GIAN course: Singular optimal control

Slide collection 2

Basic facts on linear ODE and matrix exponential

Michael Karow, TU Berlin, December 2016

karow@math.tu-berlin.de

Linear ODE, existence and uniqueness of solutions

Linear initial value problem:

$$\dot{X}(t) = A(t)\,X(t) + F(t), \quad t \in [a,b], \qquad X(t_0) = X_0,$$

where $X(t), F(t) \in \mathbb{C}^{n\times m}$, $A(t) \in \mathbb{C}^{n\times n}$, and A, F are piecewise continuous $[a,b]$.

Theorem: The initial value problem has a unique solution.

Proof. on the following slide.

Linear ODE, existence and uniqueness of solutions

The existence and uniqueness proof.

W.l.o.g. $t_0 = 0$. The function X solves the initial value problem

$$\dot{X}(t) = A(t)\,X(t) + F(t), \quad X(0) = X_0$$

iff

$$X(t) = \Psi(X)(t) := X_0 + \int_0^t A(s)\,X(s) + F(s)\,ds.$$

We show that the operator $X \mapsto \Psi(X)$ is a contraction with respect to the norm

$$|X| := \max_{t\in[a,b]} \|e^{-Lt}X(t)\|, \quad \text{where} \quad L := 2\sup_{t\in[a,b]} \|A(t)\|, \quad \|\cdot\| \text{ any submultiplicative norm.}$$

Then the existence and uniqueness statement follows from the contraction mapping theorem. For two matrix functions X_1, X_2,

$$\begin{aligned}
\|\Psi(X_1)(t) - \Psi(X_2)(t)\| &= \|\int_0^t A(s)(X_1(s) - X_2(s))\,ds\| \\
&\le \int_0^t \frac{L}{2} e^{Ls}\, e^{-Ls}\|X_1(s) - X_2(s)\|\,ds \\
&\le \frac{L}{2}\int_0^t e^{Ls}\,|X_1 - X_2|\,ds \\
&\le \frac{L}{2}\left[e^{Ls}/L\right]_{s=0}^{s=t}\,|X_1 - X_2| \le \frac{1}{2}\,e^{Lt}|X_1 - X_2|.
\end{aligned}$$

Multiplying the latter inequality with e^{-Lt} and maximizing yields

$$|\Psi(X_1) - \Psi(X_2)| \le \frac{1}{2}\,|X_1 - X_2|.$$

Fundamental solution, transition map, variation of parameters formula.

The unique solution $\Phi : [a, b] \to \mathbb{C}^{n\times n}$ of

$$\dot{\Phi}(t) = A(t)\,\Phi(t), \qquad \Phi(t_0) = I \qquad \text{(identity matrix)}$$

is called the **fundamental solution** at t_0. $\Phi(t)$ is invertible for all t. Suppose not. Then $y(t) := \Phi(t)z = 0$ for some $z \in \mathbb{C}^n \setminus \{0\}$ and some t. The function y satisfies $\dot{y} = Ay$, $y(t) = 0$. By uniqueness, $y \equiv 0$. In particular $0 = y(t_0) = \Phi(t_0)z = z$, contradiction.

Define the **transition map** by

$$\Phi(t, s) := \Phi(t)\Phi^{-1}(s), \qquad s, t \in [a, b].$$

This map does not depend on t_0. Then the solution of

$$\dot{X}(t) = A(t)\,X(t) + F(t), \quad t \in [a, b], \qquad X(t_0) = X_0, \qquad (*)$$

is given by

$$\begin{aligned} X(t) &= \Phi(t)\left(X_0 + \int_{t_0}^{t} \Phi(s)^{-1}\,F(s)\,ds\right) \\ &= \Phi(t, t_0)\,X_0 + \int_{t_0}^{t} \Phi(t, s)\,F(s)\,ds. \end{aligned}$$

<u>Proof.</u> The ansatz $X(t) = \Phi(t)\,C(t)$ for the solution of $(*)$ yields

$$\dot{C}(t) = \Phi(t)^{-1}\,F(t), \qquad C(t_0) = X_0. \qquad \Rightarrow \quad C(t) = X_0 + \int_{t_0}^{t} \Phi(s)^{-1}\,F(s), ds.$$

Linear time invariant ODE and matrix exponential

The unique solution of

$$\dot{\Phi}(t) = A\,\Phi(t), \qquad \Phi(0) = I \qquad (A \text{ constant})$$

is given by the matrix exponential

$$\Phi(t) = e^{At} := \exp(At) := \sum_{k=0}^{\infty} \frac{(At)^k}{k!} = I + At + \frac{(At)^2}{2} + \frac{(At)^3}{6} + \ldots$$

The solution of

$$\dot{X}(t) = A\,X(t) + F(t), \quad t \in \mathbb{R}, \qquad X(0) = X_0, \qquad (*)$$

is given by

$$X(t) \;=\; e^{At}\left(X_0 + \int_0^t e^{-As}\,F(s)\,ds\right) = e^{A\,t}\,X_0 + \int_0^t e^{A(t-s)}\,F(s)\,ds.$$

More generally, for any $t_0 \in \mathbb{R}$, the solution of $(*)$ satisfies

$$X(t) \;=\; e^{A(t-t_0)}\left(X(t_0) + \int_{t_0}^t e^{A(t_0-s)}\,F(s)\,ds\right) = e^{A(t-t_0)}\,X(t_0) + \int_{t_0}^t e^{A(t-s)}\,F(s)\,ds.$$

Matrix exponential of commuting matrices

Proposition. Let $A, B \in \mathbb{C}^{n\times n}$, $t \in \mathbb{R}$.

$$AB = BA \quad \Rightarrow \quad e^{(A+B)\,t} = e^{A\,t}\,e^{B\,t} = e^{B\,t}\,e^{A\,t}.$$

Proof. From $AB = BA$ it follows that $A^k B = BA^k$, $k \in \mathbb{N}$, whence $e^{A\,t}\,B = B\,e^{A\,t}$.
Let $X(t) = e^{A\,t}\,e^{B\,t}$. Then $X(0) = I$,and by the product rule

$$\dot{X}(t) = A\,e^{A\,t}\,e^{B\,t} + e^{A\,t}\,B\,e^{B\,t} = A\,e^{A\,t}\,e^{B\,t} + B\,e^{A\,t}\,e^{B\,t} = (A+B)\,X(t).$$

By uniqueness of the solution of linear initial value problems: $X(t) = e^{(A+B)\,t}$.

Remark. If A and B do not commute,

$$e^{A\,t}\,e^{B\,t} = e^{Z(t)}, \quad Z(t) = (A+B)\,t + (AB - BA)\frac{t^2}{2} + \text{infinite series of commutators.}$$

(Baker-Campbell-Hausdorff formula)

The case of diagonalizable matrix

Proposition. Suppose that $A \in \mathbb{C}^{n\times n}$ has a basis $V = [v_1, \dots, v_n]$ of eigenvectors s.t.

$$A v_k = \lambda_k \, v_k. \qquad (*)$$

Then

$$\boxed{e^{At} = V \operatorname{diag}(e^{\lambda_1 t}, \dots, e^{\lambda_n t}) V^{-1} = \textstyle\sum_{k=1}^{n} e^{\lambda_k t} \, v_k w_k^\top, \qquad \text{where } V^{-1} = \begin{bmatrix} w_1^\top \\ \vdots \\ w_n^\top \end{bmatrix}}.$$

Remark. The row vectors $w_k^\top$ are the left eigenvectors $w_k^\top A = \lambda_k \, w_k^\top$.
The matrices $P_k = v_k \, w_k^\top \in \mathbb{C}^{n\times n}$ are called spectral projectors.

Proof. The relations $(*)$ can be written as

$$AV = V\Lambda, \qquad \text{where } \Lambda := \operatorname{diag}(\lambda_1, \dots, \lambda_n).$$

Since V is nonsingular: $A = V \Lambda V^{-1}$. Furthermore, we have

$$e^{\Lambda t} = \operatorname{diag}(e^{\lambda_1 t}, \dots, e^{\lambda_n t}).$$

Let $X(t) := V e^{\Lambda t} V^{-1}$. Then $X(0) = I$ and

$$\dot X(t) = V \Lambda e^{\Lambda t} V^{-1} = V \Lambda V^{-1} \, V e^{\Lambda t} V^{-1} = A \, X(t).$$

By uniqueness: $X(t) = e^{At}$.

The case $\lambda I + N$, N nilpotent

Let $N \in \mathbb{C}^{n\times n}$ be nilpotent of index ν, i.e. $N^{\nu-1} \neq 0 = N^{\nu}$. Then for any $\lambda \in \mathbb{C}$,

$$e^{(\lambda I+N)t} = e^{\lambda t}\left(I + Nt + N^2\frac{t^2}{2} + N^3\frac{t^3}{3!} + \ldots + N^{\nu-1}\frac{t^{\nu-1}}{(\nu-1)!}.\right) \qquad (*)$$

Proof. Let $X(t)$ denote the right hand side of $(*)$. Verify that $X(0) = I$, $\dot{X} = AX$.

Corollary. For a Jordan block,

$$J = \begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & & 1 \\ & & & & \lambda \end{bmatrix} \qquad \Rightarrow \qquad e^{Jt} = e^{\lambda t}\begin{bmatrix} 1 & t & \frac{t^2}{2} & \cdots & & \frac{t^{\ell-1}}{(\ell-1)!} \\ & 1 & t & & & \\ & & & & \ddots & \\ & & & \ddots & \ddots & \frac{t^2}{2} \\ & & & & 1 & t \\ & & & & & 1 \end{bmatrix}.$$

The case of a 2×2 matrix

For $A \in \mathbb{C}^{2\times 2}$ with distinct eigenvalues $\lambda_1, \lambda_2 \in \mathbb{C}$,

$$e^{At} = e^{\lambda_1 t}\frac{A-\lambda_2 I}{\lambda_1-\lambda_2} + e^{\lambda_2 t}\frac{A-\lambda_1 I}{\lambda_2-\lambda_1} = \frac{\lambda_2 e^{\lambda_1 t}-\lambda_1 e^{\lambda_2 t}}{\lambda_2-\lambda_1} I + \frac{e^{\lambda_1 t}-e^{\lambda_1 t}}{\lambda_2-\lambda_1} A.$$

Proof. Diagonalization or use methods explained later on.

For $A \in \mathbb{C}^{2\times 2}$ with only one eigenvalue λ,

$$e^{At} = e^{\lambda t}\left(I + (A-\lambda I)\,t\right).$$

Proof. $A - \lambda I$ is nilpotent.

Real 2×2 **Example.** Let $\alpha, \omega \in \mathbb{R}$.

$$A = \begin{bmatrix} \alpha & -\omega \\ \omega & \alpha \end{bmatrix} \qquad \Rightarrow \qquad e^{At} = \underbrace{e^{\alpha t}\begin{bmatrix} \cos(\omega t) & -\sin(\omega t) \\ \sin(\omega t) & \cos(\omega t) \end{bmatrix}}_{=:X(t)}.$$

.

Proof. Verify, that $X(0) = I$, $\dot{X}(t) = A\,X(t)$.

The general real 2×2 case. Let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathbb{R}^{2\times 2}.$$

The eigenvalues of A are the zeros of its characteristic polynomial

$$\chi(\lambda) = \lambda^2 - \frac{a+b}{2}\lambda + (ad - bc),$$

i.e.

$$\lambda_\pm = \frac{(a+b) \pm \sqrt{disc}}{2}, \qquad disc := (a-d)^2 + 4bc.$$

Define $\omega = \frac{1}{2}\sqrt{|disc|}$. Then

if $disc > 0$ (real simple eigenvalues)

$$e^{At} = e^{\frac{(a+d)t}{2}} \begin{bmatrix} \cosh(\omega\, t) + \frac{a-d}{2}\,\frac{\sinh(\omega\, t)}{\omega} & b\,\frac{\sinh(\omega\, t)}{\omega} \\ c\,\frac{\sinh(\omega\, t)}{\omega} & \cosh(\omega\, t) - \frac{a-d}{2}\,\frac{\sinh(\omega\, t)}{\omega} \end{bmatrix},$$

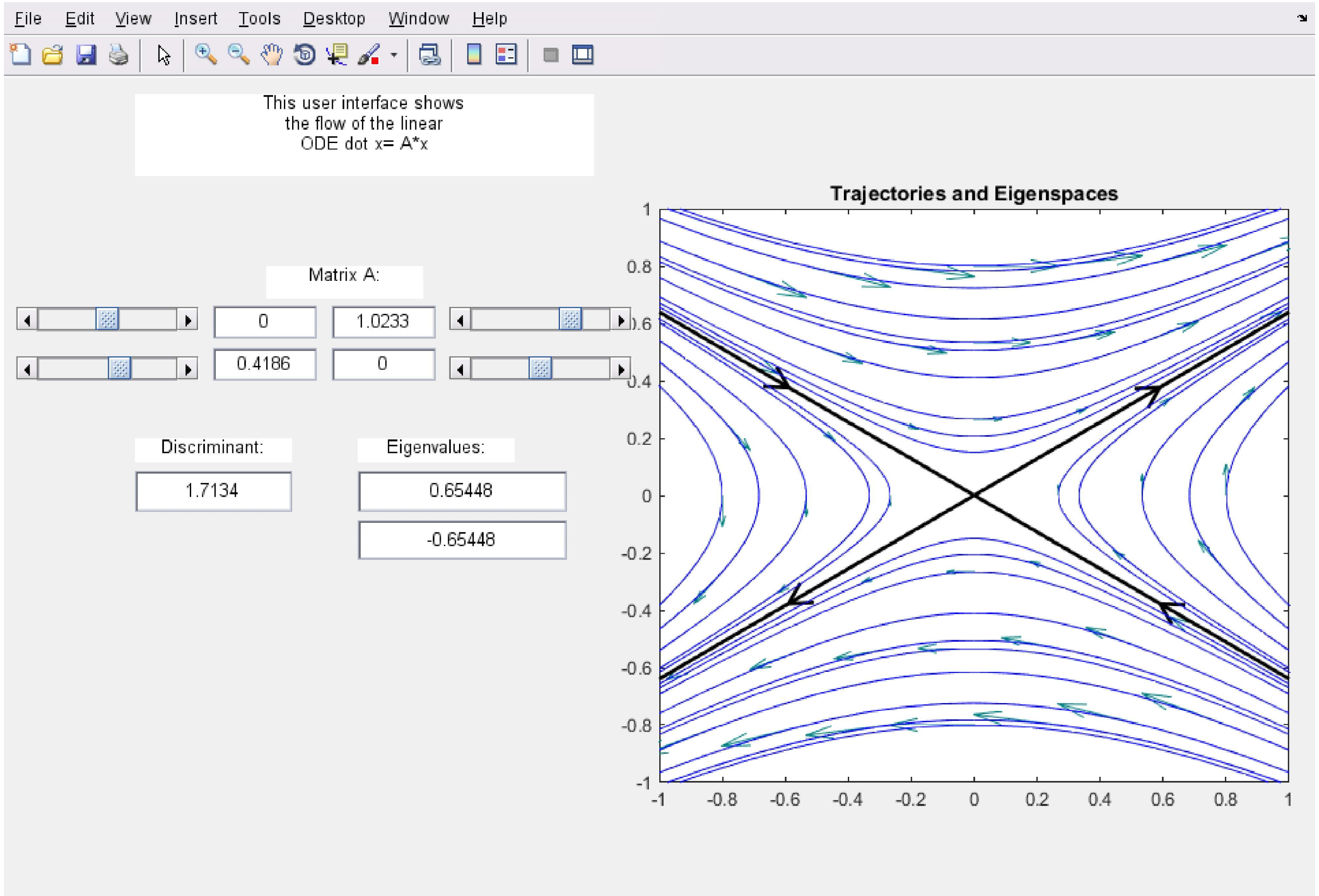
if $disc < 0$ (nonreal simple eigenvalues)

$$e^{At} = e^{\frac{(a+d)t}{2}} \begin{bmatrix} \cos(\omega\, t) + \frac{a-d}{2}\,\frac{\sin(\omega\, t)}{\omega} & b\,\frac{\sin(\omega\, t)}{\omega} \\ c\,\frac{\sin(\omega\, t)}{\omega} & \cos(\omega\, t) - \frac{a-d}{2}\,\frac{\sin(\omega\, t)}{\omega} \end{bmatrix},$$

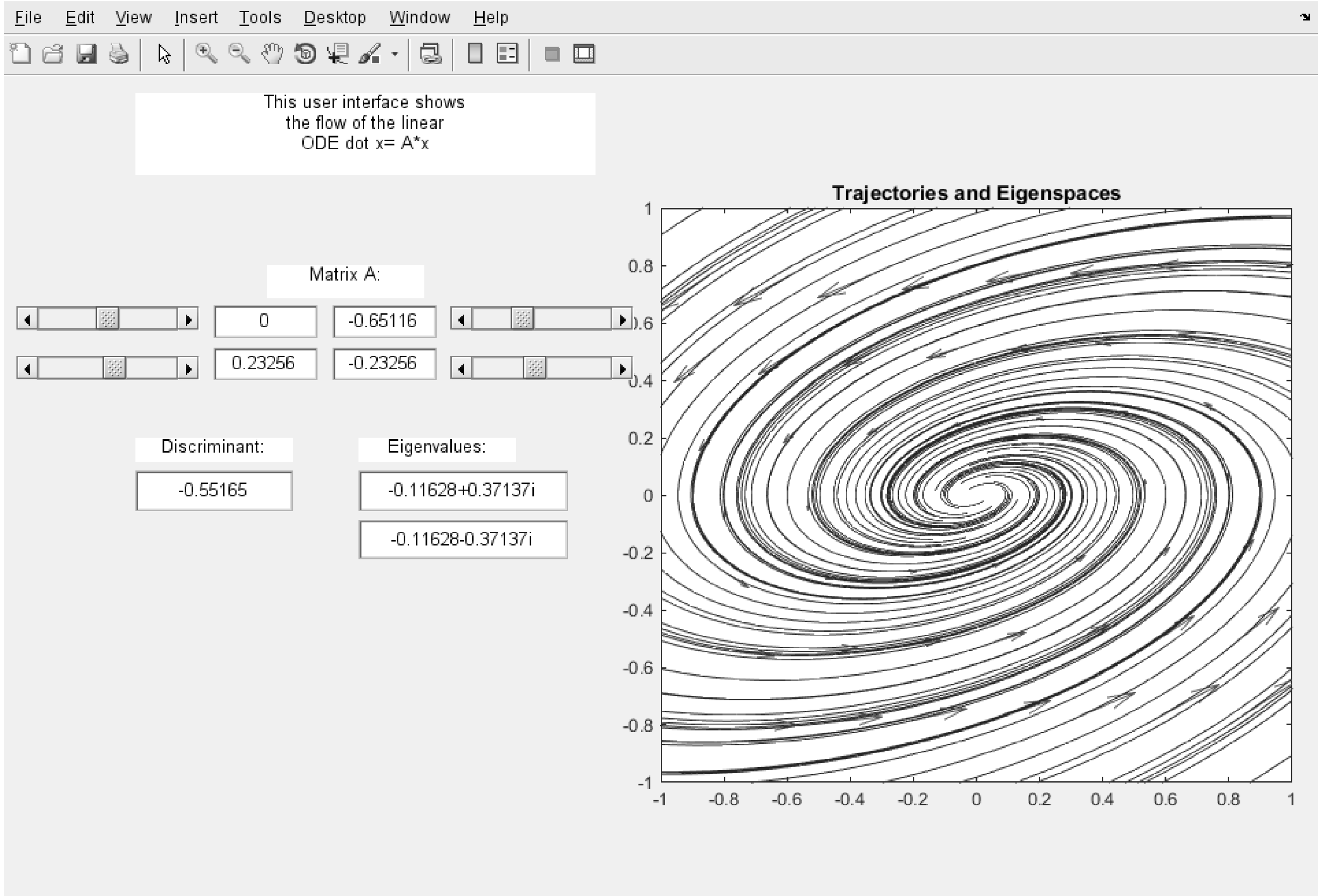
if $disc = 0$ (equal eigenvalues)

$$e^{At} = e^{\frac{(a+d)t}{2}} \begin{bmatrix} 1 + \frac{(a-d)t}{2} & b\,t \\ c\,t & 1 - \frac{(a-d)t}{2} \end{bmatrix}.$$

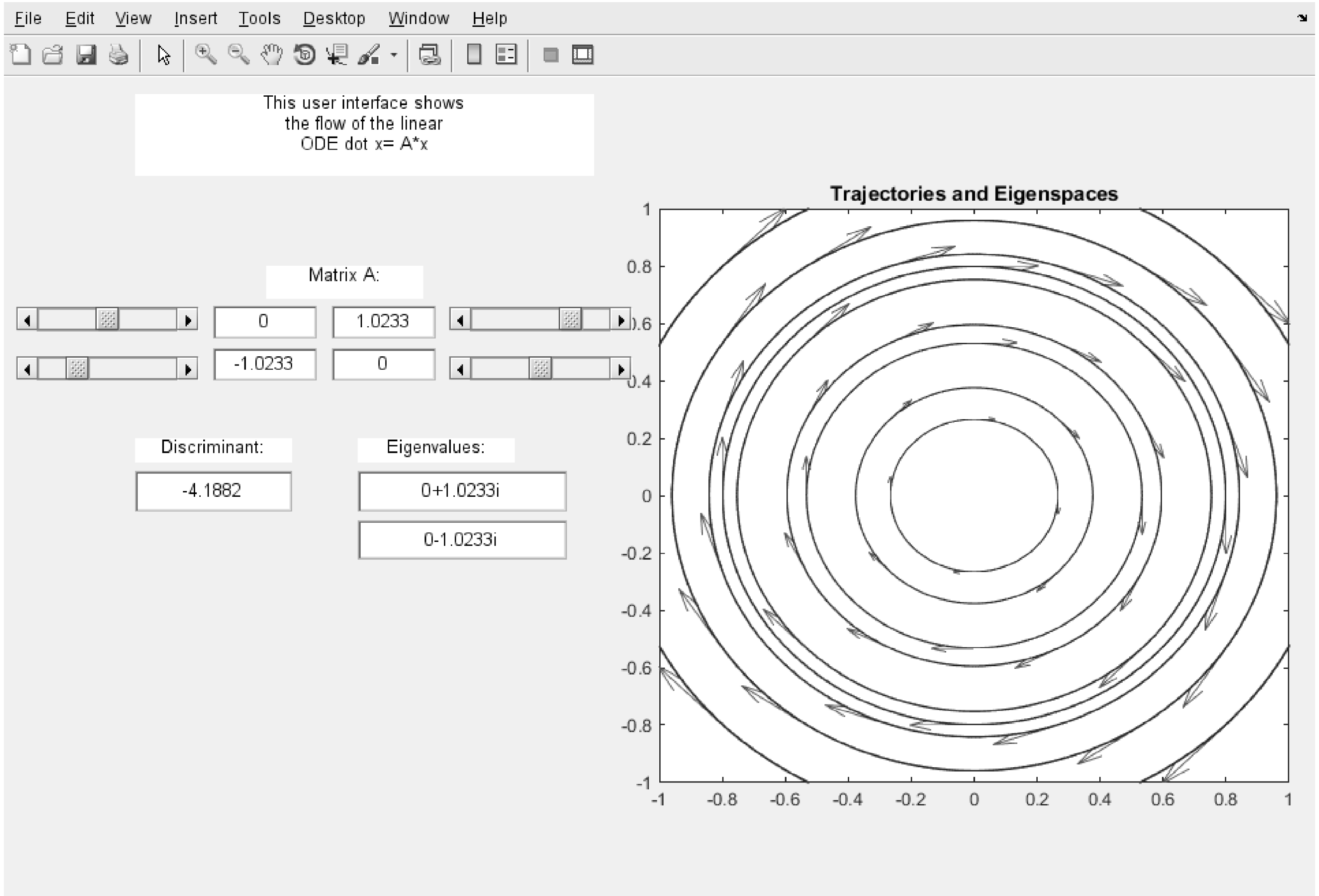
Phase portrait of $\dot{x} = Ax$, where $A \in \mathbb{R}^{2\times 2}$



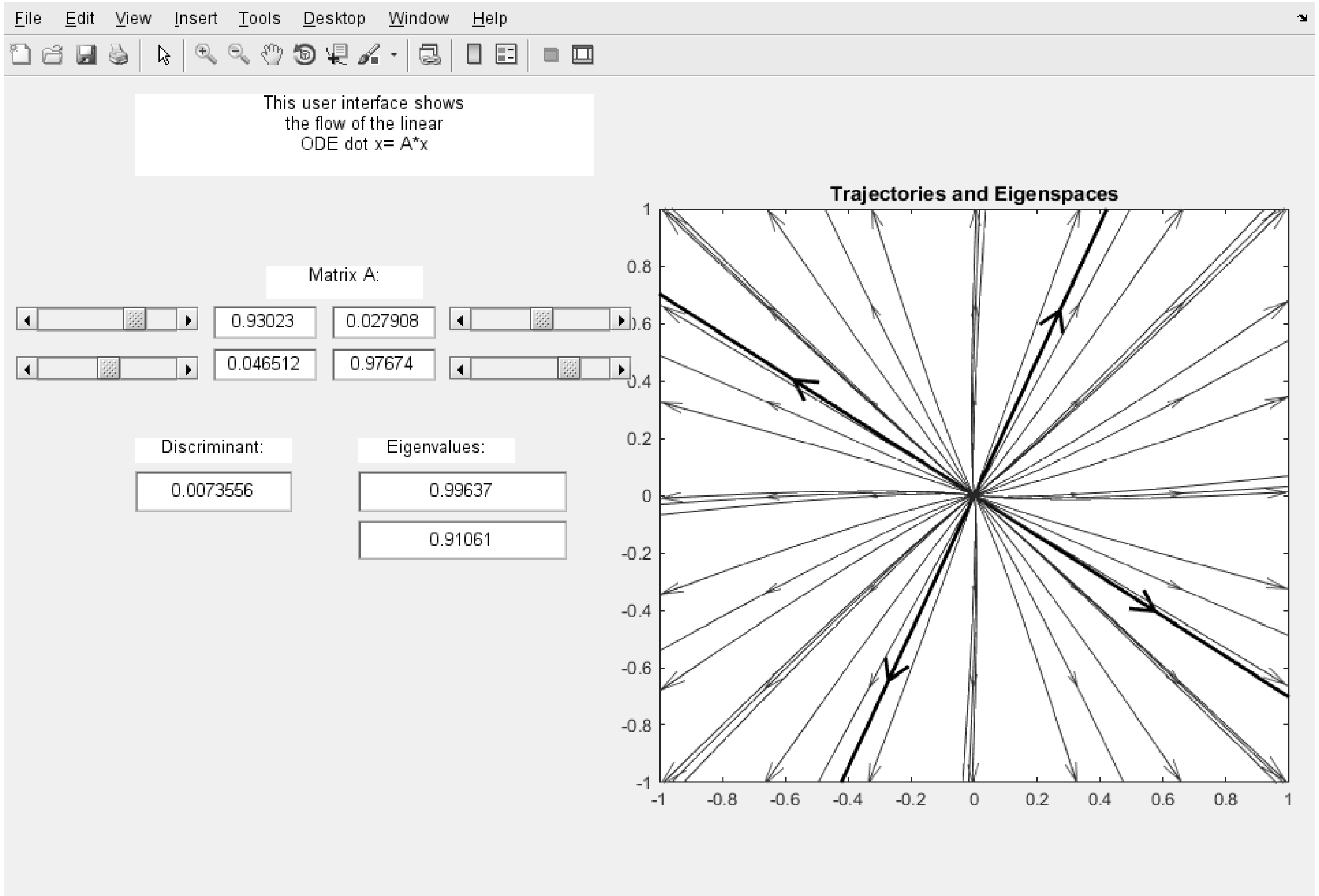
Phase portrait of $\dot{x} = Ax$, where $A \in \mathbb{R}^{2\times 2}$



Phase portrait of $\dot{x} = Ax$, where $A \in \mathbb{R}^{2\times 2}$



Phase portrait of $\dot{x} = Ax$, where $A \in \mathbb{R}^{2\times 2}$



The scalar Riccati differential equation.

Suppose

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \qquad (a, b, c, d \text{ may be time dependent})$$

Set

$$s(t) := \frac{y(t)}{x(t)} \qquad \text{(slope of the vector).}$$

Then

$$\begin{aligned} \dot{s} &= \frac{\dot{y}\,x - y\,\dot{x}}{x^2} \\ &= \frac{(cx + dy)\,x - y\,(ax + by)}{x^2} \\ &= c + (d - a)\frac{y}{x} - b\,\frac{y^2}{x^2} \\ &= c + (d - a)\,s - b\,s^2. \end{aligned}$$

This is a quadratic ODE for s, called **Riccati ODE**.
Note that the solution may have finite escape time ($x(t)$ may become 0 for some t).

The solution of the Riccati ODE can be found via the associated linear ODE.

The scalar time variant case

The solutions of the scalar homogeneous ODE

$$\dot{x}(t) = a(t)\,x(t)$$

satisfy

$$x(t) = x(t_0)\,e^{A(t)}, \qquad \text{where } A(t) = \int_{t_0}^{t} a(s)\,ds.$$

Proof. Direct verification, since $\dot{A}(t) = a(t)$.

(Simplified) Gronwall inequality: Suppose

$$\dot{x}(t) \le a(t)\,x(t)$$

Then for $t \ge t_0$,

$$x(t) \le x(t_0)\,e^{A(t)}.$$

Proof. We have

$$\frac{d}{dt}\left(x(t)\,e^{-A(t)}\right) = \dot{x}(t)e^{-A(t)} - x(t)\,a(t)\,e^{-A(t)} \le 0.$$

Thus

$$x(t)\,e^{-A(t)} \le x(t_0)\,e^{-A(t_0)} = x(t_0).$$

Digression: invariant subspaces

Let $A : \mathcal{V} \to \mathcal{V}$ be a linear operator on the vector space $\mathcal{V}$.

A subspace $\mathcal{U}$ is said to be A-**invariant** if

$$A\mathcal{U} \subset \mathcal{U}.$$

Examples. kernel and image of A are A-invariant:

$\operatorname{Ker} A := \{ v \mid Av = 0 \}$. $\quad A \operatorname{Ker} A \subset \operatorname{Ker} A$

$\operatorname{Im} A := \{ Av \mid v \in \mathcal{V} \}$. $\quad A \operatorname{Im} A \subset \operatorname{Im} A$.

Suppose $U = [u_1, \dots u_r]$ is a basis of the invariant subspace $\mathcal{U}$. Then

$$A\, u_k = \sum_{j=1}^{r} \ell_{jk}\, u_j = \sum_{j=1}^{r} u_j\, \ell_{jk} \qquad k = 1, \dots r,$$

where ℓ_{jk} are elements of the underlying field. Let $L = [\ell_{jk}]$. Then, in matrix notation,

$$AU = UL.$$

L is said to be the representation of A with respect to the basis U.

Invariant subspaces and linear ODE

Suppose

$$AU = UL \qquad \text{where } A, U, L \text{ are complex matrices.}$$

Then for any $k = 1, 2, \ldots,$

$$A^k U = U\ L^k.$$

Summation yields

$$e^{At}U\ = U\, e^{Lt}.$$

Suppose $x_0 \in \operatorname{Im} U$, i.e. $x_0 = U\xi$ for some ξ. Then

$$e^{At}x_0 = e^{At}U\xi = U\, e^{Lt}\xi\ \in \operatorname{Im} U$$

Thus:
if the solution to the homgeneous linear ODE $\dot{x} = Ax$ starts in an A-invariant subspace, then it stays in that subspace for all time.

Note: The 1-dimensional invariant subspaces of A are its eigenspaces.

Linear operators

and

polynomials

Linear operators and polynomials

In the sequel $\mathbb{K}[x]$ denotes the set of polynomials

$$p(x) = p_n\,x^n + p_{n-1}\,x^{n-1} + \ldots + p_1\,x + p_0$$

with coefficients in the field $\mathbb{K}$. $\mathcal{V}$ is a vector space over $\mathbb{K}$, and $\mathcal{L}(\mathcal{V},\mathcal{V})$ is the set of linear operators $A : \mathcal{V} \to \mathcal{V}$. If $\mathcal{V}$ is finite dimensional A is identified with its matrix with respect to the standard basis. Definition:

$$p(A) = p_n\,A^n + p_{n-1}\,A^{n-1} + \ldots + p_1\,A + p_0\,I$$

Then

(1) $1(A) = I, \qquad x(A) = A,$

(2) $(p \pm q)(A) = p(A) \pm q(A), \qquad (pq)(A) = p(A)q(A) = q(A)p(A),$

(3) if $AB = BA$ then for any $p \in \mathbb{K}[X]$, $\operatorname{Ker} p(A)$ and $\operatorname{Im} p(A)$ are B-invariant.

Proof of (3).

$0 = p(A)v \;\Rightarrow\; 0 = Bp(A)v = p(A)Bv. \qquad y = p(A)v \;\Rightarrow\; By = Bp(A)v = p(A)Bv.$

Linear operators and polynomials

Definition: A nonempty subset $\mathcal{I} \subset \mathbb{K}[x]$ is called an **Ideal** if

(1) $p, q \in \mathcal{I} \quad \Rightarrow \quad p \pm q \in \mathcal{I}$, (2) $p \in \mathcal{I} \quad \Rightarrow \quad \alpha\, p \in \mathcal{I}$ for all $\alpha \in \mathbb{K}[x]$.

Basic facts:

(a) Let $\mathcal{I}$ be an ideal. Let $p \in \mathcal{I}$ be a polynomial of minimum degree. Then

$$\mathcal{I} = \mathbb{K}[x]\, p = \{\, \alpha\, p \mid \alpha \in \mathbb{K}[x]\,\}.$$

(b) Let $\mathcal{I} = \left\{ \sum_{k=1}^{r} \alpha_k\, p_k \;\middle|\; \alpha_k \in \mathbb{K}[x] \right\}$ be the ideal generated by $p_1, \ldots, p_r \in \mathbb{K}[x]$.
Let $d \in \mathbb{K}[x]$ be a greatest common divisor of the polynomials p_k. Then

$$\mathcal{I} = \mathbb{K}[x]\, d$$

In particular, there are polynomials α_k such that

$$d = \sum_{k=1}^{r} \alpha_k\, p_k. \qquad \text{(Bezout identity)} \qquad (*)$$

Proof. (a) A polynomial $p \in \mathcal{I}$ of mininum degree in $\mathcal{I}$ divides any other polynomial $q \in \mathcal{I}$ since otherwise we could write $q = \alpha\, p + r$ with degree(r) <degree(p), $r \neq 0$. $\Rightarrow \quad r = q - \alpha\, p \in \mathcal{I}$. Contradiction.
(b) By (a) a polynomial $d \in \mathcal{I}$ of minimum degree divides any polynomial in $\mathcal{I}$. In particular, it is a common divisor of the p_k. From $(*)$ it follows that any common divisor of the p_k divides d.

Remark: The factors α_k in $(*)$ can be computed by the Euclidean algorithm.

Linear operators and polynomials

Kernel intersection Lemma: Let d be the greatest common divisor of $p, q \in \mathbb{K}[x]$. Then for $A \in \mathcal{L}(\mathcal{V}, \mathcal{V})$,

$$\operatorname{Ker} d(A) = \operatorname{Ker} p(A) \cap \operatorname{Ker} q(A).$$

Proof. There exist $\alpha, \beta \in \mathbb{K}[x]$ such that $d = \alpha\, p + \beta\, q$ (Bezout-identity) . Thus

$$d(A) = \alpha(A)\, p(A) + \beta(A)\, q(A).$$

Thus, $p(A)v = q(A)v = 0$ implies $d(A)v = 0$, whence $\operatorname{Ker} p(A) \cap \operatorname{Ker} q(A) \subset \operatorname{Ker} d(A)$.

On the other hand let $\gamma\, d = p$. Then $\gamma(A)\, d(A) = p(A)$. So $d(A)v = 0$ implies $p(A)v = 0$, whence $\operatorname{Ker} d(A) \subset \operatorname{Ker} p(A)$.

Linear operators and polynomials

Kernel decomposition theorem.

Let $p=\prod_{k=1}^{r} p_k$, where $p_1,\dots p_r \in \mathbb{K}[x]$ are pairwise coprime. Then for $A\in\mathcal{L}(\mathcal{V},\mathcal{V})$,

(1) there exist $\pi_k \in \mathbb{K}[x]$ such that $v\in \operatorname{Ker} p_j(A)$ implies $\pi_k(A)v = \begin{cases} v & \text{if } j=k \\ 0 & \text{otherwise.}\end{cases}$,

(2) $\operatorname{Ker} p(A) = \bigoplus_{k=1}^{r} \operatorname{Ker} p_k(A)$.

Proof: The polynomials $\ell_k = \prod_{j=1,j\neq k}^{r} p_j$, $k=1,\dots,r$ have no common factor. Thus, there exist $f_k\in\mathbb{K}[x]$ such that

$$1=\sum_{k=1}^{r}\overbrace{\alpha_k \prod_{j=1,j\neq k}^{r} p_j}^{=:\pi_k}.$$

It follows that for all $v\in\mathcal{V}$,

$$v = I\,v = \sum_k \pi_k(A)v. \qquad (*)$$

(1) Each π_k, $j\neq k$, contains the factor p_j. Thus $v\in \operatorname{Ker} p_j(A)$ implies $\pi_k(A)v=0$ for $j\neq k$, and then by $(*)$, $v=\pi_j(A)v$.

(2) We have $p_k(A)\pi_k(A) = \alpha_k(A)\,p(A)$ Thus $v\in \operatorname{Ker} p(A)$ implies $\pi_k(A)v\in \operatorname{Ker} p_k(A)$. Thus, $\operatorname{Ker} p(A)$ is the sum of the subspaces $\operatorname{Ker} p_j(A)$. The sum is direct: applying $\pi_k(A)$ on the relation

$$0=\sum_j v_j, \qquad v_j\in \operatorname{Ker} p_j(A)$$

yields $v_k=0$.

Linear operators and polynomials

Kernel decomposition theorem.

Let $A \in \mathcal{L}(\mathcal{V}, \mathcal{V})$, where $\mathcal{V}$ is vector space over an arbitrary field $\mathbb{K}$.

Let $p = \prod_{k=1}^{r} p_k$, where $p_1, \ldots p_r \in \mathbb{K}[x]$ are pairwise coprime.

(1) Then there exist $\alpha_k \in \mathbb{K}[x]$ such that

$$1 = \sum_{k=1}^{r} \alpha_k \overbrace{\prod_{j=1, j\neq k}^{r} p_j}^{=:\pi_k} . \qquad (*)$$

(2) The operators $\pi_k(A)$ satisfy:

$$(a)\ \ I = \sum_{k=1}^{r} \pi_k(A), \qquad (b)\ \ v \in \operatorname{Ker} p_k(A) \ \text{ implies } \ \pi_i(A)v = \begin{cases} v & \text{if } i = k \\ 0 & \text{otherwise.} \end{cases}$$

(3) $\operatorname{Ker} p(A) = \bigoplus_{k=1}^{r} \operatorname{Ker} p_k(A)$.

<u>Proof:</u> The polynomials $\ell_k = \prod_{j=1, j\neq k}^{r} p_j$, $k = 1, \ldots, r$ have no common factor. $\Rightarrow$ (1).

(2a) is immediate from (1). (2b). Each π_i, $i \neq k$, contains the factor p_k.
Thus $v \in \operatorname{Ker} p_k(A)$ implies $\pi_i(A)v = 0$ for $i \neq k$, and then by (2a), $v = \pi_k(A)v$.

(3) We have $p_k(A)\pi_k(A) = \alpha_k(A)\, p(A)$. Thus, $v \in \operatorname{Ker} p(A)$ implies $\pi_k(A)v \in \operatorname{Ker} p_k(A)$.
Thus, $\operatorname{Ker} p(A)$ is the sum of the subspaces $\operatorname{Ker} p_j(A)$. The sum is direct:
applying $\pi_k(A)$ on the relation $0 = \sum_i v_i$ with $v_i \in \operatorname{Ker} p_i(A)$ yields $v_k = 0$.

Linear operators and polynomials

Remarks on the polynomial identity

$$1 = \sum_{k=1}^{r} \alpha_k \overbrace{\prod_{j=1, j\neq k}^{r} p_j}^{=:\ell_k} .$$

(1) This identity is equivalent to the incomplete partial fraction expansion

$$\frac{1}{\prod_{k=1}^{r} p_k} = \sum_{k=1}^{r} \frac{\alpha_k}{p_k}.$$

(2) If $p_k(x) = (x - \lambda_k)^{m_k}$, $k = 1, \ldots, r$ with distinct $\lambda_k \in \mathbb{C}$,
then $\ell_k(x) = \prod_{j=1, j\neq k}^{r} (x - \lambda_j)^{m_j - 1}$, and

$$\alpha_k(x) = \sum_{j=0}^{m_k - 1} \frac{1}{j!} \left(\frac{d^j}{dx^j} \left. \frac{1}{\ell_k(x)} \right|_{x=\lambda_k} \right) (x - \lambda_k)^j. \qquad \text{(Taylor expansion of } 1/\ell_k(x)\text{)}$$

in particular $\deg(\ell_k) \leq m_k - 1$, and

$$\alpha_k(x) \equiv const = 1 \Big/ \prod_{j=1, j\neq k}^{r} (\lambda_k - \lambda_j)^{m_j - 1} \qquad \text{if } m_k = 1.$$

Proof. see Wikipedia (Partial Fraction Expansion)

Digression: scalar ODE of higher order with constant coefficients

Application of kernel decomposition theorem:
Solutions of homogeneous linear ODE of higher order with constant coefficients

Let $p(x) = x^n + \sum_{j=0}^{n-1} a_j\, x^j = \prod_{k=1}^{r}(x-\lambda_k)^{m_k}$, where the $\lambda_k \in \mathbb{C}$ are distinct.

All solutions of the scalar ODE

$$0 = p\left(\frac{d}{dt}\right) y(t) = y^{(n)}(t) + p_{n-1}\, y^{(n-1)}(t) + \ldots + p_1\, \dot{y}(t) + p_0\, y(t).$$

are given by

$$y(t) = \sum_{k=1}^{r} q_k(t)\, e^{\lambda_k t}, \qquad q_k \text{ polynomial of degree} \leq m_k - 1. \qquad (*)$$

<u>Proof.</u> By the kernel decomposition theorem, $\quad \mathrm{Ker}\, p\left(\frac{d}{dt}\right) = \bigoplus_{k=1}^{r} \mathrm{Ker}\left(\frac{d}{dt} - \lambda_k\right)^{m_k}$.

Any differentiable function y_k can be written in the form $y_k(t) = q_k(t)\, e^{\lambda_k t}$. We have

$$\left(\frac{d}{dt} - \lambda_k\right) y_k(t) = \dot{q}_k(t) e^{\lambda_k t}, \qquad \Rightarrow \qquad \left(\frac{d}{dt} - \lambda_k\right)^{m_k} y_k(t) = q_k^{(m_k)}(t) e^{\lambda_k t}.$$

Thus,

$$y_k \in \mathrm{Ker}\left(\frac{d}{dt} - \lambda_k\right)^{m_k} \quad \Leftrightarrow \quad q_k^{(m_k)} \equiv 0 \quad \Leftrightarrow \quad q_k \text{ is polynomial of degree} \leq m_k - 1.$$

Terminology: functions of the form $(*)$ are called **Bohl functions**.

Digression: scalar ODE of higher order with constant coefficients

Suplement:

solution formula for inhomogeneous initial value problem of higher order

Let $\phi_0, \ldots, \phi_{n-1} : \mathbb{R} \to \mathbb{C}$ be the solutions of the homogeneous initial value problem

$$\phi_k^{(n)}(t) + p_{n-1}\,\phi_k^{(n-1)}(t) + \ldots + p_1\,\dot{\phi}_k(t) + p_0\,\phi_k(t) = 0, \qquad \phi_k^{(j)}(0) = \delta_{jk}, \quad j = 0, \ldots, n-1$$

(Bohl functions). Then

$$y : \mathbb{R} \to \mathbb{C}, \qquad y(t) = \sum_{k=0}^{n-1} \phi_k(t-t_0)\,y_k + \int_{t_0}^{t} \phi_{n-1}(t-s)\,f(s)\,ds$$

is the unique solution of the initial value problem

$$y^{(n)}(t) + p_{n-1}\,y^{(n-1)}(t) + \ldots + p_1\,\dot{y}(t) + p_0\,y(t) = f(t), \qquad y^{(k)}(t_0) = y_k.$$

Question: Is there a nice way for computing the functions ϕ_k ?

Answer: Yes. Method will be shown after some remarks on the next slides.

Digression: scalar ODE of higher order with constant coefficients

Example to the solution formula

Consider the ODE of second order (harmonic oscillator):

$$\ddot{y}(t) + \omega^2\, y(t) = f(t), \qquad \omega > 0. \qquad (*)$$

The functions

$$\phi_0(t) = \cos(\omega\, t), \qquad \phi_1(t) = \omega^{-1} \sin(\omega\, t)$$

solve the homogeneous ODE

$$\ddot{\phi}(t) + \omega^2\, \phi(t) = 0$$

with initial conditions $\phi_k^{(j)} = \delta_{jk}$. Thus, the solutions of $(*)$ are

$$y(t) = y(t_0)\, \cos(\omega\,(t - t_0)) + \dot{y}(t_0)\, \omega^{-1} \sin(\omega\,(t - t_0)) + \int_{t_0}^{t} \omega^{-1} \sin(\omega\,(t - s))\, f(s)\, ds.$$

Digression: scalar ODE of higher order with constant coefficients

Remark on the integral part of the solution formula

For a cont. function $g(t,s)$ which is continuosly differentiable in t the following holds:

$$\frac{d}{dt}\int_{t_0}^{t} g(t,s)\,ds = g(t,t) + \int_{t_0}^{t} \frac{d}{dt} g(t,s)\,ds.$$

Thus,

$$\begin{aligned}
\frac{d}{dt}\int_{t_0}^{t} \phi_{n-1}(t-s)\,f(s)\,ds &= \overbrace{\phi_{n-1}(t-t)}^{=0}\,f(t) + \int_{t_0}^{t} \dot\phi_{n-1}(t-s)\,f(s)\,ds \\
\frac{d^2}{dt^2}\int_{t_0}^{t} \phi_{n-1}(t-s)\,f(s)\,ds &= \frac{d}{dt}\int_{t_0}^{t} \dot\phi_{n-1}(t-s)\,f(s)\,ds \\
&= \underbrace{\dot\phi_{n-1}(t-t)}_{=0}\,f(t) + \int_{t_0}^{t} \ddot\phi_{n-1}(t-s)\,f(s)\,ds \\
&\;\;\vdots \\
\frac{d^n}{dt^n}\int_{t_0}^{t} \phi_{n-1}(t-s)\,f(s)\,ds &= \frac{d}{dt}\int_{t_0}^{t} \phi_{n-1}^{(n-1)}(t-s)\,f(s)\,ds \\
&= \underbrace{\phi_{n-1}^{(n-1)}(t-t)}_{=1}\,f(t) + \int_{t_0}^{t} \phi_{n-1}^{(n)}(t-s)\,f(s)\,ds
\end{aligned}$$

Summing up:

$$\sum_{k=0}^{n} p_k\,\frac{d^k}{dt^k}\int_{t_0}^{t} \phi_{n-1}(t-s)\,f(s)\,ds = f(t) + \int_{t_0}^{t} \underbrace{\sum_{k=1}^{n} p_k\,\phi_{n-1}^{(k)}(t-s)}_{=0}\;f(s)\,ds$$

Digression: scalar ODE of higher order with constant coefficients

The basis solutions ϕ_k can be computed in the following way.

Step 1. Perform partial fraction expansion

$$\frac{1}{p(x)} = \frac{1}{\prod_{k=1}^{r}(x-\lambda_k)^{m_k}} = \sum_{k=1}^{r}\sum_{j=1}^{m_k}\frac{\alpha_{kj}}{(x-\lambda_k)^j}.$$

Set

$$\phi_{n-1}(t) := \sum_{k=1}^{r} e^{\lambda_k t}\sum_{j=1}^{m_k}\alpha_{kj}\,\frac{t^{j-1}}{(j-1)!}.$$

Step 2. Set

$$\begin{aligned}
\phi_{n-2} &:= p_{n-1}\,\phi_{n-1} + \dot{\phi}_{n-1},\\
\phi_{n-3} &:= p_{n-2}\,\phi_{n-1} + \dot{\phi}_{n-2},\\
&\;\vdots\\
\phi_0 &:= p_1\,\phi_{n-1} + \dot{\phi}_1.
\end{aligned}$$

Then $p(\frac{d}{dt})\phi_k = 0$ and $\phi_k^{(j)}(0) = \delta_{jk}$.

Proof. The function ϕ_{n-1} defined in Step 1 solves the ODE by the kernel decomposition theorem. The conditions $\phi_{n-1}^{(j)}(0) = \delta_{j,n-1}$ can be verified by direct computation or by Laplace transformation. Obviously, the functions of Step 2 solve the ODE. The initial condition properties follow by differentiating $p(\frac{d}{dt})\phi_{n-1} = 0$ and setting $t = 0$.

Digression: scalar ODE of higher order with constant coefficients

Example to the solution formula method for ODE of higher order

Let

$$p(x) = x^3 + p_2\,x^2 + p_1\,x + p_0 = (x-\lambda_1)(x-\lambda_2)(x-\lambda_3), \qquad \lambda_k \in \mathbb{C} \text{ distinct.}$$

Partial fraction expansion yields

$$\frac{1}{p(x)} = \frac{a_1}{x-\lambda_1} + \frac{a_2}{x-\lambda_2} + \frac{a_3}{x-\lambda_3},$$

where

$$1/a_1 = (\lambda_1-\lambda_2)(\lambda_1-\lambda_3), \quad 1/a_2 = (\lambda_2-\lambda_1)(\lambda_2-\lambda_3), \quad 1/a_3 = (\lambda_3-\lambda_1)(\lambda_3-\lambda_2).$$

Thus, the general solution of

$$y^{(3)} + p_2\,\ddot{y} + p_1\,\dot{y} + p_0\,y = f$$

is

$$y(t) = \phi_0(t-t_0)\,y(t_0) + \phi_1(t-t_0)\,\dot{y}(t_0) + \phi_2(t-t_0)\,\ddot{y}(t_0) + \int_{t_0}^{t} \phi_2(t-s)\,f(s)\,ds,$$

where

$$\begin{aligned}
\phi_2(t) &= a_1\,e^{\lambda_1 t} + a_2, e^{\lambda_2 t} + a_3\,e^{\lambda_3 t} \\
&= \frac{e^{\lambda_1 t}}{(\lambda_1-\lambda_2)(\lambda_1-\lambda_3)} + \frac{e^{\lambda_2 t}}{(\lambda_2-\lambda_1)(\lambda_2-\lambda_3)} + \frac{e^{\lambda_3 t}}{(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)}, \\
\phi_1(t) &= p_2\,\phi_2(t) + \dot{\phi}_2(t), \\
\phi_0(t) &= p_1\,\phi_2(t) + \dot{\phi}_1(t).
\end{aligned}$$

Matrix exponential and ODE of higher order

Proposition. Suppose $A \in \mathbb{C}^{N\times N}$ satisfies the polynomial identity

$$0 = p(A) = A^n + p_{n-1}\, A^{n-1} + \ldots + p_1\, A + p_0\, I. \qquad \text{(We may have } N \neq n\text{)}$$

Then

$$e^{At} = \sum_{k=0}^{n-1} \phi_k(t)\, A^k,$$

where the scalar functions $\phi_k : \mathbb{R} \to \mathbb{C}$ satisfy

$$p\left(\frac{d}{dt}\right) \phi_k = 0, \qquad \phi_k^{(j)}(0) = \delta_{jk}, \qquad j, k = 0, \ldots, n-1.$$

<u>Proof.</u> We have $\left(\frac{d}{dt}\right)^k e^{At} = A^k\, e^{At}$. Thus, $p\left(\frac{d}{dt}\right) e^{At} = p(A)\, e^{At} = 0$.

Thus (by our result on ODE of higher order):

$$e^{At} = \sum_{k=0}^{n-1} \phi_k(t)\, Y_k \qquad \text{with the matrices} \quad Y_k = \left(\frac{d}{dt}\right)^k e^{At}\Big|_{t=0} = A^k.$$

Corollary (Solution formula without matrix exponential).

$$\dot{x}(t) = A\, x(t) \quad \Rightarrow \quad x(t) = \textstyle\sum_{k=0}^{n-1} \phi_k(t-t_0)\, A^k x(t_0) \quad \text{and} \quad p\left(\frac{d}{dt}\right) x(t) = 0.$$

Matrix exponential with respect to polynomial basis I

The theorem below generalizes the statement from the slide before.

Representation theorem. Suppose $A \in \mathbb{C}^{N\times N}$ satisfies the polynomial identity

$$0 = p(A) = A^n + p_{n-1}\,A^{n-1} + \ldots + p_1\,A + p_0\,I. \qquad \text{(We may have } N \neq n\text{)}$$

Let $\beta_0(x), \beta_1(x), \ldots, \beta_{n-1}(x)$ be a basis of the set of polynomials of degree $\leq n-1$.
Let $a_k, c_k, m_{ik} \in \mathbb{C}$ be coefficients such that (polynomial division)

$$x\,\beta_k(x) = a_k\,p(x) + \sum_{i=0}^{n-1} m_{ik}\,\beta_i(x), \qquad k = 0, \ldots, n-1,$$

and $\sum_{k=0}^{n-1} c_k\,\beta_k(x) = 1$. Then,

$$e^{A\,t} = \sum_{k=0}^{n-1} \psi_k(t)\;\beta_k(A), \qquad (*)$$

where the scalar functions ψ_k solve the initial value problem

$$\frac{d}{dt}\begin{bmatrix} \psi_0(t) \\ \psi_1(t) \\ \vdots \\ \psi_{n-1}(t) \end{bmatrix} = \underbrace{\begin{bmatrix} m_{00} & \ldots & m_{0,n-1} \\ \vdots & & \vdots \\ \vdots & & \vdots \\ m_{n-1,0} & \ldots & m_{n-1,n-1} \end{bmatrix}}_{=:M_\beta} \begin{bmatrix} \psi_0(t) \\ \psi_1(t) \\ \vdots \\ \psi_{n-1}(t) \end{bmatrix}, \qquad \begin{bmatrix} \psi_0(0) \\ \psi_1(0) \\ \vdots \\ \psi_{n-1}(0) \end{bmatrix} = \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{n-1} \end{bmatrix}.$$

Proof. Show that the right hand side of $(*)$ solves the same initial value problem as $e^{A\,t}$.

Matrix exponential with respect to polynomial basis II

Special cases of the representation theorem.

(1) For $\beta_k(x) = x^k$, $k = 0, \dots, n-1$, we have

$$x\,\beta_k(x) = \beta_{k+1}(x), \;\; k = 0, \dots, n-2, \qquad x\,\beta_{n-1}(x) = x^n = p(x) - \sum_{i=0}^{i-1} p_i\,\beta_i(x).$$

Thus,

$$M_\beta = \begin{bmatrix} 0 & & & -p_0 \\ 1 & & & -p_1 \\ & \ddots & & \vdots \\ & & 1 & -p_{n-1} \end{bmatrix}. \qquad \text{(companion matrix)}$$

In this case the functions ψ_k equal the function ϕ_k from our considerations of ODE of higher order.

(2) Let $\lambda_1, \dots, \lambda_n$ be the zeros of p counting multiplicities. For the (Newton)basis

$$\beta_0(x) = 1, \quad \beta_1(x) = x - \lambda_1, \quad \beta_2 = (x-\lambda_1)(x-\lambda_2), \; \dots, \; \beta_{n-1}(x) = \prod_{j=1}^{n-1}(x - \lambda_j)$$

we have $1 = \beta_0(x) = \beta_0(x) + \sum_{k \geq 1} 0\,\beta_k(x) \;\Rightarrow\; \psi_0(0) = 1, \psi_k(0) = 0$ for $k \geq 1$, and

$$x\,\beta_k(x) = \beta_{k+1}(x) + \lambda_k\,\beta_k(x), \quad k = 0, \dots, n-2, \qquad x\,\beta_{n-1}(x) = p(x) + \lambda_n\,\beta_{n-1}(x).$$

In this case

$$M_\beta = \begin{bmatrix} \lambda_1 & & & \\ 1 & \lambda_2 & & \\ & \ddots & \ddots & \\ & & 1 & \lambda_n \end{bmatrix} \quad \Rightarrow \quad \begin{cases} \dot\psi_0 = \lambda_1\,\psi_0 \\ \dot\psi_k = \lambda_{k+1}\,\psi_k + \psi_{k-1}, \;\; k \geq 1 \end{cases} \qquad \text{(Putzer algorithm)}$$

Matrix exponential with respect to polynomial basis III

Example to the Putzer algorithm.

Suppose $A \in \mathbb{C}^{n\times n}$ has minimal polynomial

$$\mu_A(x) = (x-\lambda_1)(x-\lambda_2)^2, \quad \lambda_1 \neq \lambda_2.$$

The associated Newton Basis is

$$\beta_0(x) = 1, \quad \beta_1(x) = (x-\lambda_1), \quad \beta_2(x) = (x-\lambda_1)(x-\lambda_2).$$

Thus,

$$e^{At} = \psi_0(t)\, I + \psi_1(t)\, (A - \lambda_1\, I) + \psi_2(t)\, (A-\lambda_1\, I)(A-\lambda_2\, I),$$

where

$$\psi_0(0) = 1, \quad \dot\psi_0 = \lambda_1\, \psi_0, \quad \Rightarrow \quad \psi_0(t) = e^{\lambda_1 t},$$

$$\psi_1(0) = 0, \quad \dot\psi_1 = \lambda_2\, \psi_1 + \psi_0, \quad \Rightarrow \quad \psi_1(t) = \frac{e^{\lambda_1 t} - e^{\lambda_2 t}}{\lambda_1 - \lambda_2},$$

$$\psi_2(0) = 0, \quad \dot\psi_2 = \lambda_2\, \psi_2 + \psi_1, \quad \Rightarrow \quad \psi_2(t) = \frac{e^{\lambda_1 t} - e^{\lambda_2 t}}{(\lambda_1 - \lambda_2)^2} - \frac{e^{\lambda_2 t}}{\lambda_1 - \lambda_2}\, t.$$

Linear operators and polynomials

Polynomial identities, minimal polynomial, characteristic polynomial

Suppose the linear operator $A \in \mathcal{L}(\mathcal{V}, \mathcal{V})$ satisfies a **polynomial identity:**

$$p(A) = 0.$$

Then the Ideal $\mathcal{I}_A := \{ \, q \in \mathbb{K}[x] \, | \, q(A) = 0 \, \}$ contains a unique monic polynomial μ_A of minimum degree (the **minimal polynomial**). Thus $q(A) = 0$ iff $\mu_A | q$. Let

$$\mu_A = \prod_{k=1}^{r} f_k^{m_k}, \qquad f_k \text{ irreducible over } \mathbb{K}.$$

By the kernel decomposition theorem:

(1) $\mathcal{V} = \operatorname{Ker} \mu_A(A) = \bigoplus_{k=1}^{r} \operatorname{Ker} f_k^{m_k}(A)$,

(2) there are $\pi_k \in \mathbb{K}[x]$ such that $\pi_k(A) : \mathcal{V} \to \operatorname{Ker} f_k^{m_k}(A)$ are the associated projectors.

Furthermore, by the kernel intersection lemma and its proof (homework):

(3) $m \geq m_k \quad \Rightarrow \quad \operatorname{Ker} f_k^{m}(A) = \operatorname{Ker} f_k^{m_k}(A)$.

(4) $q(A)$ is invertible and there is a polynomial ϕ with $q(A) = \phi(A)^{-1}$ iff $f_k \nmid q$ for any k.

Cayley-Hamilton theorem: Let $\chi_A(x) := \det(x\, I - A)$ be the characteristic polynomial of $A \in \mathbb{K}^{n \times n}$. Then $\chi_A(A) = 0$. Thus, $\mu_A | \chi_A$.

Digression: Proof of the Cayley-Hamilton Theorem

Let $A = [a_{ik}] = [a_1, \ldots, a_n] \in \mathbb{K}^{n\times n}$. Let $e_j \in \mathbb{K}^n$ be the jth canonical basis vector. Then

$$\begin{aligned}
\sum_j \underbrace{\det(a_1, \ldots, a_{i-1}, e_j, a_{i+1}, \ldots, a_n)}_{=:a^\sharp_{ij}} a_{jk} &= \det(a_1, \ldots, a_{i-1}, \sum_j e_j\, a_{jk}, a_{i+1}, \ldots, a_n) \\
&= \det(a_1, \ldots, a_{i-1}, a_k, a_{i+1}, \ldots, a_n) \\
&= \det(A)\, \delta_{ik}.
\end{aligned}$$

In matrix notation: $A^\sharp A = \det(A)\, I$, where $A^\sharp = [a^\sharp_{ij}]$ (matrix of cofactors).

Replacing A by $x\, I - A$ with $x \in \mathbb{K}$, we have

$$(x\, I - A)^\sharp\, (x\, I - A) = \det(x\, I - A)\, I = \left(\sum p_j\, x^j\right) I. \qquad (*)$$

The entries of $(x\, I - A)^\sharp$ are polynomials of degree $n - 1$. Thus, $(x\, I - A)^\sharp = \sum_{j=1}^{n-1} B_j\, x^j$.

Equating coefficients of x^j in $(*)$ we obtain

$$-B_0\, A = p_0\, I, \qquad B_{j-1} - B_j\, A = p_j\, I \quad \text{for} \quad j = 1, \ldots, n-1, \qquad B_{n-1} = p_n\, I.$$

Multiplying the jth equation with A^j gives

$$-B_0\, A = p_0\, I, \qquad B_{j-1}\, A^j - B_j\, A^{j+1} = p_j\, A^j \quad \text{for} \quad j = 1, \ldots, n-1, \qquad B_{n-1}\, A^n = p_n\, A^n, .$$

Summing up these equations we obtain

$$0 = -B_0\, A + (B_0\, A - B_1\, A^2) + \ldots + (B_{n-2}\, A^{n-1} - B_{n-1}\, A^n) + B_{n-1}\, A^n = \sum p_j\, A^j.$$

2 Methods to compute the spectral projectors $\pi_k(A)$. Let

$$\mu_A = \prod_{k=1}^{r} f_k^{m_k}, \quad f_k \text{ irreducible over } \mathbb{K}, \qquad \ell_k = \frac{\mu_A}{f_k^{m_k}} = \prod_{j=1, j\neq k}^{r} f_j^{m_j}, \quad k = 1, 2, \ldots, r.$$

<u>Method 1 (already discussed).</u> Incomplete partial fraction expansion:

$$\frac{1}{\mu_A} = \sum_{k=1}^{r} \frac{\alpha_k}{\ell_k} \qquad \Rightarrow \qquad \pi_k(A) = \alpha_k(A)\,\ell_k(A).$$

<u>Method 2.</u> Let $\sigma(x) = \sum_{k=1}^{r} \ell_k(x)$. Then $\sigma(A)^{-1}$ exists and is a polynomial of A. We have

$$\pi_k(A) = \sigma(A)^{-1}\,\ell_k(A) \qquad\qquad (*)$$

<u>Proof.</u> The sum σ is not divisible by any f_j since $f_j | \ell_k$ for $j \neq k$ and $f_j \not| \ell_j$.
Hence μ_A and σ have no common factor, and there exist polynomials α, β with

$$\alpha\,\sigma + \beta\,\mu_A = 1 \qquad \Rightarrow \qquad \alpha(A)\,\sigma(A) = I \qquad \Rightarrow \quad \alpha(A) = \sigma(A)^{-1}.$$

The matrices $\pi_k(A)$ of $(*)$ satisfy $\sum \pi_k(A) = I$, $\pi_k(A)\pi_j(A) = \delta_{jk}\, I$, $f_k^{m_k}(A)\,\pi_k(A) = 0$.
This yields a second proof of the kernel decomposition theorem.

Linear operators and polynomials

The spectral decomposition theorem

Let $A \in \mathcal{L}(\mathcal{V}, \mathcal{V})$, where $\mathcal{V}$ is a vector space over $\mathbb{C}$. Suppose

$$p(A) = 0, \qquad \text{where} \quad p(X) = \prod_{k=1}^{r} (x - \lambda_k)^{m_k}, \quad \lambda_k \text{ distinct.}$$

Let $P_k := \pi_k(A) : \mathcal{V} \to \operatorname{Ker}(A - \lambda_k I)^{m_k}$, $k = 1, \ldots, r$ be the projectors for the decomp.

$$\mathcal{V} = \bigoplus_{k=1}^{r} \operatorname{Ker}(A - \lambda_k I)^{m_k}.$$

Furthermore, let $N_k := (A - \lambda_k I) P_k$. Then

$$\boxed{A = \textstyle\sum_{k=1}^{r} (\lambda_k P_k + N_k), \quad \text{and } N_k^{m_k} = 0.}$$

i.e. N_k is nilpotent. All these operators commute with each other:
$P_j P_k = P_k P_j = \delta_{jk} P_k$, $N_j P_k = P_k N_j = \delta_{jk} N_j$, $N_j N_k = N_k N_j$.

<u>Proof.</u> From $I = \sum_{k=1}^{r} \pi_k(A) = \sum_{k=1}^{r} P_k$ it follows that

$$A = \sum_{k=1}^{r} A P_k = \sum_{k=1}^{r} \lambda_k P_k + \underbrace{(A - \lambda_k I) P_k}_{N_k}.$$

All operators commute because they are polynomials of A. We have
$N_k^{m_k} = (A - \lambda_k I)^{m_k} P_k^{m_k} = (A - \lambda_k I)^{m_k} P_k = 0$, since P_k is a projector onto $\operatorname{Ker}(A - \lambda_k I)^{m_k}$.

The spectral decomposition and the matrix exponential

Let $A \in \mathbb{C}^{n\times n}$ have the spectral decomposition

$$A = \sum_{k=1}^{r} (\lambda_k P_k + N_k), \qquad N_k^{m_k - 1} \neq 0, \quad N_k^{m_k} = 0$$

Then

$$e^{At} = \sum_{k=1}^{r} e^{\lambda_k t} \left(P_k + N_k t + \frac{(N_k t)^2}{2} + \frac{(N_k t)^3}{3!} + \ldots + \frac{(N_k t)^{m_k - 1}}{(m_k - 1)!} \right) \qquad (*)$$

Proof. Let Y denote the right hand side of $(*)$. Using the relations

$$P_j P_k = P_k P_j = \delta_{jk} P_k, \quad N_j P_k = P_k N_j = \delta_{jk} N_j, \quad N_j N_k = N_k N_j$$

it is easily verified that $\dot{Y} = AY$. Furthermore, $Y(0) = I$.

Corollary. All components of e^{At} are Bohl functions.

Spectral decomposition and block diagonalization

Let $V = [V_1, V_2, \ldots, V_r]$ where V_k is a matrix whose columns form a basis of $\operatorname{Ker}(A - \lambda_k I)^{m_k}$. Then there are nilpotent matrices $\widehat{N}_k$ such that

$$A\,V_k = V_k\,(\lambda_k\,I + \widehat{N}_k).$$

Thus,

$$\begin{aligned} A &= V \begin{bmatrix} \lambda_1\,I + \widehat{N}_1 & & & \\ & \lambda_2\,I + \widehat{N}_2 & & \\ & & \ddots & \\ & & & \lambda_r\,I + \widehat{N}_r \end{bmatrix} V^{-1} \\ &= \sum_{k=1}^{r} V_k\,(\lambda_k\,I + \widehat{N}_k)\,W_k^\top = \sum_{k=1}^{r} \lambda_k\,\underbrace{V_k\,W_k^\top}_{P_k} + \underbrace{V_k\,\widehat{N}_k\,W_k^\top}_{N_k}, \qquad \text{where} \quad \begin{bmatrix} W_1^\top \\ W_2^\top \\ \vdots \\ W_r^\top \end{bmatrix} = V^{-1}. \end{aligned}$$

Furthermore,

$$e^{At} = V \begin{bmatrix} e^{(\lambda_1\,I+\widehat{N}_1)\,t} & & & \\ & e^{(\lambda_2\,I+\widehat{N}_2)\,t} & & \\ & & \ddots & \\ & & & e^{(\lambda_r\,I+\widehat{N}_r)\,t} \end{bmatrix} V^{-1}, \qquad e^{(\lambda_k\,I+\widehat{N}_k)\,t} = e^{\lambda_k\,t} \sum_{j=0}^{m_k-1} \frac{t^j}{j!}\widehat{N}_k^j.$$

The spectral decomposition and the resolvent $(x\,I - A)^{-1}$

Let $A \in \mathbb{C}^{n\times n}$ have the spectral decomposition

$$A = \sum_{k=1}^{r} (\lambda_k\, P_k + N_k), \qquad N_k^{m_k-1} \neq 0, \quad N_k^{m_k} = 0.$$

Then, as already mentioned,

$$e^{At} = \sum_{k=1}^{r} e^{\lambda_k\, t} \left(P_k + N_k\, t + \frac{(N_k\, t)^2}{2} + \frac{(N_k\, t)^3}{3!} + \ldots + \frac{(N_k\, t)^{m_k-1}}{(m_k-1)!} \right).$$

Using the properties of the spectral factors it can be shown that for any $x \in \mathbb{C}$,

$$(x\,I - A)^{-1} = \sum_{k=1}^{r} \left(\frac{P_k}{x-\lambda_k} + \frac{N_k}{(x-\lambda_k)^2} + \frac{N_k^2}{(x-\lambda_k)^3} + \ldots + \frac{N_k^{m_k-1}}{(x-\lambda_k)^{m_k}} \right),$$

(this is a partial fraction expansion).

Note, that $(x\,I - A)^{-1}$ is the Laplace transformation of e^{At},

$$\mathcal{L}\{e^{At}\}(x) = \int_{t=0}^{\infty} e^{-x\,t}\, e^{At}\, dt = (x\,I - A)^{-1}, \qquad \Re(x) > \max_k \Re(\lambda_k).$$

The spectral decomposition and matrix functions

From the spectral decomposition

$$A = \sum_{k=1}^{r} (\lambda_k P_k + N_k), \qquad N_k^{m_k - 1} \neq 0, \quad N_k^{m_k} = 0.$$

it follows that for any polynomial p,

$$p(A) = \sum_{k=1}^{r} \left(p(\lambda_k)\, P_k + N_k\, p'(\lambda_k) + \frac{N_k^2\, p''(\lambda_k)}{2} + \frac{N_k^3\, p^{(3)}(\lambda_k)}{3!} + \dots + \frac{N_k^{m_k-1}\, p^{(m_k-1)}(\lambda_k)}{(m_k-1)!} \right).$$

Taking limits, it follows that for any analytic function defined on neighborhood of the spectrum of A,

$$f(A) = \sum_{k=1}^{r} \left(f(\lambda_k)\, P_k + N_k\, f'(\lambda_k) + \frac{N_k^2\, f''(\lambda_k)}{2} + \frac{N_k^3\, f^{(3)}(\lambda_k)}{3!} + \dots + \frac{N_k^{m_k-1}\, f^{(m_k-1)}(\lambda_k)}{(m_k-1)!} \right).$$

The spectral decomposition theorem and stability

Definition: $A \in \mathbb{C}^{n\times n}$ is said to be (Hurwitz) stable if $\lim_{t\to\infty} e^{At} = 0$.

From

$$e^{At} = \sum_{k=1}^{r} e^{\lambda_k t} \left(P_k + N_k\, t + \frac{(N_k\, t)^2}{2} + \frac{(N_k\, t)^3}{3!} + \ldots + \frac{(N_k\, t)^{m_k-1}}{(m_k-1)!} \right)$$

it follows that A is **stable** iff all of its eigenvalues have negative real part.

Reason: for any polynomial p and $\lambda \in \mathbb{C}$,

$$\lim_{t\to\infty} |e^{\lambda\, t} p(t)| = \lim_{t\to\infty} e^{\Re(\lambda)\, t} |p(t)| = \begin{cases} \infty & \Re(\lambda) > 0 \text{ or } (\Re(\lambda) = 0 \text{ and degree}(p) > 0), \\ |p| & \Re(\lambda) = 0 \text{ and degree}(p) \leq 0, \\ 0 & \Re(\lambda) < 0. \end{cases} .$$

Corollary. If $\beta > \max\{\, \Re(\lambda_k) \mid k = 1, \ldots, r \,\}$ (spectral abscissa) then to any norm $\|\cdot\|$ there is a constant $M \geq 1$ such that

$$\|e^{At}\| \leq M\, e^{\beta\, t} \quad \text{for } t \geq 0. \qquad (\ (M, \beta)\text{-stability})$$

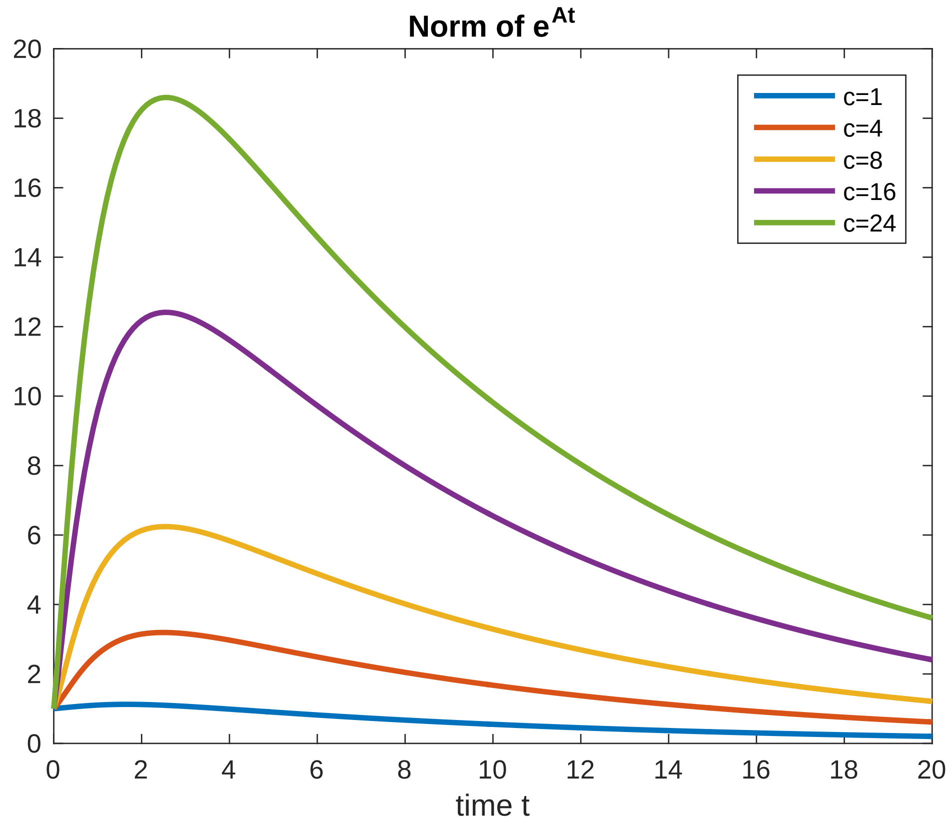
<u>Proof.</u> $A - \beta\, I$ is stable. Thus $\|e^{(A-\beta\, I)t}\| \leq M$ for some M and all $t \geq 0$.

Examples: transient behaviour of stable matrices

The picture below shows the function $t \mapsto \|e^{At}\|_2$ for the stable matrix

$$A = \begin{bmatrix} -0.6 & c \\ 0 & -1 \end{bmatrix}.$$

and several values of $c \in \mathbb{R}$.

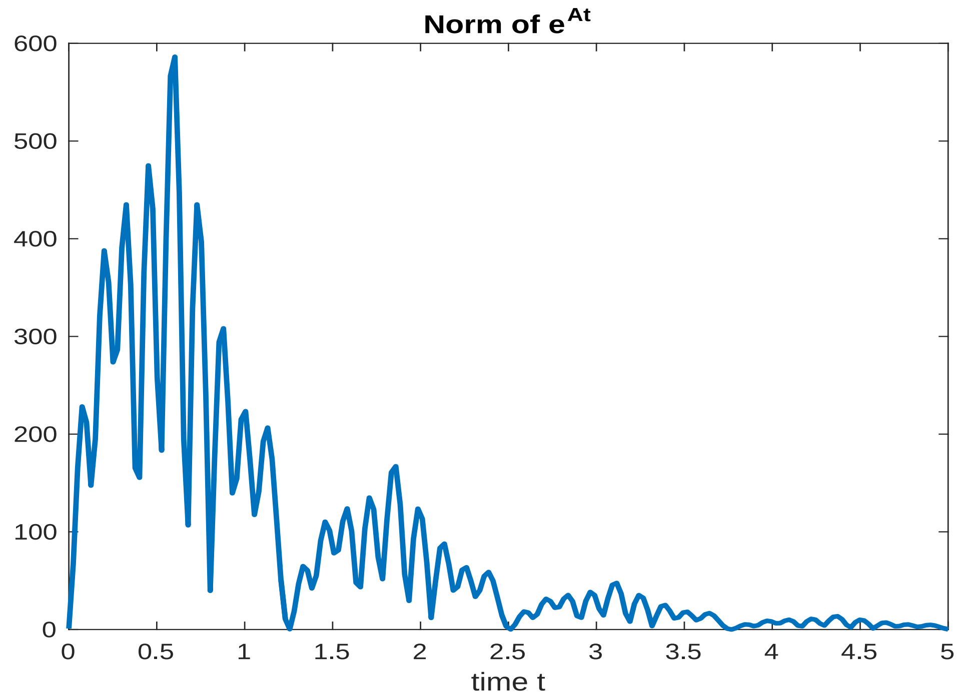


Examples: transient behaviour of stable matrices

The picture below shows the function $t \mapsto \|e^{At}\|_2$ for the stable matrix

$$A = \begin{bmatrix} -1 & -100 & 0 & -150 & 0 & 200 & -1000 \\ 1 & -1 & 1 & -10 & 25 & 11 & -200 \\ 0 & 0 & -1 & 400 & -30 & 0 & 250 \\ 0 & 0 & -1 & -1 & 5 & 5 & 200 \\ 0 & 0 & 0 & 0 & -1 & -2 & 30 \\ 0 & 0 & 0 & 0 & 0 & -1 & -625 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix}.$$

The eigenvalues of A are $-1, -1 \pm 10i, -1 \pm 20i, -1 \pm 25i$.



Growth bounds for $\|e^{At}\|$ (More growth bounds on the next slide collection).

Let $\|\cdot\|$ be a submultiplicative norm with $\|I\| = 1$. Especially,

$$\|A\|_1 = \max_{k=1}^{n} \sum_{i=1}^{n} |a_{ik}|, \qquad \|A\|_2 = \sqrt{\lambda_{\max}(A^*A)}, \qquad \|A\|_\infty = \max_{i=1}^{n} \sum_{k=1}^{n} |a_{ik}|.$$

Then we have for $t \geq 0$:

(1) $\|e^{A\,t}\| \leq e^{\|A\|\,t}$.

(2) $\widehat{\beta} := \min\{\, \beta > 0 \mid \|e^{A\,t}\| \leq e^{\beta\,t},\ t \geq 0 \,\} \qquad \Rightarrow \qquad \widehat{\beta} = \lim_{t \searrow 0} \dfrac{\|I + A\,t\| - 1}{t}$.

(3) $\widehat{\beta}_1 = \max_k \left(\Re(a_{kk}) + \sum_{i \neq k} |a_{ik}| \right), \quad \widehat{\beta}_2 = \lambda_{\max}\left(\frac{A+A^*}{2} \right), \quad \widehat{\beta}_\infty = \max_i \left(\Re(a_{ii}) + \sum_{k \neq i} |a_{ik}| \right)$

(4) Suppose $A = V\,\mathrm{diag}(\lambda_1, \ldots, \lambda_n)\,V^{-1}$. Let $\alpha = \max\{\Re(\lambda_1), \ldots, \Re(\lambda_n)\}$. Then

$$\|e^{A\,t}\| \leq \|V\|\,\|V^{-1}\|\,e^{\alpha\,t}$$

($\|V\|\,\|V^{-1}\|$ = condition number of V, α = spectral abscissa).

Proof. (1) and (4) are trivial. For (2) and (3) see

Hinrichsen, Pritchard: Mathematical Systems Theory 1.

Stability and bounded inputs

Let $A \in \mathbb{C}^{n\times n}$ be stable such that

$$\|e^{A\,t}\| \le M\,e^{\beta\,t} \qquad \beta < 0,$$

where $\|\cdot\|$ satisfies $\|Xv\| \le \|X\|\,\|v\|$ for matrices X and vectors v. Then the solution of

$$\dot{x} = A\,x, \qquad x(0) = 0$$

satisfies

$$\|x(t)\| = \left\| \int_0^t e^{A\,(t-s)}\, f(s)\, ds \right\| \le \begin{cases} \frac{M}{|\beta|} \sup_{s\in[0,t]} \|f(s)\|, \\ M\left(\frac{1}{q\,|\beta|}\right)^{1/q} \left(\int_0^t \|f(s)\|^p\, ds\right)^{1/p}, \quad p,q \ge 1,\ p^{-1} + q^{-1} = 1. \end{cases}$$

<u>Proof.</u> We have for $q \ge 1$,

$$\int_0^t \|e^{A(t-s)}\|^q\, ds = \int_0^t \|e^{A\,s}\|^q\, ds \le M^q \int_0^t e^{-q\,|\beta|\,s}\, ds = M^q\, \frac{1 - e^{-q\,|\beta|\,t}}{q\,|\beta|} \le \frac{M^q}{q\,|\beta|}.$$

Now, using Hölder inequality,

$$\left\| \int_0^t e^{A\,(t-s)}\, f(s)\, ds \right\| \le \int_0^t \|e^{A\,(t-s)}\|\,\|f(s)\|\, ds \le \left(\int_0^t \|e^{A\,(t-s)}\|^q\, ds \right)^{1/q} \left(\int_0^t \|f(s)\|^p\, ds \right)^{1/p}.$$

Stability and periodic inputs

Let $A \in \mathbb{C}^{n\times n}$ be stable, and let $f : [0,\infty) \to \mathbb{C}^n$ be T-periodic, i.e. $f(t+T) = f(t)$ for some $T > 0$.

Then any solution of $\quad \dot{x} = Ax + f \quad$ converges for $t \to \infty$ to the T-periodic function

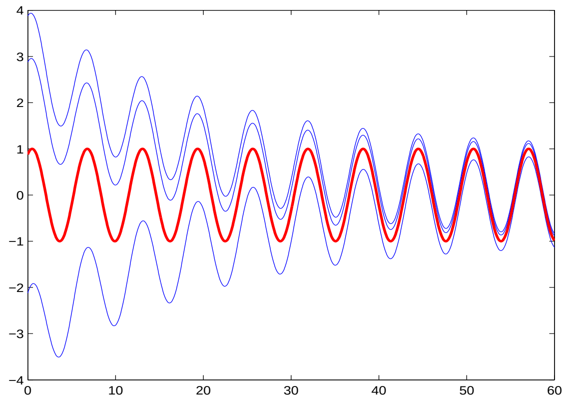
$$\widehat{x}(t) = \int_0^\infty e^{A\,s} f(t-s)\,ds$$

In particular, for any $\omega \in \mathbb{R}$ and any vector v the solutions of

$$\dot{x} = Ax + e^{i\,\omega\,t}\,v$$

converge to

$$\widehat{x}(t) = e^{i\omega\,t}\,v_\omega \quad \text{where} \quad v_\omega = (i\,\omega\,I - A)^{-1}v.$$



Proof. We have

$$x(t) = e^{A\,t}\,x(0) + \int_0^t e^{A(t-s)}\,f(s)\,ds = \underbrace{e^{A\,t}\,x(0)}_{\to 0} + \int_0^t e^{A\,t}\,f(t-s)\,ds \;\to \int_0^\infty e^{A\,t}\,f(t-s)\,ds,$$

and for $f(t) = e^{i\omega\,t}\,v$,

$$\int_0^\infty e^{As}e^{i\omega\,(t-s)}\,v = e^{i\,\omega\,t}\int_0^\infty e^{(A-i\omega\,I)\,s}\,v\,ds = e^{i\,\omega\,t}(A-i\omega\,I)^{-1}\int_0^\infty \underbrace{(A-i\omega\,I)e^{(A-i\omega\,I)\,s}}_{=\frac{d}{ds}e^{(A-i\omega\,I)\,s}}\,v\,ds = \widehat{x}(t).$$

Stability and small nonlinear feedback

Theorem. Let $A \in \mathbb{C}^{n\times n}$ be stable. Let

$$r = \frac{1}{\max_{\omega\in\mathbb{R}} \|(i\omega\, I - A)^{-1}\|_2} = \min_{\omega\in\mathbb{R}} \sigma_{\min}(i\omega\, I - A) \qquad \text{(stability radius)}$$

Let f be differentiable and such that

$$\|f(t,x)\|_2 < r\,\|x\|_2.$$

Then the solutions of

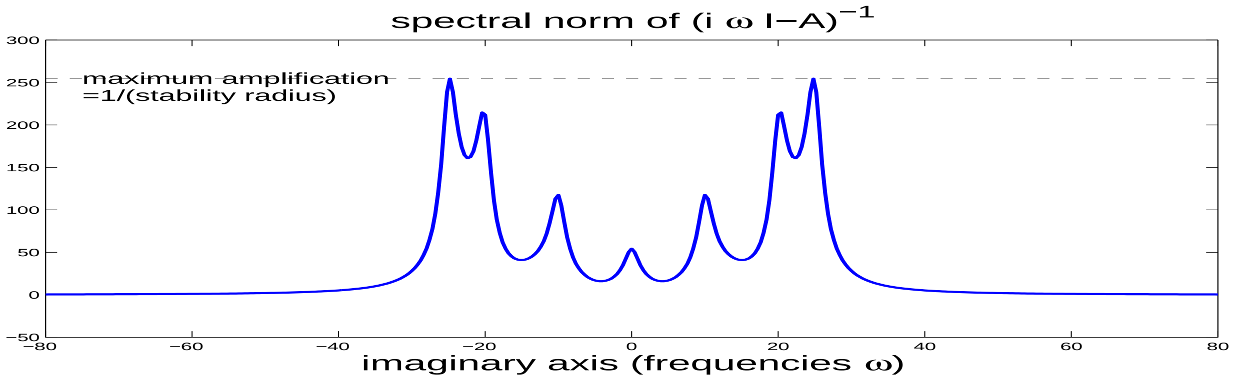
$$\dot{x}(t) = A\,x(t) + f(t, x(t))$$

satisfy

$$\lim_{t\to\infty} x(t) = 0.$$

Proof. This follows from the KYP-Lemma. A more general result will be given later.

Relationship between stability radius and frequency response



Recall: for stable A the solutions of

$$\dot{x} = Ax + e^{i\,\omega\,t}\,v \qquad .$$

converge to

$$\widehat{x}(t) = e^{i\omega\,t}\,(i\,\omega\,I - A)^{-1}v.$$

Signal amplification: $a = \frac{\|(i\,\omega\,I - A)^{-1}v\|_2}{\|v\|_2}$, in worst case $a = \|(i\,\omega\,I - A)^{-1}\|_2$.

Stability radius:

$$r = \frac{1}{\max_{\omega\in\mathbb{R}}\|(i\omega\,I - A)^{-1}\|_2} = \frac{1}{\text{maximum amplification factor}}.$$

Jordan canonical form

The Jordan chains

A finite sequence $v_1, v_2, \dots, v_\ell \in \mathcal{V} \setminus \{0\}$ is said to be a **Jordan chain** of $A \in \mathcal{L}(\mathcal{V}, \mathcal{V})$ to the eigenvalue $\lambda \in \mathbb{K}$ if

$$\begin{array}{lcl} Av_1 = \lambda\, v_1 & \Leftrightarrow & (A - \lambda\, I)v_1 = 0 \\ Av_2 = \lambda\, v_2 + v_1 & \Leftrightarrow & (A - \lambda\, I)v_2 = v_1 \\ \vdots & & \vdots \\ Av_\ell = \lambda\, v_\ell + v_{\ell-1} & \Leftrightarrow & (A - \lambda\, I)v_\ell = v_{\ell-1} \end{array}$$

For a Jordan chain we have

$$A\,[v_1\ v_2\ \dots v_\ell] = [v_1\ v_2\ \dots v_\ell]\ \underbrace{\begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & & 1 \\ & & & & \lambda \end{bmatrix}}_{\mathcal{J}_\ell(\lambda),\ \text{Jordan block of size } \ell}$$

and

$$v_k = (A - \lambda\, I)^{\ell-k} v_\ell, \qquad (A - \lambda\, I)^{\ell} v_\ell = (A - \lambda\, I)v_1 = 0,$$

Thus

$$v_k \in \mathsf{Ker}\,(A - \lambda\, I)^k \subseteq \mathsf{Ker}\,(A - \lambda\, I)^\ell.$$

It is easily verified that the vectors of a Jordan chain are linearly independent and span an A-invariant subspace.

Jordan canonical form

The Jordan canonical form theorem

Let $A \in \mathbb{C}^{n\times n}$ with minimal polynomial $\mu_A(x) = \prod_{k=1}^{r}(x-\lambda_k)^{m_k}$.
Then $\mathbb{C}^n$ has a basis $V = [v_1, v_2, \ldots v_n]$ consisting of Jordan chains, i.e.

$$V^{-1}AV = \text{diag}(\mathcal{J}_1, \ldots, \mathcal{J}_r), \qquad \mathcal{J}_k = \text{diag}(\mathcal{J}_{p_{k1}}(\lambda_k), \ldots, \mathcal{J}_{p_{ki}}(\lambda_k), \ldots)$$

The λ_k are the eigenvalues of A.
The sizes of the Jordan blocks are unique (up to ordering).
The maximum size of a Jordan block to the eigenvalue λ_k
equals the multiplicity m_k of λ_k in μ_A.

<u>Proof.</u> omitted.

Jordan canonical form

Jordan canonical form and matrix exponential

Fact 1: Let $v_1, \ldots, v_\ell$ be a Jordan chain for A to the eigenvalue λ such that

$$A v_1 = \lambda\, v_1, \qquad A v_k = \lambda\, v_k + v_{k-1}, \quad k = 2, \ldots, \ell.$$

Then the functions

$$\begin{aligned}
x_1(t) &= e^{\lambda_k t} v_1, \\
x_2(t) &= e^{\lambda_k t}(v_2 + t\, v_1), \\
x_3(t) &= e^{\lambda_k t}(v_3 + t\, v_2 + \frac{t^2}{2} v_1), \\
&\vdots \\
x_\ell(t) &= e^{\lambda_k t}(v_\ell + t\, v_{\ell-1} + \frac{t^2}{2} v_{\ell-2} + \ldots + \frac{t^{\ell-1}}{(\ell-1)!} v_1)
\end{aligned}$$

fulfill the homogeneous ODE

$$\dot{x}_j(t) = A\, x_j(t).$$

<u>Proof.</u> direct computation.

Fact 2: All solutions of $\dot{x} = A\, x$ are linear combinations of the above.

Jordan canonical form

Jordan canonical form and matrix exponential

Fact 3. Exponential of Jordan block:

$$J = \begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & & 1 \\ & & & & \lambda \end{bmatrix} \qquad \Rightarrow \qquad e^{Jt} = e^{\lambda t} \begin{bmatrix} 1 & t & \frac{t^2}{2} & \cdots & & \frac{t^{\ell-1}}{(\ell-1)!} \\ & 1 & t & & & \\ & & & & \ddots & \\ & & & \ddots & \ddots & \frac{t^2}{2} \\ & & & & 1 & t \\ & & & & & 1 \end{bmatrix}.$$

<u>Proof.</u> The matrix on the right fulfills $\dot{Y} = J\,Y$, $Y(0) = I$.

Fact 4. For all powers $k = 1, 2, \ldots,$

$$A = V\,\text{diag}(A_1, \ldots, A_r)\,V^{-1} \quad \Rightarrow \quad A^k = V\,\text{diag}(A_1^k, \ldots, A_r^k)\,V^{-1}.$$

By summation it follows that

$$A = V\,\text{diag}(A_1, \ldots, A_r)\,V^{-1} \quad \Rightarrow \quad e^{At} = V\,\text{diag}(e^{A_1 t}, \ldots, e^{A_r t})\,V^{-1}.$$

This in particular holds if the matrices A_j are Jordan blocks.

Jordan form and realization

Basic question of realization theory:

For a prescribed function of the form (Bohl function)

$$G(t) = \sum_{k=1}^{r} e^{\lambda_k t} \sum_{j=0}^{m_k-1} G_{kj} \frac{t^j}{j!}, \qquad G_{kj} \in \mathbb{C}^{p,q}$$

find matrices A, B, C such that the output y of the linear system

$$\dot{x} = Ax + Bu, \quad y = Cx, \quad x(0) = 0 \qquad (*)$$

is given by

$$y(t) = \int_0^t G(t-s)\, u(s)\, ds = \int_0^t Ce^{A(t-s)}B\, u(s)\, ds.$$

This can always be achieved by a matrix A consisting of Jordan blocks (though this might not be the optimal way).
An example is given on the following page.

Note that the **Laplace transform** of $(*)$ is

$$\widehat{y}(z) = \widehat{G}(z)\, \widehat{u}(z), \qquad \widehat{G}(z) = \sum_{k=1}^{r} \sum_{j=0}^{m_k-1} \frac{G_{kj}}{(z-\lambda_k)^{j+1}}$$

$\widehat{G}$ is called the **transfer function** of the system $(*)$.

Jordan form and realization

Realization example. Suppose

$$G(t) = e^{\lambda_1 t} G_1 + e^{\lambda_2 t}\left(G_{20} + G_{21}\, t + G_{22}\,\frac{t^2}{2}\right), \qquad G_{kj} \in \mathbb{C}^{p,q}.$$

Let

$$A = \begin{bmatrix} \lambda_1 I & & & \\ & \lambda_2 I & I & \\ & & \lambda_2 I & I \\ & & & \lambda_2 I \end{bmatrix}, \quad B = \begin{bmatrix} I \\ 0 \\ 0 \\ I \end{bmatrix}, \quad C = \begin{bmatrix} G_1 & G_{20} & G_{21} & G_{22} \end{bmatrix}.$$

Then

$$e^{A\,t}\,B = \begin{bmatrix} e^{\lambda_1 t}\, I & & & \\ & e^{\lambda_2 t}\, I & e^{\lambda_2 t}\, t\, I & e^{\lambda_2 t}\,\frac{t^2}{2}\, I \\ & & e^{\lambda_2 t}\, I & e^{\lambda_2 t}\; t\, I \\ & & & e^{\lambda_2 t}\, I \end{bmatrix} \begin{bmatrix} I \\ 0 \\ 0 \\ I \end{bmatrix} = \begin{bmatrix} e^{\lambda_1 t}\, I \\ e^{\lambda_2 t}\,\frac{t^2}{2}\, I \\ e^{\lambda_2 t}\; t\, I \\ e^{\lambda_2 t}\, I \end{bmatrix}$$

and

$$C\, e^{A\,t}\, B = \begin{bmatrix} G_1 & G_{20} & G_{21} & G_{22} \end{bmatrix} \begin{bmatrix} e^{\lambda_1 t}\, I \\ e^{\lambda_2 t}\,\frac{t^2}{2}\, I \\ e^{\lambda_2 t}\; t\, I \\ e^{\lambda_2 t}\, I \end{bmatrix} = G(t).$$

Perturbation of matrix exponential

For $A, E \in \mathbb{C}^{n\times n}$, $t \in \mathbb{R}$,

$$e^{(A+E)\,t} \;=\; e^{A\,t} + \int_0^t e^{A(t-s)} E e^{(A+E)\,s}\,ds \;=\; e^{A\,t} + \int_0^t e^{A(t-s)} E e^{A\,s}\,ds + \mathcal{O}(\|E\|^2).$$

<u>Proof.</u> Let $X(t) = e^{(A+E)\,t}$. Then

$$X(0) = I, \quad \dot{X}(t) = (A+E)\,X(t) = A\,X(t) + F(t), \quad F(t) = E\,e^{(A+E)\,t}.$$

Thus,

$$\begin{aligned} e^{(A+E)\,t} \;=\; X(t) \;&=\; e^{A\,t}I + \int_0^t e^{A\,(t-s)} F(s)\,ds \\ &=\; e^{A\,t} + \int_0^t e^{A\,(t-s)} E\, \underbrace{e^{(A+E)\,s}}_{G(s)}\,ds. \end{aligned}$$

By the same formula with t replaced with s,

$$G(s) = e^{(A+E)\,s} = e^{A\,s} + \int_0^s e^{A\,(s-\tau)} E\,e^{(A+E)\,\tau}\,d\tau.$$

Thus,

$$e^{(A+E)\,t} \;=\; e^{A\,t} + \int_0^t e^{A(t-s)} E e^{A\,s}\,ds + \underbrace{\int_0^t e^{A\,(t-s)} E \int_0^s e^{A\,(s-\tau)} E\,e^{(A+E)\,\tau}\,d\tau\,ds.}_{\mathcal{O}(\|E\|^2)}$$

Numerical computation of the matrix exponential

Method, implemented as MATLAB function `expm(A)`:

Pade approximation combined with scaling and squaring.

Pade approximation (rational approximation) of scalar exponential:

$$e^x \approx r_m(x) = \frac{p_m(x)}{p_m(-x)}, \quad \text{where } p_m(x) = \sum_{j=0}^{m} \frac{(2m-j)!\, m!}{(m-j)!\,(2m)!\, j!}\, x^j \qquad (\text{good for } |x| \text{ small})$$

For any integer m, $e^{At} = (e^{At/m})^m$. Thus (scaling and squaring):

$$e^A = (e^{A/2^s})^{2^s} = B_s, \qquad \text{where} \qquad B_1 = e^{A/2^s},\ B_{k+1} = B_k^2,\ \ k = 1, \ldots, s-1.$$

MATLAB:

$$e^A \approx r_6(A/2^s)^{2^s}, \quad \|A/2^s\|_\infty \le 1/2.$$

For more sophisticated methods see book and talks (WWW) by N. Higham.

Final comments:

Literature on numerical computation of matrix exponential:

1) Cleve Moler, Charles Van Loan.
   Paper: Dubious Ways to Compute the Exponential of a Matrix,
   Twenty-Five Year Later. SIAM review, 45, No 1, (2003)

2) N. J. Higham.
   Book: Functions of Matrices: Theory and Computation. SIAM (2008)

3) N. J. Higham.
   Slides: How and How Not to Compute the Exponential of a Matrix
   `http://www.maths.manchester.ac.uk/~higham/talks/exp10.pdf`

Matrix exponential in MATLAB: `expm(A)`

Examples of $\|e^{At}\|$ are from Hinrichsen, Pritchard: Mathematical Systems Theory 1

'Ansatz' (german) in math. context: preparation, basic approach

Piers Bohl (1865 1921) was a Latvian mathematician