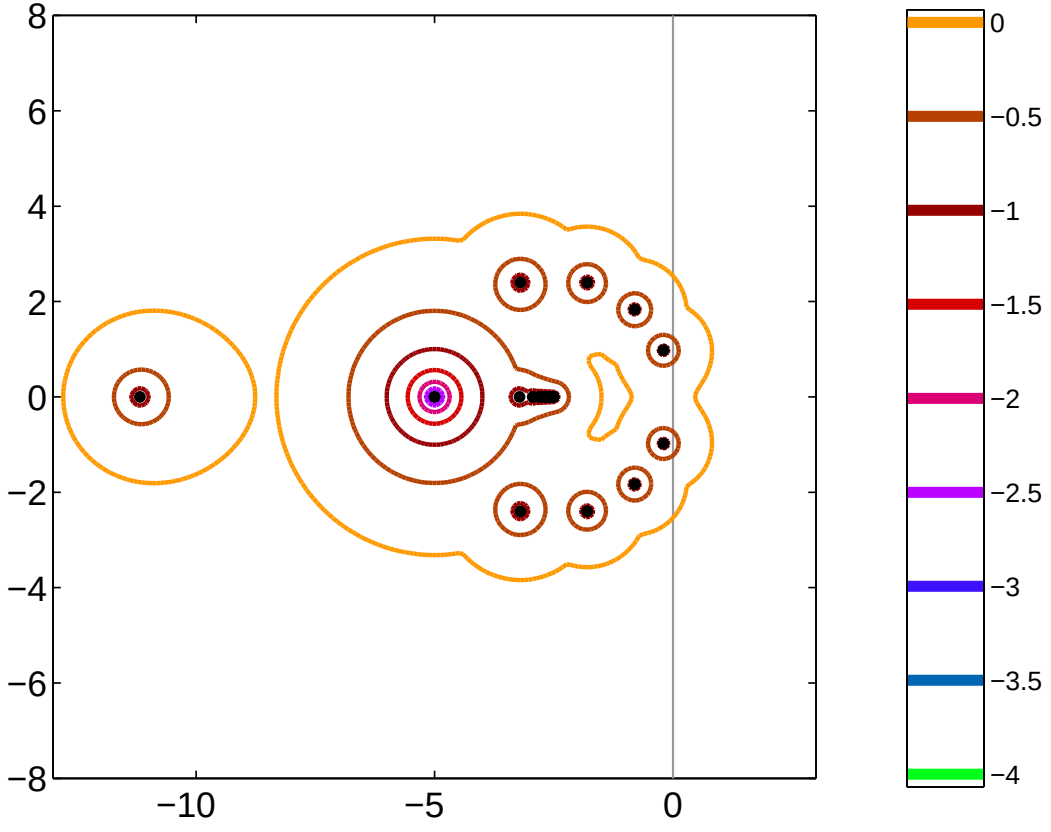
Wave Equation with Optimal Damping: Transient Behavior



- Transient slow convergence before eventual rapid decay.
- Note (approximate) extinction at $t = 2$.

Kelvin–Voigt damping: $G(x)v(x) = (a(x)v_t(x))_x$

Pseudospectra for $a(x) \equiv 2/5$:

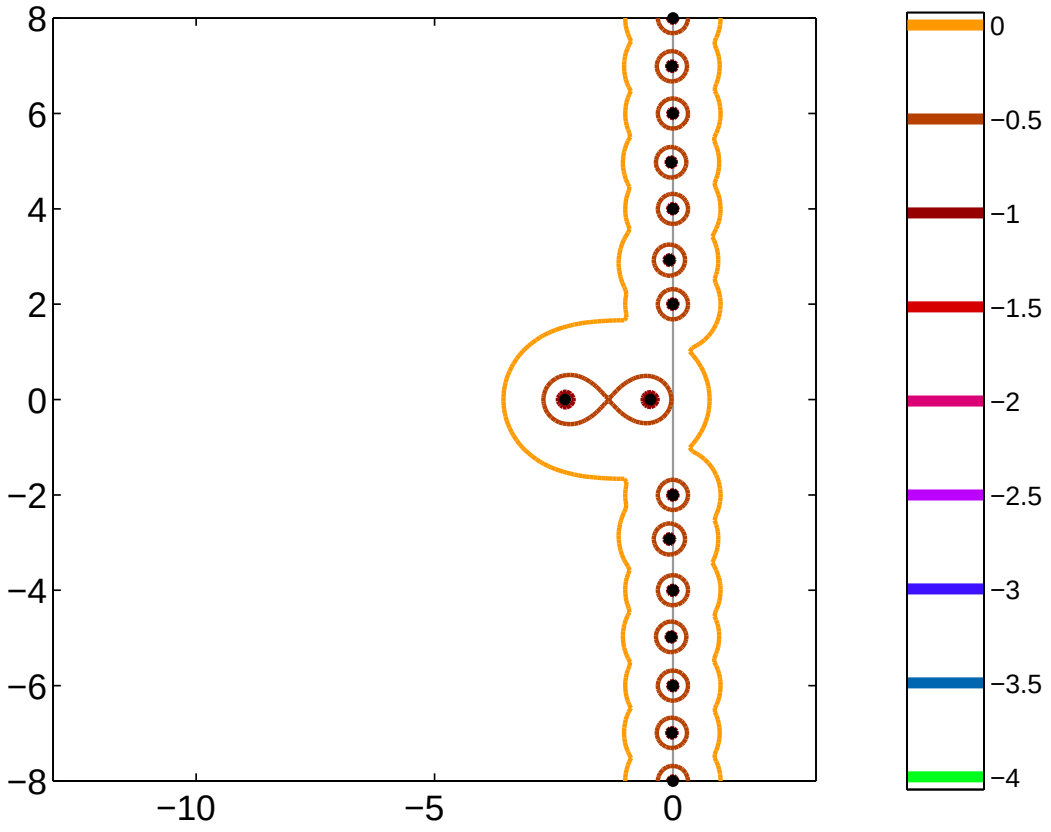


When a is constant, $\sigma(A) = \left\{ \dfrac{-ak^2 \pm \sqrt{a^2k^4 - 4k^2}}{2} : k = 1, 2, \ldots \right\}$.

See, e.g., [Liu and Liu 2002; Renardy 2004]

Magnetic damping: $G(x)v(x) = -a(x)\int_0^\pi a(s)v(s)\,\mathrm{d}s$

Pseudospectra for $a(x) \equiv 1$:



Inverse problem for Viscous Damping

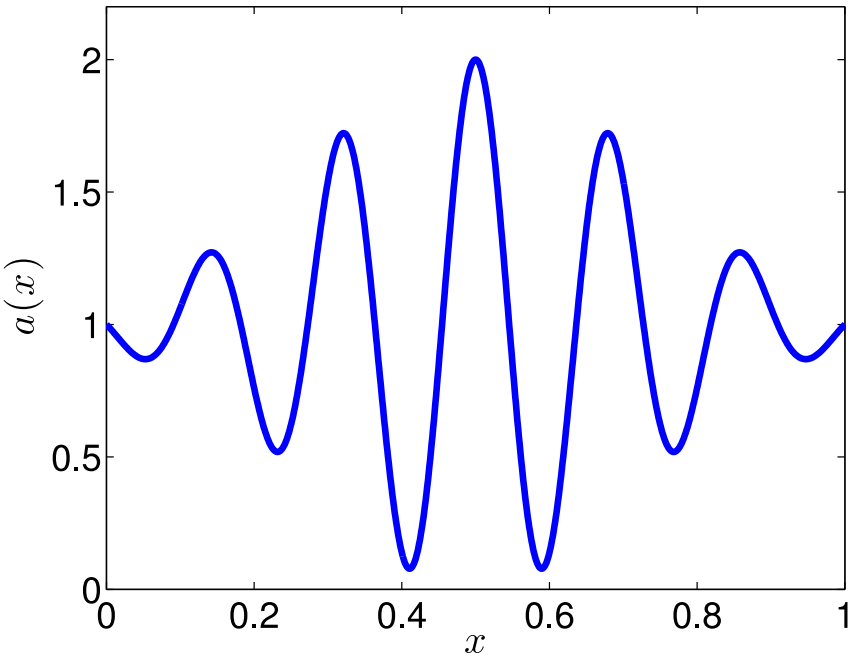
- We are interested in *designing strings* that sound a certain way.
- How can you distribute the damping material to give a certain sound?
- If you can "listen" to a string, can you recover its spectrum?
- From the spectrum, can you determine the damping coefficient?

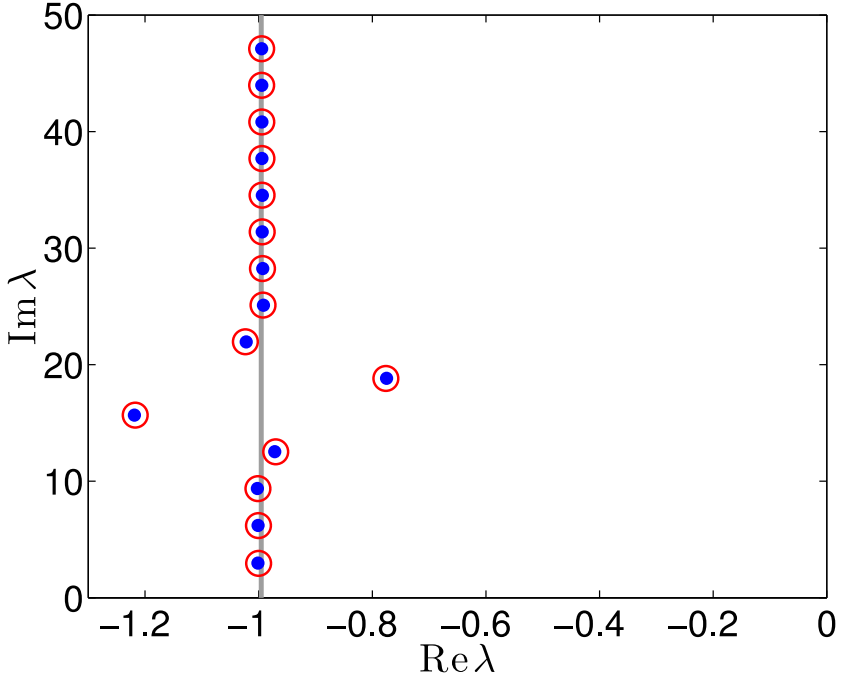
Inverse problem for Viscous Damping

- We are interested in *designing strings* that sound a certain way.
- How can you distribute the damping material to give a certain sound?
- If you can "listen" to a string, can you recover its spectrum?
- From the spectrum, can you determine the damping coefficient?

We have an asymptotic method to estimate an *even* $a(x)$ from eigenvalue data [Cox & E., 2011].

Accuracy of Spectral Asymptotics for Damped String





- true eigenvalues
- asymptotic expression

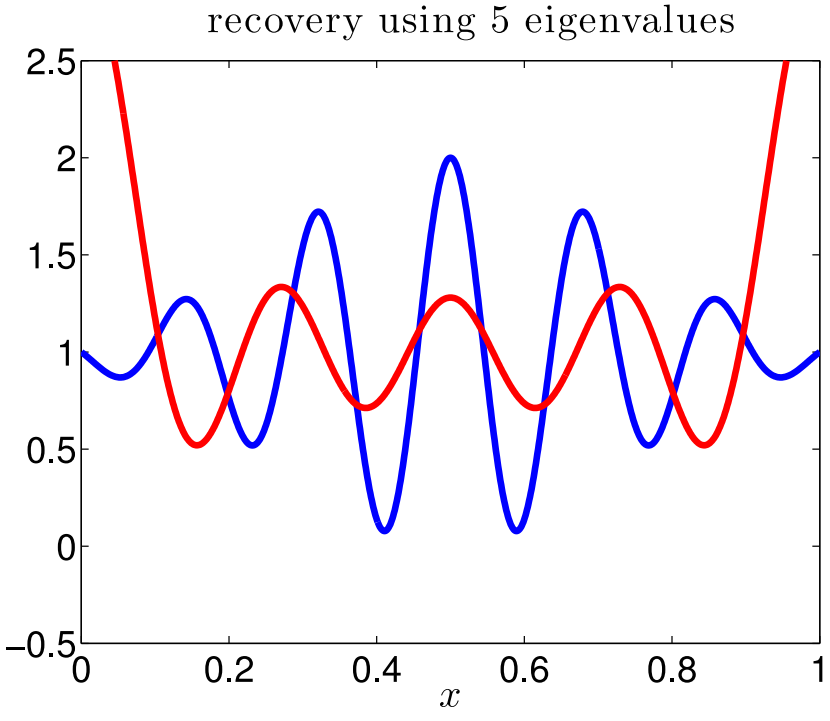
Accuracy of Spectral Asymptotics for Damped String

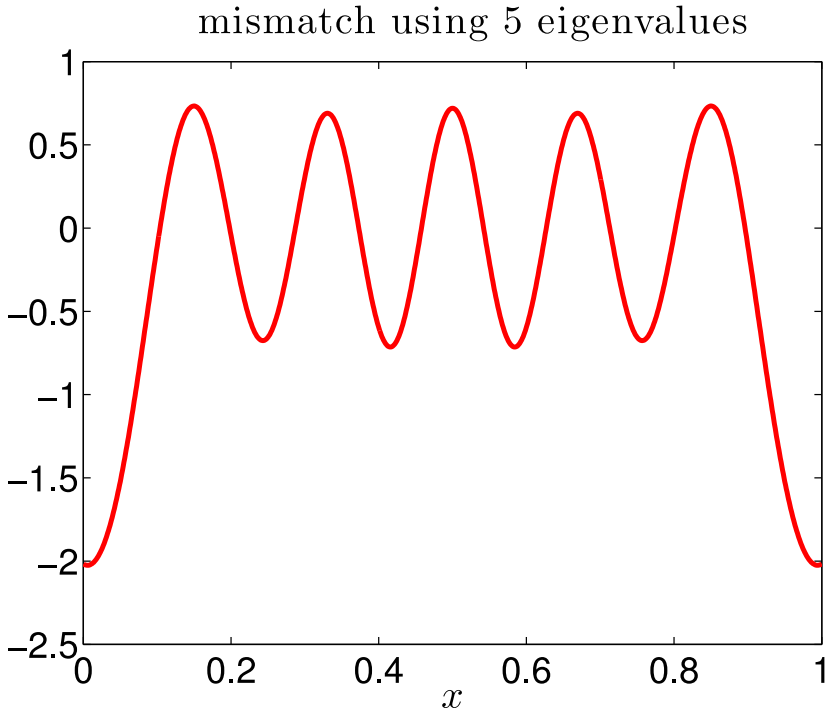
Even at low frequencies, our asymptotic formula often does well.

For the damping function on the previous slide:

k	$\lambda_k(a)$	*error*
1	$-1.00122 + 2.97964\,i$	0.0269879
2	$-1.00106 + 6.20514\,i$	0.0167440
3	$-1.00228 + 9.37615\,i$	0.0144677
4	$-0.97180 + 12.53868\,i$	0.0197670
5	$-1.21888 + 15.66851\,i$	0.0022341
6	$-0.77494 + 18.82612\,i$	0.0080114
7	$-1.02197 + 21.95436\,i$	0.0097917
8	$-0.99174 + 25.10548\,i$	0.0036935
9	$-0.99287 + 28.25021\,i$	0.0026511
10	$-0.99374 + 31.39954\,i$	0.0005894

Recovery of Even Viscous Damping

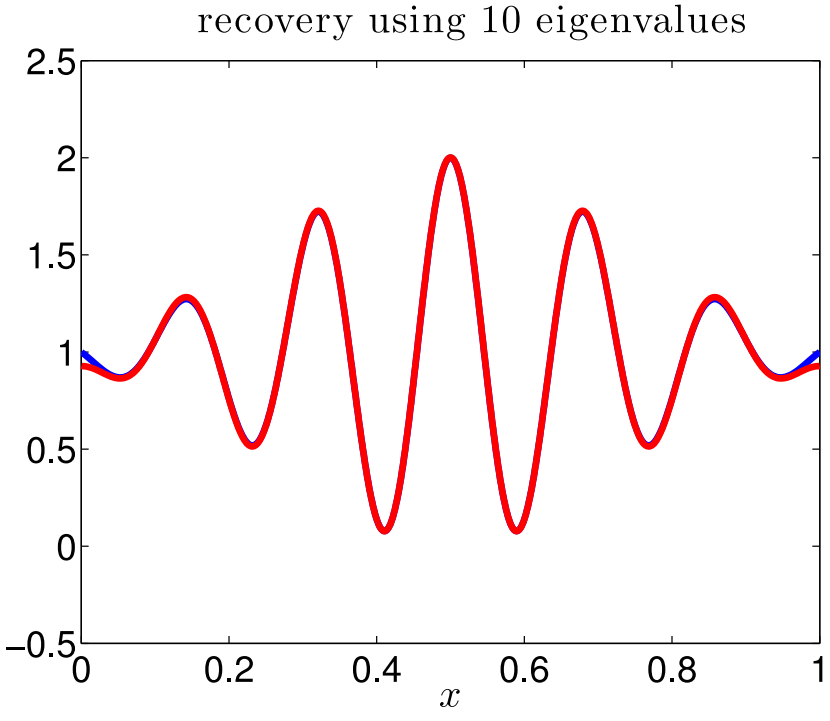


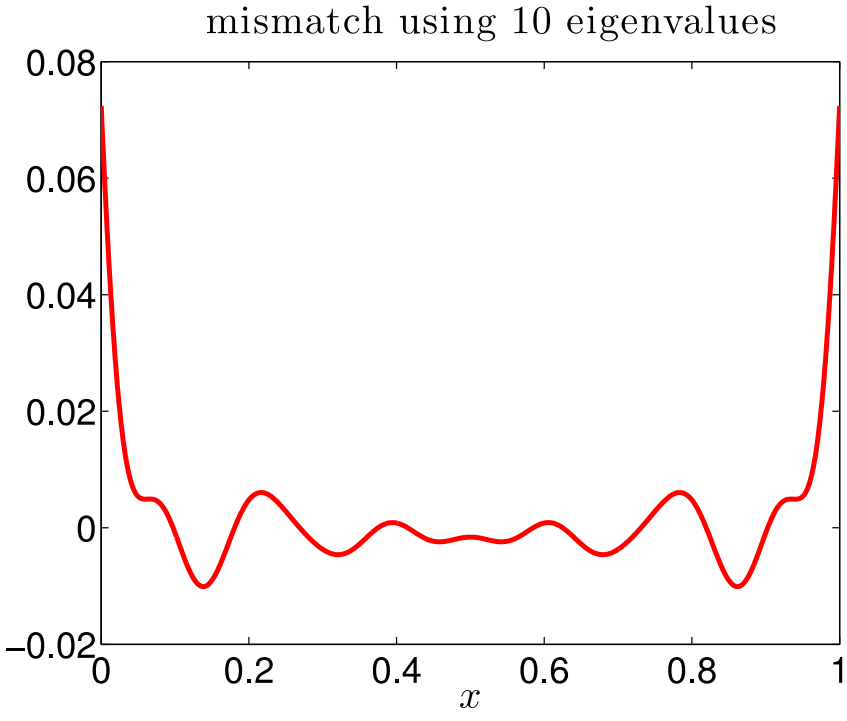


damping function

reconstructed from 5 eigenvalues

Recovery of Even Viscous Damping

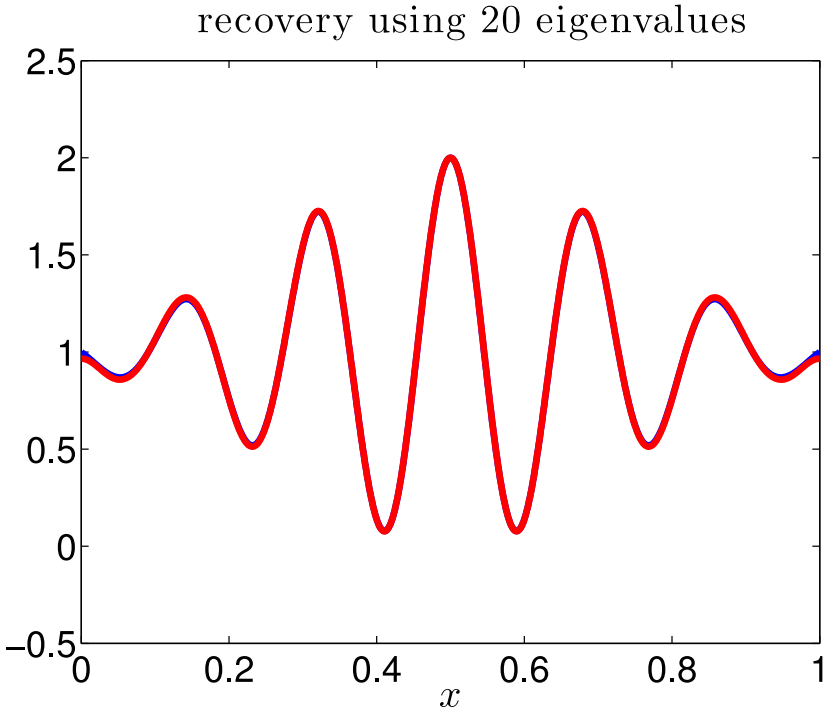


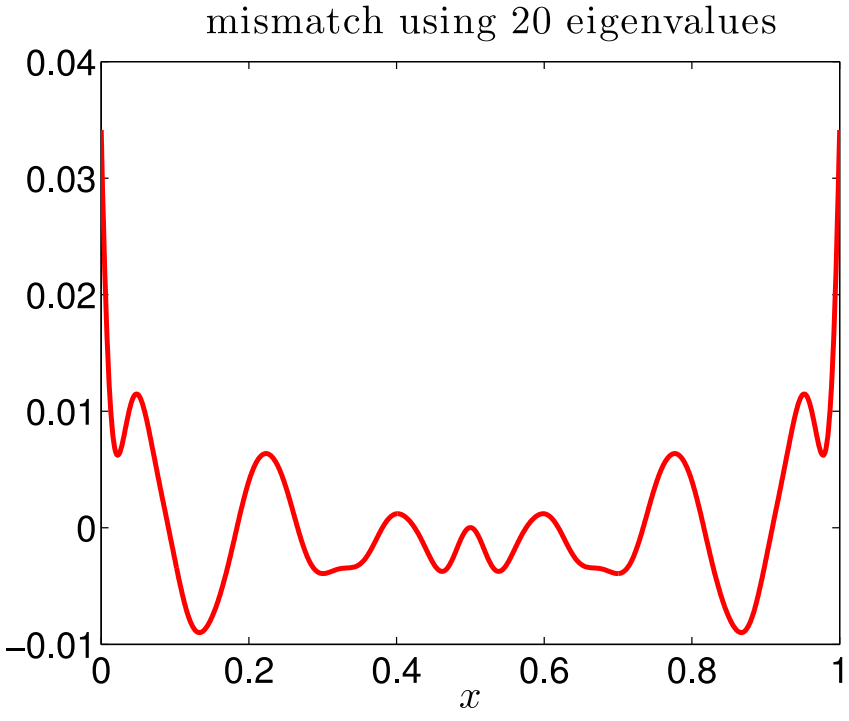


damping function

reconstructed from 10 eigenvalues

Recovery of Even Viscous Damping

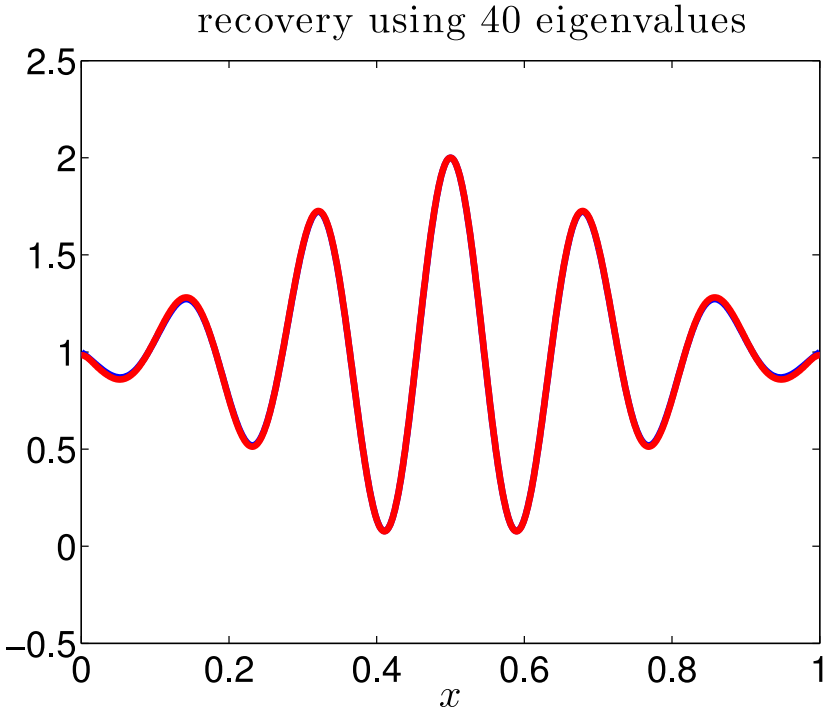


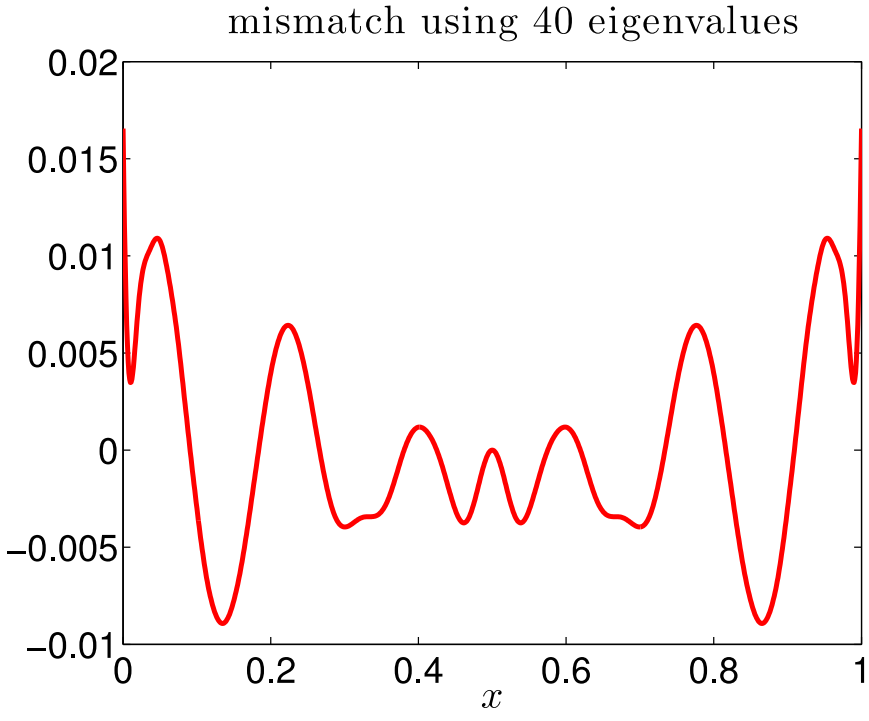


damping function

reconstructed from 20 eigenvalues

Recovery of Even Viscous Damping

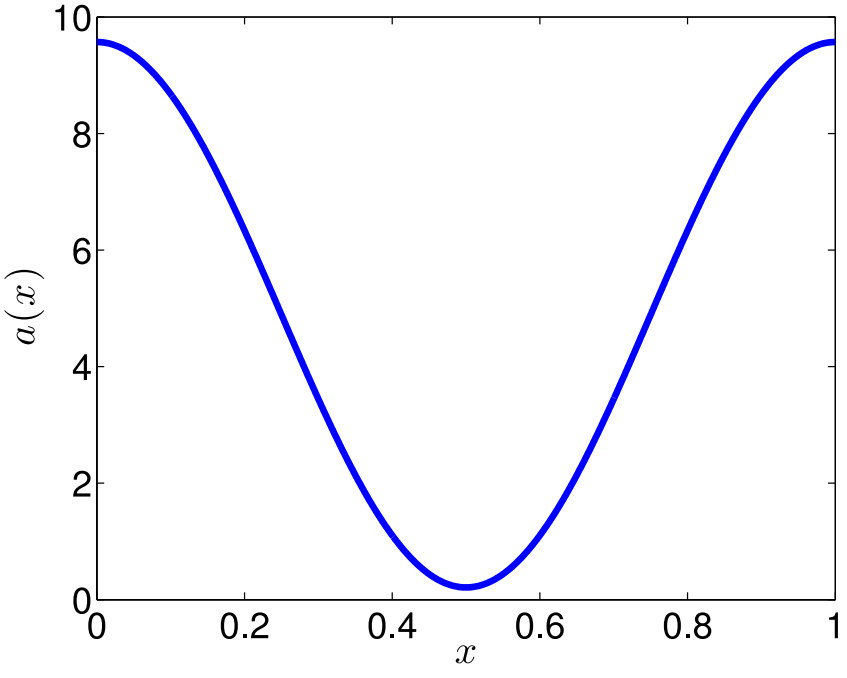


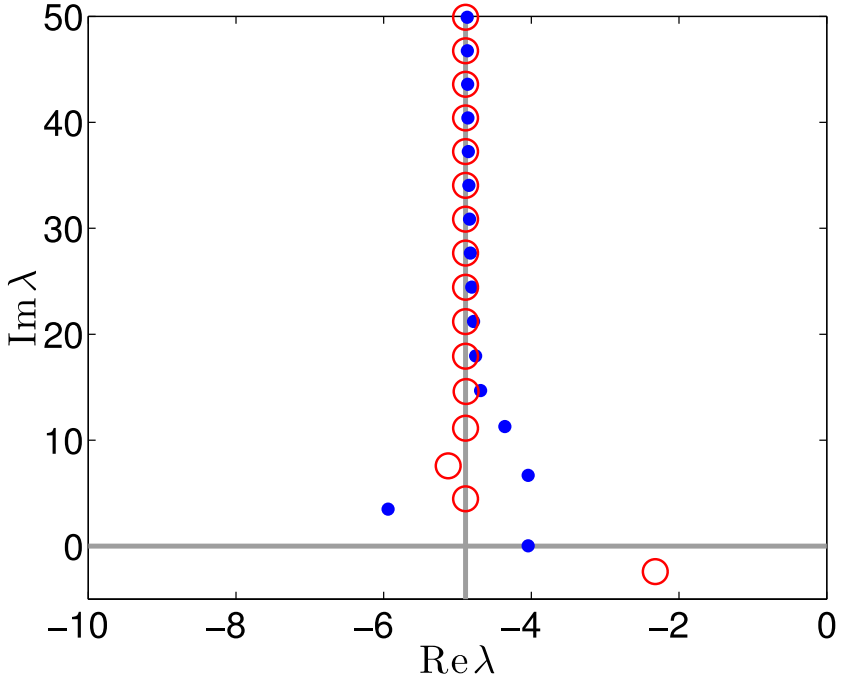


damping function
reconstructed from 40 eigenvalues

A Cautionary Example for Inverse Eigenvalue Problems

A damping function of Freitas [1998] that beat the best constant conjecture.

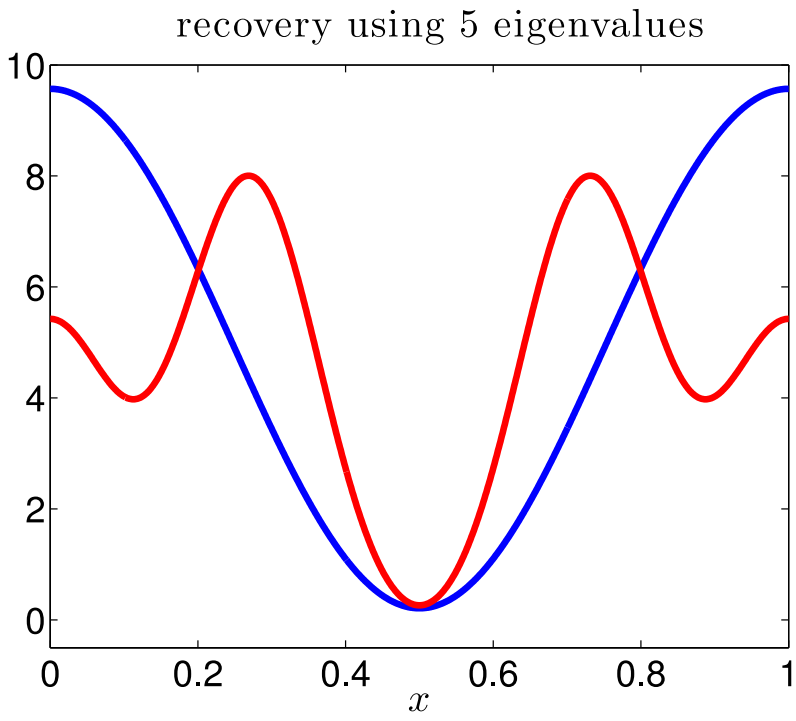


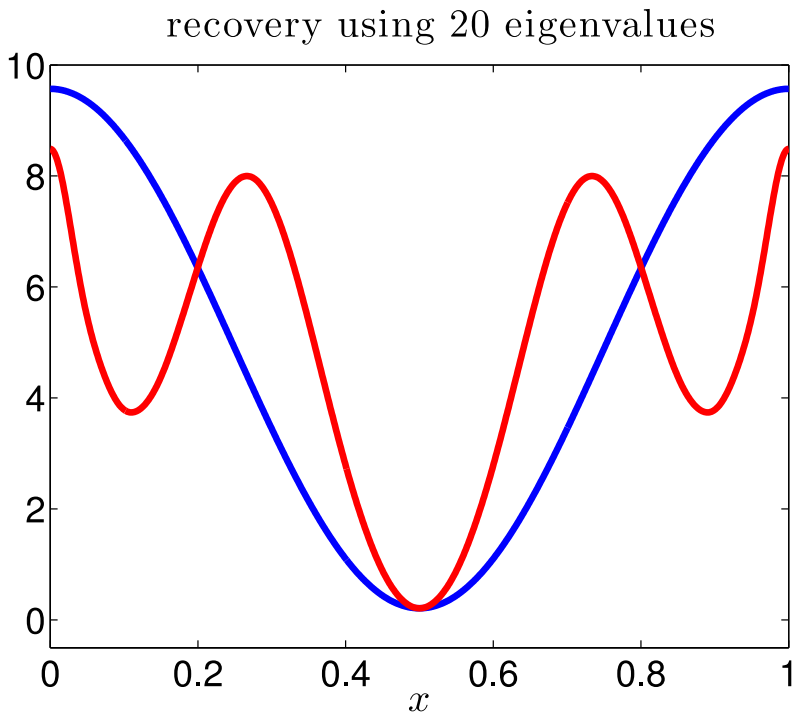


- true eigenvalues
- asymptotic expression

A Cautionary Example for Inverse Eigenvalue Problems

A damping function of Freitas [1998] that beat the best constant conjecture.



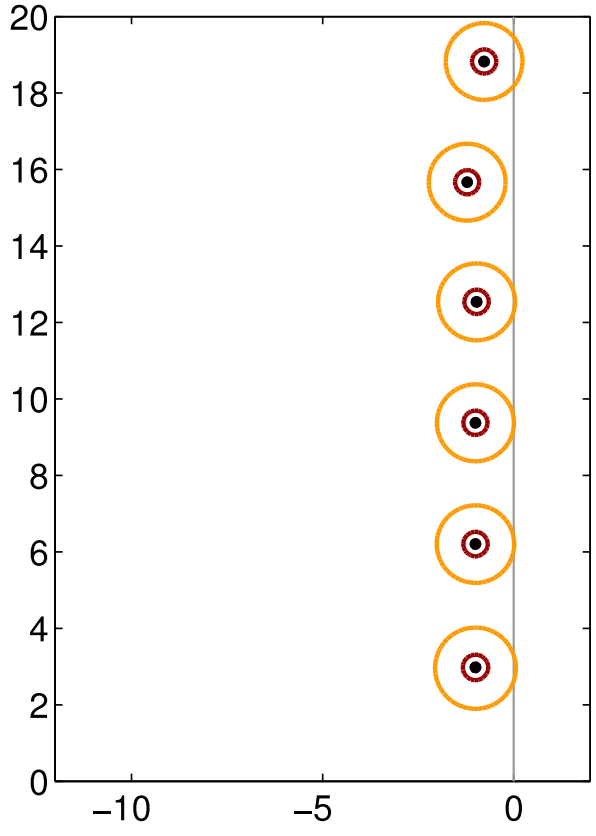


damping function

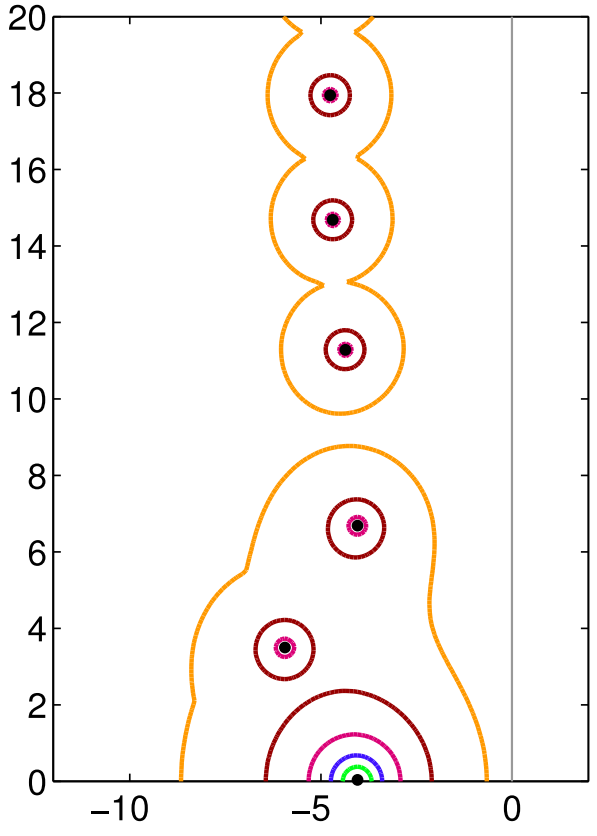
reconstructed from 5 eigenvalues

A Cautionary Example

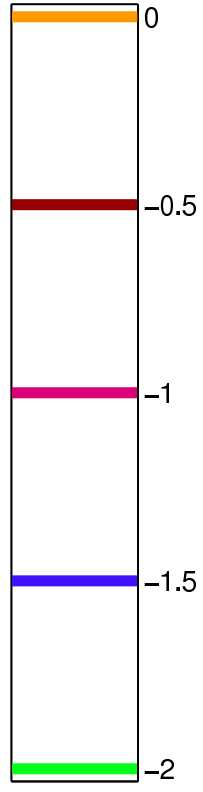
Pseudospectra hint at the problem. . . .



good reconstruction



poor reconstruction



5(e) Applications

Overview of the Course

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

By the end of the week, you will have a thorough understanding of the sources of nonnormality, its affects on the behavior of dynamical systems, tools for assessing its impact, and examples where it arises.

You will be able to understand phenomena that many people find quite mysterious when encountered in the wild.

Hopefully I have fulfilled this promise.

Thank you for your engagement and interest!