Introduction to Pseudospectra and their Applications

Arne Jensen

Department of Mathematical Sciences
Aalborg University

Aalborg, September 10, 2009

Motivation

Consider the problem

$$\frac{du}{dt} = Au, \qquad u(0) = u_0.$$

Solution is formally

$$u(t) = e^{tA}u_0.$$

Interpretation depends on the context.

- A operator on a finite dimensional space $\mathcal{H}$.

Motivation

Consider the problem

$$\frac{du}{dt} = Au, \qquad u(0) = u_0.$$

Solution is formally

$$u(t) = e^{tA}u_0.$$

Interpretation depends on the context.

- A operator on a finite dimensional space $\mathcal{H}$.
- A a bounded operator on an infinite dimensional Hilbert or Banach space.

Motivation

Consider the problem

$$\frac{du}{dt} = Au, \qquad u(0) = u_0.$$

Solution is formally

$$u(t) = e^{tA}u_0.$$

Interpretation depends on the context.

- A operator on a finite dimensional space $\mathcal{H}$.
- A a bounded operator on an infinite dimensional Hilbert or Banach space.
- A unbounded and generator of a semigroup on an infinite dimensional Hilbert or Banach space.

Motivation

Consider the problem

$$\frac{du}{dt} = Au, \qquad u(0) = u_0.$$

Solution is formally

$$u(t) = e^{tA}u_0.$$

Interpretation depends on the context.

- A operator on a finite dimensional space $\mathcal{H}$.
- A a bounded operator on an infinite dimensional Hilbert or Banach space.
- A unbounded and generator of a semigroup on an infinite dimensional Hilbert or Banach space.

I will limit the discussion to the first two cases.

Motivation

A result in the finite dimensional case. A an $N \times N$ matrix with

$$\operatorname{Re}\lambda < 0 \quad \text{for all } \lambda \in \sigma(A).$$

Then

$$u(t) = e^{tA}u_0 \to 0 \quad \text{as} \quad t \to \infty$$

for all initial conditions u_0.

Some questions.

- How fast will the solution approach zero.

Motivation

A result in the finite dimensional case. A an $N \times N$ matrix with

$$\operatorname{Re} \lambda < 0 \quad \text{for all } \lambda \in \sigma(A).$$

Then

$$u(t) = e^{tA} u_0 \to 0 \quad \text{as} \quad t \to \infty$$

for all initial conditions u_0.

Some questions.

- How fast will the solution approach zero.
- How large will the solution get before approaching zero.

Motivation

A result in the finite dimensional case. A an $N \times N$ matrix with

$$\operatorname{Re} \lambda < 0 \quad \text{for all } \lambda \in \sigma(A).$$

Then

$$u(t) = e^{tA} u_0 \to 0 \quad \text{as} \quad t \to \infty$$

for all initial conditions u_0.

Some questions.

- How fast will the solution approach zero.
- How large will the solution get before approaching zero.

It is this last quantative question we will try to answer.

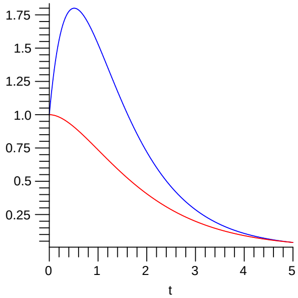
Motivation

An example:

$$A = \begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix}, \qquad B = \begin{bmatrix} -1 & 5 \\ 0 & -2 \end{bmatrix}.$$

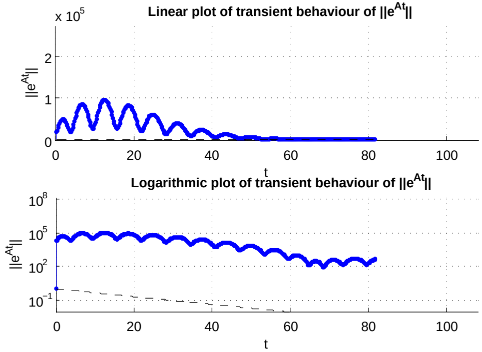
We plot $\|e^{tA}\|$ and $\|e^{tB}\|$ as functions of t.

Which is which?



Motivation

Here is an example from real life: Boing aircraft wing flutter. Matrix is 55×55.



Motivation—Resolvent and Spectrum

Operator equation $Ax - \lambda x = b, \quad \lambda \notin \sigma(A)$. Solution

$$x = (A - \lambda I)^{-1} b.$$

Perturb right hand side, $\|u\| \leq \varepsilon$, i.e. consider $Ay - \lambda y = b + u$. We have

$$\|x - y\| \leq \varepsilon \|(A - \lambda I)^{-1}\|.$$

Thus the effect depends on the norm of the resolvent. It may be large far from the spectrum. Recall general result:

$$\|(A - zI)^{-1}\| \geq \frac{1}{\operatorname{dist}(z, \sigma(A))}$$

Definition of Pseudospectrum

Definition

Let $A \in \mathcal{B}(\mathcal{H})$ and $\varepsilon > 0$. The ε-pseudospectrum of A is given by

$$\sigma_\varepsilon(A) = \sigma(A) \cup \{z \in \mathbf{C} \setminus \sigma(A) \mid \|(A - zI)^{-1}\| > \varepsilon^{-1}\}.$$

Definition of Pseudospectrum

Definition

Let $A \in \mathcal{B}(\mathcal{H})$ and $\varepsilon > 0$. The ε-pseudospectrum of A is given by

$$\sigma_\varepsilon(A) = \sigma(A) \cup \{z \in \mathbf{C} \setminus \sigma(A) \mid \|(A - zI)^{-1}\| > \varepsilon^{-1}\}.$$

Theorem

Let $A \in \mathcal{B}(\mathcal{H})$ and $\varepsilon > 0$. Then the following three statements are equivalent.

(i) $z \in \sigma_\varepsilon(A)$.

(ii) *There exists $B \in \mathcal{B}(\mathcal{H})$ with $\|B\| < \varepsilon$ such that $z \in \sigma(A + B)$.*

(iii) *$z \in \sigma(A)$ or there exists $v \in \mathcal{H}$ with $\|v\| = 1$ such that $\|(A - zI)v\| < \varepsilon$.*

Definition of Pseudospectrum, Finite Dimension

Let T be an $N \times N$ matrix. The eigenvalues of $(T^*T)^{1/2}$ are called the singular values of T. The smallest singular value is denoted $s_{\min}(T)$.

Theorem

Assume that $\mathcal{H}$ is finite dimensional and $T \in \mathcal{B}(\mathcal{H})$. Let $\varepsilon > 0$. Then $z \in \sigma_\varepsilon(T)$ if and only if $s_{\min}(T - zI) < \varepsilon$.

Definition of Pseudospectrum, Finite Dimension

Let T be an $N \times N$ matrix. The eigenvalues of $(T^*T)^{1/2}$ are called the singular values of T. The smallest singular value is denoted $s_{\min}(T)$.

Theorem

Assume that $\mathcal{H}$ is finite dimensional and $T \in \mathcal{B}(\mathcal{H})$. Let $\varepsilon > 0$. Then $z \in \sigma_\varepsilon(T)$ if and only if $s_{\min}(T - zI) < \varepsilon$.

Since the singular values of a matrix can be computed numerically, this result provides a method for plotting the pseudospectra of a given matrix. One chooses a finite grid of points in the complex plane, and evaluates $s_{\min}(T - zI)$ at each point. Plotting level curves for these points provides a picture of the pseudospectra of T.

Properties of $\sigma_\varepsilon(A)$

Define $D_\delta = \{z \in \mathbf{C} \mid |z| < \delta\}$.

Proposition

Let $A \in \mathcal{B}(\mathcal{H})$. Each $\sigma_\varepsilon(A)$ is a bounded open subset of $\mathbf{C}$. We have $\sigma_{\varepsilon_1}(A) \subset \sigma_{\varepsilon_2}(A)$ for $0 < \varepsilon_1 < \varepsilon_2$. Furthermore, $\cap_{\varepsilon>0}\sigma_\varepsilon(A) = \sigma(A)$. For $\delta > 0$ we have $D_\delta + \sigma_\varepsilon(A) \subseteq \sigma_{\varepsilon+\delta}(A)$. We have $\sigma_\varepsilon(A^*) = \overline{\sigma_\varepsilon(A)}$.

Properties of $\sigma_\varepsilon(A)$

Define $D_\delta = \{z \in \mathbf{C} \mid |z| < \delta\}$.

Proposition

Let $A \in \mathcal{B}(\mathcal{H})$. Each $\sigma_\varepsilon(A)$ is a bounded open subset of $\mathbf{C}$. We have $\sigma_{\varepsilon_1}(A) \subset \sigma_{\varepsilon_2}(A)$ for $0 < \varepsilon_1 < \varepsilon_2$. Furthermore, $\cap_{\varepsilon>0}\sigma_\varepsilon(A) = \sigma(A)$. For $\delta > 0$ we have $D_\delta + \sigma_\varepsilon(A) \subseteq \sigma_{\varepsilon+\delta}(A)$. We have $\sigma_\varepsilon(A^*) = \overline{\sigma_\varepsilon(A)}$.

Proposition

Let $A \in \mathcal{B}(\mathcal{H})$ and assume that $V \in \mathcal{B}(\mathcal{H})$ is invertible. Let $\kappa = \mathrm{cond}(V)(= \|V\| \cdot \|V^{-1}\|)$. Let $B = VAV^{-1}$. Then

$$\sigma(B) = \sigma(A),$$

and for $\varepsilon > 0$ we have

$$\sigma_{\varepsilon/\kappa}(A) \subseteq \sigma_\varepsilon(B) \subseteq \sigma_{\kappa\varepsilon}(A).$$

Properties of $\sigma_\varepsilon(A)$

Proposition

Let $A \in \mathcal{B}(\mathcal{H})$ *and* $\varepsilon > 0$. *Then*

$$\{z \mid \operatorname{dist}(z, \sigma(A)) < \varepsilon\} \subseteq \sigma_\varepsilon(A).$$

If A *is normal, then*

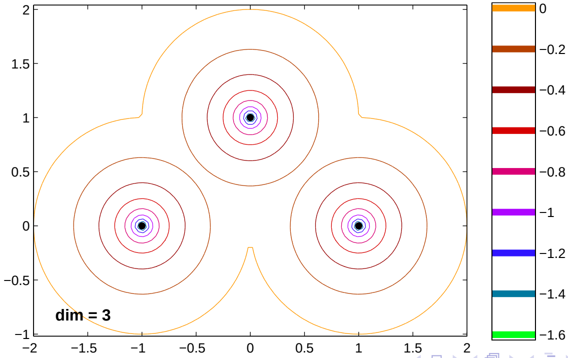
$$\sigma_\varepsilon(A) = \{z \mid \operatorname{dist}(z, \sigma(A)) < \varepsilon\}.$$

Example 1

Take

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

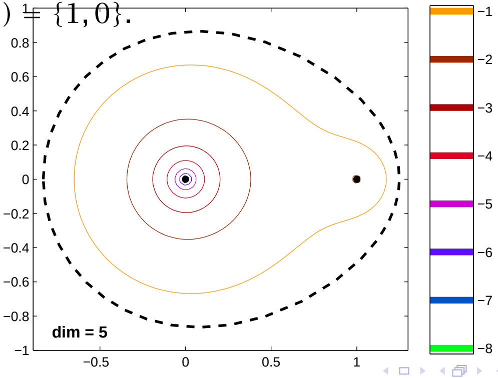
with $\sigma(A) = \{1, -1, i\}$.



Example 2

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

We have $\sigma(B) = \{1, 0\}$.



Example 2

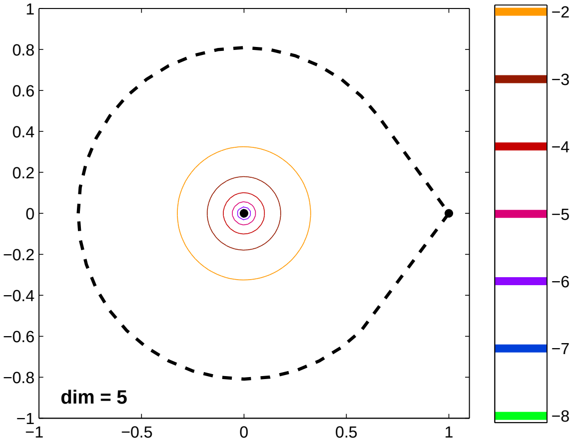
Let us look at the Jordan canonical form of this matrix. We have $J = Q^{-1}BQ$, where

$$J = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad Q = \begin{bmatrix} -1 & -1 & -1 & -1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

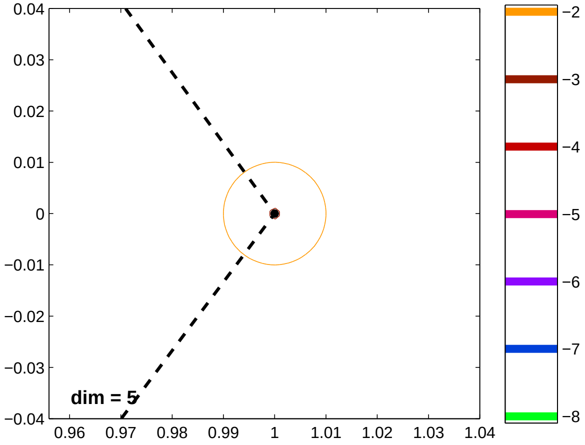
We have

$$\operatorname{cond}(Q) = 3 + 2\sqrt{2} \approx 5.828427125.$$

Example 2



Example 2


0.04
0.03
0.02
0.01
0
−0.01
−0.02
−0.03
−0.04
0.96
0.97
0.98
0.99
1
1.01
1.02
1.03
1.04
dim = 5
−2
−3
−4
−5
−6
−7
−8

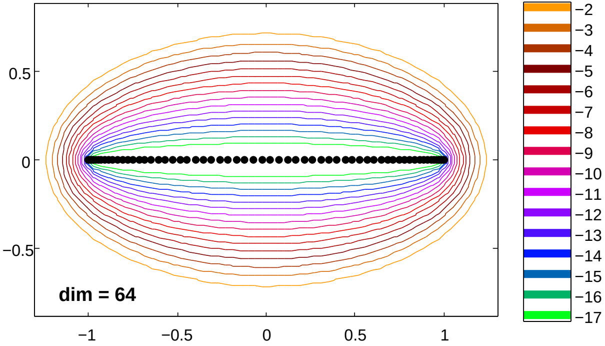
Example 3

A Toeplitz matrix

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 1/4 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 1/4 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 1/4 & 0 \end{bmatrix}$$

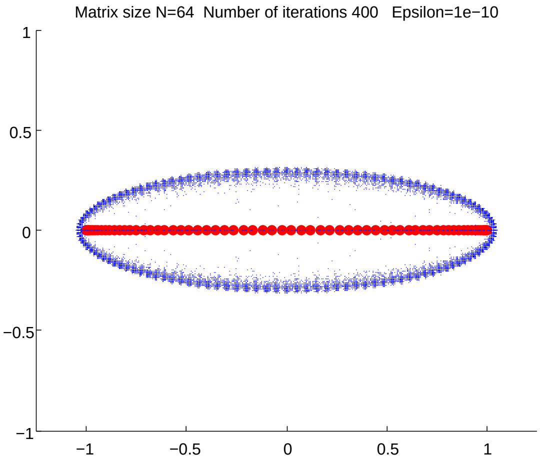
We have $A = SDS^{-1}$ with D diagonal.

Example 3



Example 3

$A + E$, E random matrix with $\|E\| < 10^{-10}$. Plot of spectra: blue. Spectrum of A: red.



Some definitions

Definition

Let $A \in \mathcal{B}(\mathcal{H})$. The numerical range of A is the set

$$W(A) = \{\langle u, Au\rangle \mid \|u\| = 1\}.$$

Theorem (Toeplitz-Hausdorff)

The numerical range $W(A)$ is always a convex set. If $\mathcal{H}$ is finite dimensional, then $W(A)$ is a compact set.

Proposition

Let $A \in \mathcal{B}(\mathcal{H})$. Then $\sigma(A) \subseteq \mathrm{cl}(W(A))$.

Some definitions

Definition

We define the following quantities.

$$\begin{aligned} \alpha(A) &= \sup\{\operatorname{Re} z \mid z \in \sigma(A)\}, \\ \alpha_\varepsilon(A) &= \sup\{\operatorname{Re} z \mid z \in \sigma_\varepsilon(A)\}, \\ \omega(A) &= \sup\{\operatorname{Re} z \mid z \in W(A)\}. \end{aligned}$$

We recall the relations

$$(A - zI)^{-1} = \int_0^\infty e^{-tz} e^{tA} dt \quad \text{for } \operatorname{Re} z \text{ sufficiently large.}$$

$$e^{tA} = \frac{-1}{2\pi i} \int_\gamma e^{tz} (A - zI)^{-1} dz$$

where γ is a simple closed contour enclosing $\sigma(A)$.

Estimates on $\|e^{tA}\|$

We have the following results from semigroup theory:

Theorem

$$\alpha(A) = \lim_{t\to\infty} \frac{1}{t} \log\|e^{tA}\|$$

Estimates on $\|e^{tA}\|$

We have the following results from semigroup theory:

Theorem

$$\alpha(A) = \lim_{t\to\infty} \frac{1}{t} \log\|e^{tA}\|$$

Theorem

$$\omega(A) = \frac{d}{dt}\,\|e^{tA}\|\Big|_{t=0} = \lim_{t\downarrow 0} \frac{1}{t} \log\|e^{tA}\|$$

Estimates on $\|e^{tA}\|$

Some simple estimates:

$$\|e^{tA}\| \geq e^{t\alpha(A)} \quad \text{for all } t \geq 0$$

$$\|e^{tA}\| \leq e^{t\omega(A)} \quad \text{for all } t \geq 0$$

An estimate based on pseudospectra:

Theorem

For all $\varepsilon > 0$ we have

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

Estimates on $\|e^{tA}\|$

Define the Kreiss constant

$$\mathcal{K}(A) = \sup_{\varepsilon>0} \frac{\alpha_\varepsilon(A)}{\varepsilon}$$

Theorem (Kreiss Matrix Theorem)

If A is an $N \times N$ matrix, then we have

$$\|e^{tA}\| \leq eN\mathcal{K}(A).$$

Estimates on $\|e^{tA}\|$

Some further results

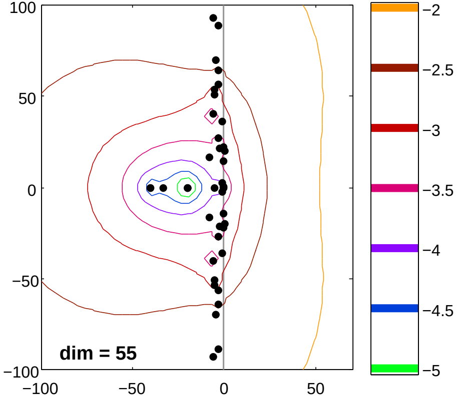
Theorem

Let $a = \operatorname{Re} z$. *Let* $K = \operatorname{Re} z \|(A - zI)^{-1}\|$. *Then for* $\tau > 0$ *we have*

$$\sup_{0 \le t \le \tau} \|e^{tA}\| \ge e^{a t} \left(1 + \frac{e^{at} - 1}{K}\right)^{-1}$$

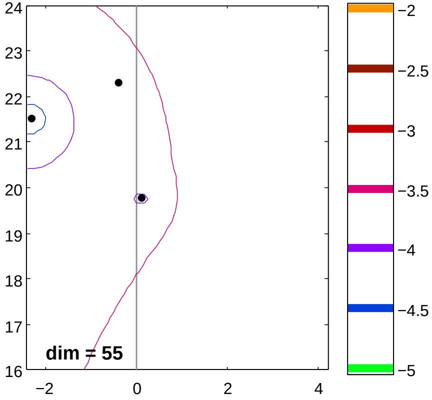
Some examples

Unstable Boing wing flutter matrix.



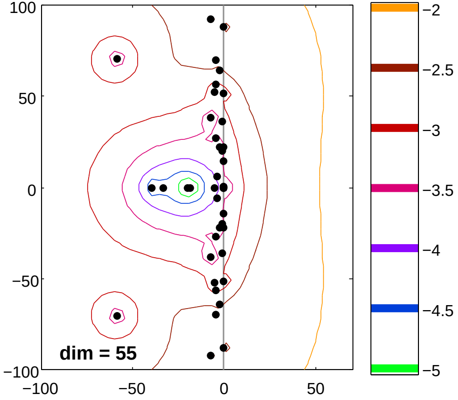
Some examples

Unstable Boing wing flutter matrix. Close-up near eigenvalue with Re $z > 0$



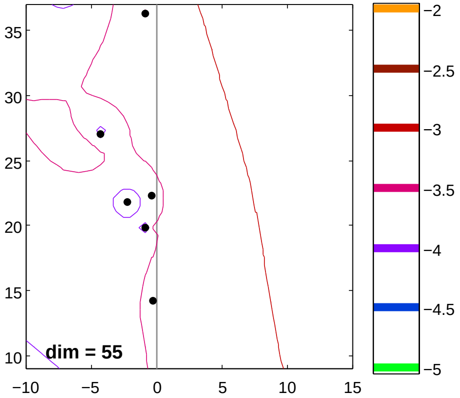
Some examples

Stable Boing wing flutter matrix.



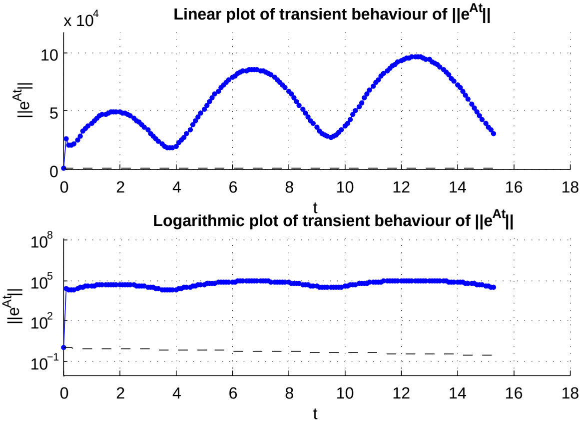
Some examples

Stable Boing wing flutter matrix. Close-up near eigenvalue with $\mathrm{Re}\, z \sim 0$



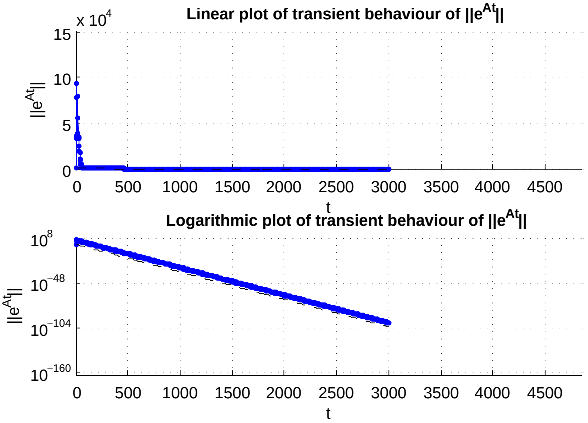
Some examples

Stable Boing wing flutter matrix. Initial behavior.



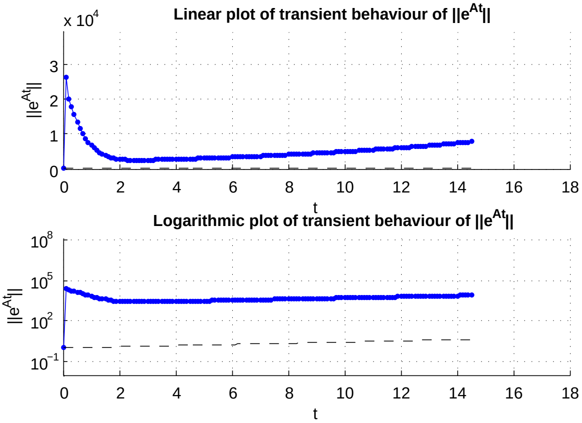
Some examples

Stable Boing wing flutter matrix. Long time behavior.



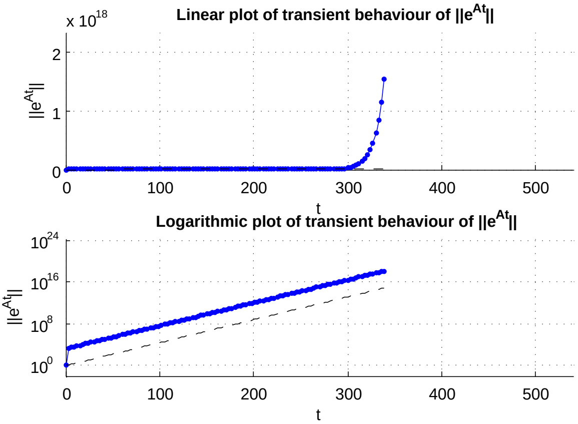
Some examples

Unstable Boing wing flutter matrix. Initial behavior.



Some examples

Unstable Boing wing flutter matrix. Long time behavior.

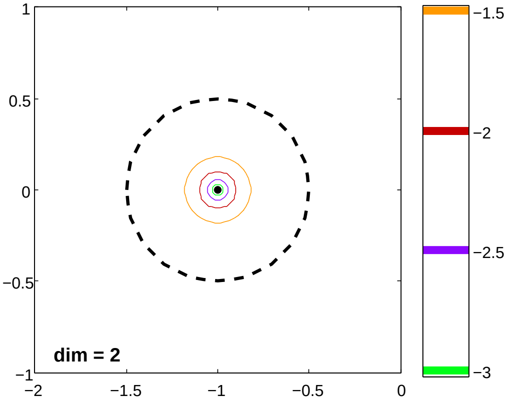


The first examples

Matrix

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

Boundary of numerical range included.



The first examples

Matrix

$$B = \begin{bmatrix} -1 & 5 \\ 0 & -2 \end{bmatrix}$$

Boundary of numerical range included.