# Solutions of Linear Systems

- Analysis of autonomous linear systems
- Stability
- A bit of Lyapunov theory
- Variation-of-Constants Formula
- Transfer Matrices

## Related Reading

[AM]: Chapters 4.3-4.4, 5.1-5.3, 8.1-8.3 and [F]: Chapters 3.1-3.6

## Autonomous Systems

A dynamical state-space system without inputs/outputs is described as

$$\dot{x} = f(x) \quad \text{with} \quad f : \mathbb{R}^n \to \mathbb{R}^n.$$

Understanding the response of such systems to initial conditions $x_0 \in \mathbb{R}^n$ is a rich field called dynamical systems theory. For us the main tool is numerical simulation.

The system is linear if $f$ is a linear map. All such maps are described as $f(x) = Ax$ with some matrix $A \in \mathbb{R}^{n\times n}$. Hence a linear autonomous system is described as

$$\dot{x} = Ax \quad \text{with} \quad A \in \mathbb{R}^{n\times n}.$$

The responses of such systems **can be completely understood** by using tools from linear algebra.

# Diagonal $A$

Suppose that $A$ is a diagonal matrix:

$$A = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix} \quad \text{with} \quad \lambda_1, \ldots, \lambda_n \in \mathbb{R}.$$

Then $\dot{x} = Ax$ splits up into the $n$ scalar differential equations

$$\dot{x}_1 = \lambda_1 x_1, \ \ldots, \ \dot{x}_n = \lambda_n x_n.$$

These can be solved independently. All solutions are

$$x_1(t) = e^{\lambda_1 t} x_{01}, \ \ldots, \ x_n(t) = e^{\lambda_n t} x_{0n} \quad \text{with} \quad x_{01}, \ldots, x_{0n} \in \mathbb{R}.$$

This can be compactly written as

$$x(t) = \begin{pmatrix} e^{\lambda_1 t} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{\lambda_n t} \end{pmatrix} x_0 \quad \text{with} \quad x_0 = \begin{pmatrix} x_{01} \\ \vdots \\ x_{0n} \end{pmatrix}.$$

Observe that $x(0) = x_0$.

## State-Coordinate Change

In practice $A$ will not be diagonal. Can we still exploit what we have seen? For this purpose we introduce the following key concept.

An arbitrary non-singular (or invertible) matrix $T \in \mathbb{R}^{n\times n}$ defines a **state-coordinate transformation** $z = Tx$.

- Suppose that $x(t)$ satisfies $\dot{x}(t) = Ax(t)$. Then $z(t) = Tx(t)$ satisfies

$$\dot{z}(t) = T\dot{x}(t) = TAx(t) = (TAT^{-1})z(t) = \tilde{A}z(t).$$

In the new coordinates the system is described with $\tilde{A} = TAT^{-1}$.

- Conversely suppose that $z(t)$ is any trajectory of $\dot{z}(t) = \tilde{A}z(t)$. Then

$$x(t) = T^{-1}z(t) \;\text{ satisfies }\; \dot{x}(t) = Ax(t).$$

If we understand the solutions of $\dot{z} = \tilde{A}z$ then we understand those of $\dot{x} = Ax$, since they are linearly transformed into each other by $T$.

# Diagonalizable $A$

In many practical cases $A$ can be transformed into a diagonal matrix $\tilde{A}$ by a suitably chosen state-coordinate change.

Suppose that $TAT^{-1} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Then the unique solution of $\dot{x} = Ax,\ x(0) = x_0$ is given by

$$x(t) = [T^{-1}\,\mathrm{diag}(e^{\lambda_1 t}, \ldots, e^{\lambda_n t})T]x_0.$$

**Proof.** Just combine the two previous slides. $x(t)$ satisfies

$$\dot{x}(t) = Ax(t) \ \text{ with } \ x(0) = x_0$$

if and only if $z(t) = Tx(t)$ and $z_0 = Tx_0$ satisfy

$$\dot{z}(t) = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)z(t) \ \text{ with } \ z(0) = z_0.$$

The latter problem has the unique solution

$$z(t) = \mathrm{diag}(e^{\lambda_1 t}, \ldots, e^{\lambda_n t})z_0.$$

## How to Diagonalize $A$?

Matlab does it for us with `[S,La]=eig(A)`.

Linearize Segway in upright position: $\dot{x}_\Delta = Ax_\Delta$ with

$$A = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 3.92 & -2 & -0.32 \\ 0 & 22.1 & -3.23 & -1.82 \end{pmatrix}$$

Get $\Lambda = S^{-1}AS$ with

$$S = \begin{pmatrix} 1 & -0.03 & -0.04 & 0.58 \\ 0 & -0.26 & -0.16 & -0.11 \\ 0 & -0.12 & 0.22 & -0.79 \\ 0 & -0.96 & 0.96 & 0.16 \end{pmatrix}, \quad \Lambda = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 3.72 & 0 & 0 \\ 0 & 0 & -6.16 & 0 \\ 0 & 0 & 0 & -1.38 \end{pmatrix}$$

**Qualitative insight.** Each component of any solution $x_\Delta(t)$ of $\dot{x}_\Delta = Ax_\Delta$ is a linear combination of the exponential functions

$$e^{0\,t} = 1, \quad e^{3.72\,t}, \quad e^{-6.16\,t}, \quad e^{-1.38\,t}.$$

# Example

We can now **answer all kinds of questions** about the whole set of solutions.

If $x_0$ equals the first, second, third, fourth column of $S$ the solutions are

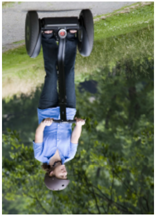
$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} e^{0\,t}, \; -\begin{pmatrix} 0.03 \\ 0.26 \\ 0.12 \\ 0.96 \end{pmatrix} e^{3.72\,t}, \; \begin{pmatrix} -0.04 \\ -0.16 \\ 0.22 \\ 0.96 \end{pmatrix} e^{-6.16\,t}, \; \begin{pmatrix} 0.58 \\ -0.11 \\ -0.79 \\ 0.16 \end{pmatrix} e^{-1.38\,t}.$$

**All other solutions are linear combinations thereof.** Please note that we have complete insight into how solutions behave for $t \to \infty$.

**Question:** For which initial conditions does the solution not blow up?

**Answer:** For all $x_0$ that can be written as $x_0 = S \begin{pmatrix} z_{01} \\ 0 \\ z_{03} \\ z_{04} \end{pmatrix}$.

# Example



For the linearization in the downright position suitable matrices $S$, $T = S^{-1}$ and $\Lambda$ are given by

$$T = \begin{pmatrix} 1 & 0 & 0.7 & 0.12 \\ 0 & -0.03 + 2.58i & 0.12 - 0.37i & 0.54 + 0.08i \\ 0 & -0.03 - 2.58i & 0.12 + 0.37i & 0.54 - 0.08i \\ 0 & -0.36 & 1.33 & 0.21 \end{pmatrix} = \begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \end{pmatrix}$$

$$S = \begin{pmatrix} 1 & 0.04i & -0.04i & -0.54 \\ 0 & -0.05 - 0.2i & -0.05 + 0.2i & 0.12 \\ 0 & -0.17 - 0.05i & -0.17 + 0.05i & 0.81 \\ 0 & 0.96 & 0.96 & -0.18 \end{pmatrix} = \begin{pmatrix} c_1 & c_2 & c_3 & c_4 \end{pmatrix}$$

$$\Lambda = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -1.17 + 4.46i & 0 & 0 \\ 0 & 0 & -1.17 - 4.46i & 0 \\ 0 & 0 & 0 & -1.49 \end{pmatrix} = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3, \lambda_4).$$

## Complex Transformations and Diagonal Matrices

Despite the fact that $T$ and $\Lambda$ are complex, the main result on slide 5 stays true **without change!**

In fact observe that $\lambda_2 = \bar{\lambda}_3$, $c_2 = \bar{c}_3$, $r_2 = \bar{r}_3$. Therefore

$$T^{-1}\,\mathrm{diag}(e^{\lambda_1 t}, e^{\lambda_2 t}, e^{\lambda_3 t}, e^{\lambda_4 t})T =$$

$$= \begin{pmatrix} c_1 e^{\lambda_1 t} & c_2 e^{\lambda_2 t} & \bar{c}_2 e^{\bar{\lambda}_2 t} & c_4 e^{\lambda_4 t} \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \\ \bar{r}_2 \\ r_4 \end{pmatrix} =$$

$$= e^{\lambda_1 t} c_1 r_1 + e^{\lambda_2 t} c_2 r_2 + e^{\bar{\lambda}_2 t} \bar{c}_2 \bar{r}_2 + e^{\lambda_4 t} c_4 r_4 =$$

$$= e^{\lambda_1 t} c_1 r_1 + 2\mathrm{Re}[e^{\lambda_2 t} c_2 r_2] + e^{\lambda_4 t} c_4 r_4$$

is always a **real** matrix. The response to initial condition $x_0 \in \mathbb{R}^4$ is

$$e^{\lambda_1 t} c_1 (r_1 x_0) + 2\mathrm{Re}[e^{\lambda_2 t} c_2 (r_2 x_0)] + e^{\lambda_4 t} c_4 (r_4 x_0).$$

# Complex Transformations and Diagonal Matrices

For a complex number $\lambda = \sigma + i\omega$ ($\sigma = \mathsf{Re}(\lambda)$, $\omega = \mathsf{Im}(\lambda)$) recall

$$e^{\lambda t} = e^{(\sigma + i\omega)t} = e^{\sigma t}[\cos(\omega t) + i\sin(\omega t)].$$

If $c$ is a complex column and $r$ a complex row then

$$cr = [\mathsf{Re}(c) + i\mathsf{Im}(c)][\mathsf{Re}(r) + i\mathsf{Im}(r)] = M + iN$$

with the real matrices

$$M = [\mathsf{Re}(c)\mathsf{Re}(r) - \mathsf{Im}(c)\mathsf{Im}(r)], \quad N = [\mathsf{Re}(c)\mathsf{Im}(r) + \mathsf{Im}(c)\mathsf{Re}(r)].$$

This leads to the explicit formula

$$\boxed{\mathsf{Re}[e^{\lambda t}cr] = e^{\sigma t}\left[\cos(\omega t)M - \sin(\omega t)N\right].}$$

The components of $\mathsf{Re}[e^{\lambda t}cr]x_0$ hence behave **qualitatively** as follows:

$\sigma = 0$, $\sigma > 0$, $\sigma < 0$: sustained, increasing, decreasing oscillation.

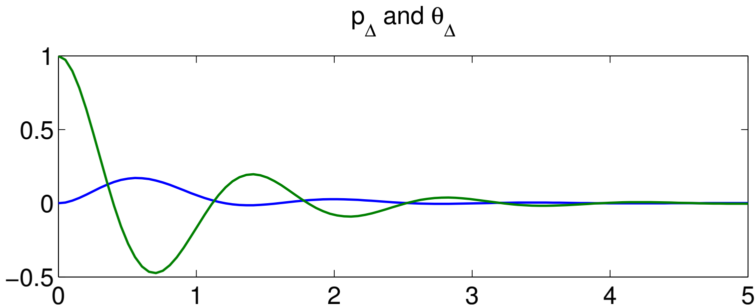
# Example: Linearization of Segway

Solution with $x_0 = \text{col}(1,0,0,0)$ is $x_\Delta(t) = \text{col}(1,0,0,0)$ for all $t \geq 0$.

Solution with $x_0 = \text{col}(0,1,0,0)$ is

$$2\,e^{-1.17\,t}\left[\cos(4.46\,t)\begin{pmatrix}-0.1\\0.52\\0.15\\-0.03\end{pmatrix} - \sin(4.46\,t)\begin{pmatrix}-0.01\\-0.13\\-0.43\\2.48\end{pmatrix}\right] + e^{-1.49\,t}\begin{pmatrix}0.2\\-0.04\\-0.29\\0.06\end{pmatrix}$$

and hence decaying in an oscillatory fashion.



# Recap

The complex number $\lambda \in \mathbb{C}$ is an **eigenvalue** of $A \in \mathbb{R}^{n \times n}$ if it is a zero of the **characteristic polynomial of** $A$:

$$\det(\lambda I - A) = 0.$$

If $\lambda$ is an eigenvalue of $A$ then any non-zero vector $e \in \mathbb{C}^n$ with

$$(\lambda I - A)e = 0 \text{ or equivalently } Ae = \lambda e$$

is a corresponding **eigenvector** of $A$. (Eigenvectors are not unique.)

## Some Facts

- If $(\lambda, e)$ is a pair of eigenvalue/eigenvector then so is $(\bar{\lambda}, \bar{e})$.
- Any matrix $A$ does have at most $n$ different eigenvalues.
- For each eigenvalue one can compute at most $n$ linearly independent eigenvectors (by computing a basis of the null space of $\lambda I - A$).
- Eigenvalues are not modified by a coordinate change.

## How to Diagonalize $A$ in Theory?

For each eigenvalue of $A \in \mathbb{R}^{n\times n}$ we can hence compute a set of linearly independent eigenvectors; if done for $\lambda$, take the complex conjugate vectors for $\bar{\lambda}$; collect these vectors in the list $e_1, \ldots, e_g$ and denote by $\lambda_1, \ldots, \lambda_g$ the corresponding (not necessarily different) eigenvalues.

We have $g \leq n$ and the vectors $e_1, \ldots, e_g$ are linearly independent. In case that $g = n$ then

$$S^{-1}AS = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) \quad \text{with} \quad S = \begin{pmatrix} e_1 & \cdots & e_n \end{pmatrix}.$$

If $g < n$ then $A$ **cannot** be diagonalized by a coordinate change.

The formula for $S^{-1}AS$ is very easy to see:

$$AS = A \begin{pmatrix} e_1 & \cdots & e_n \end{pmatrix} =$$

$$= \begin{pmatrix} Ae_1 & \cdots & Ae_n \end{pmatrix} = \begin{pmatrix} \lambda_1 e_1 & \cdots & \lambda_n e_n \end{pmatrix} =$$

$$= \begin{pmatrix} e_1 & \cdots & e_n \end{pmatrix} \mathrm{diag}(\lambda_1, \ldots, \lambda_n) = S\,\mathrm{diag}(\lambda_1, \ldots, \lambda_n).$$

## Example 1

Consider the matrix

$$A = \begin{pmatrix} 3 & 4 & 6 & -4 & 8 \\ 2 & 3 & 4 & -2 & 6 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 2 & 2 & -1 & 2 \\ -2 & -2 & -4 & 2 & -5 \end{pmatrix}$$

The characteristic polynomial of $A$ is computed with `poly(A)` to

$$\det(\lambda I - A) = (\lambda + 1)^2 (\lambda - 1)^3.$$

The LU-factorization (with command `lu`) of $1I - A$ and $-1I - A$ give

$$U_1 = \begin{pmatrix} -2 & -4 & -6 & 4 & -8 \\ 0 & 2 & 2 & -2 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad U_2 = \begin{pmatrix} -4 & -4 & -6 & 4 & -8 \\ 0 & -2 & -1 & 0 & -2 \\ 0 & 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Since the dimensions of the null-spaces of $U_1$ and $U_2$ are $3$ and $2$, we have $g = 5 = n$ and hence $A$ **can be** diagonalized.

# Example 1

Linear independent vectors in the null-space of $U_1$ and $U_2$ are

$$\begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ -1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Then

$$S = \begin{pmatrix} -1 & 0 & -2 & 1 & -1 \\ -1 & 1 & -1 & 0 & -1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix} \text{ satisfies } S^{-1}AS = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix}.$$

# Example 2

Consider the matrix

$$A = \begin{pmatrix} 1 & 7 & 7 & -8 & 6 \\ 1 & 5 & 5 & -5 & 5 \\ 1 & 0 & 2 & -1 & 1 \\ 0 & 3 & 3 & -3 & 2 \\ -1 & -4 & -5 & 5 & -4 \end{pmatrix}.$$

The characteristic polynomial of $A$ is computed with `poly(A)` to

$$\det(\lambda I - A) = (\lambda + 1)^2 (\lambda - 1)^3.$$

The LU-factorization (with command `lu`) of $1I - A$ and $-1I - A$ give

$$U_1 = \begin{pmatrix} -1 & -4 & -5 & 5 & -5 \\ 0 & -7 & -7 & 8 & -6 \\ 0 & 0 & 0 & 0.57 & 0.57 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \; U_2 = \begin{pmatrix} -2 & -7 & -7 & 8 & -6 \\ 0 & 3.5 & 0.5 & -3 & 2 \\ 0 & 0 & -2.57 & -0.57 & -0.29 \\ 0 & 0 & 0 & -0.89 & -0.44 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Since the dimensions of the null-spaces of $U_1$ and $U_2$ are $2$ and $1$, we have $g = 3 < 5 = n$ and hence $A$ is **not** diagonalizable.

# Summary

Let us be given $\dot{x} = Ax$ with $A \in \mathbb{R}^{n \times n}$:

- We can check whether it can be diagonalized as shown on slide 13. If existing one can **compute** a diagonalizing transformation $T$. Then $TAT^{-1}$ has on its diagonal all eigenvalues of $A$ (possibly repeated).
- If $A$ can be diagonalized, we can get complete insight into the solution set of the differential equation as described on slide 5.

With the transformation as on slide 13 the formula on slide 5 reads as

$$x(t) = \begin{pmatrix} e^{\lambda_1 t} e_1 & \cdots & e^{\lambda_n t} e_n \end{pmatrix} S^{-1} x_0.$$

The eigenvalues determine the dynamic characteristics of the solution. The eigenvectors provide information how these influence the state.

The eigenvalues of $A$ are often called the **modes** of the system $\dot{x} = Ax$. The corresponding eigenvectors are the **mode-shapes**.

# Matrix Exponential

If we recall the Taylor series expansion of the function $e^x$ we infer

$$T^{-1}\begin{pmatrix} e^{\lambda_1 t} \cdots & 0 \\ \vdots \quad \ddots & \vdots \\ 0 \quad \cdots & e^{\lambda_n t} \end{pmatrix} T = T^{-1}\begin{pmatrix} \sum_{k=0}^{\infty} \frac{1}{k!}(\lambda_1 t)^k \cdots & & 0 \\ \vdots & \ddots & \vdots \\ 0 & & \cdots \sum_{k=0}^{\infty} \frac{1}{k!}(\lambda_n t)^k \end{pmatrix} T =$$

$$= \sum_{k=0}^{\infty} \frac{1}{k!}\left[T^{-1}\begin{pmatrix} \lambda_1^k \cdots & 0 \\ \vdots \quad \ddots & \vdots \\ 0 \quad \cdots & \lambda_n^k \end{pmatrix} T\right] t^k = \sum_{k=0}^{\infty} \frac{1}{k!}\left[T^{-1}\begin{pmatrix} \lambda_1 \cdots & 0 \\ \vdots \quad \ddots & \vdots \\ 0 \quad \cdots & \lambda_n \end{pmatrix} T\right]^k t^k.$$

In case that $A = T^{-1}\operatorname{diag}(\lambda_1, \ldots, \lambda_n)T$ we hence conclude that

$$T^{-1}\begin{pmatrix} e^{\lambda_1 t} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{\lambda_n t} \end{pmatrix} T = \sum_{k=0}^{\infty} \frac{1}{k!}(At)^k.$$

## Matrix Exponential

For any $M \in \mathbb{R}^{n\times n}$ recall $M^0 = I$ and $M^k = MM^{k-1}$ for $k = 1, 2, \ldots$.

For any $A \in \mathbb{R}^{n\times n}$ and $t \in \mathbb{R}$ define

$$e^{At} := \lim_{N\to\infty} \sum_{k=0}^{N} \frac{1}{k!}(At)^k = \sum_{k=0}^{\infty} \frac{1}{k!}(At)^k.$$

- The series converges. Hence $e^{At}$ is a well-defined function of $t$.
- $e^{A0} = I$ and $e^{A(t+\tau)} = e^{At}e^{A\tau}$ and hence $e^{-At} = [e^{At}]^{-1}$.
- $e^{At}$ is arbitrarily often differentiable and $\frac{d}{dt}e^{At} = Ae^{At} = e^{At}A$.

The third statement follows from

$$\frac{d}{dt}\sum_{k=0}^{\infty}\frac{1}{k!}(At)^k = \sum_{k=1}^{\infty} A^k \frac{1}{k!}\left[\frac{d}{dt}t^k\right] = A\sum_{k=1}^{\infty} A^{k-1}\frac{1}{(k-1)!}t^{k-1} = Ae^{At}$$

and the observation that we can also pull $A$ out to the right.

## General $A$

For any $A \in \mathbb{R}^{n\times n}$ the unique solution of $\dot{x} = Ax,\ x(0) = x_0$ is

$$x(t) = e^{At}x_0.$$

**Proof.** To show existence we just check that $x(t)$ is a solution:

$$x(0) = e^{A0}x_0 = x_0 \ \text{ and } \ \dot{x}(t) = \frac{d}{dt}e^{At}x_0 = Ae^{At}x_0 = Ax(t).$$

To show uniqueness let $y(t)$ be another solution. Then

$$\frac{d}{dt}[e^{-At}y(t)] = [\frac{d}{dt}e^{-At}]y(t) + e^{-At}[\frac{d}{dt}y(t)] =$$

$$= -Ae^{-At}y(t) + e^{-At}Ay(t) = e^{-At}[-A + A]y(t) = 0.$$

Since $e^{-At}y(t)$ has zero derivative for all $t \in \mathbb{R}$, it must be constant. If we evaluate at $t = 0$, we observe that this constant must equal $x_0$. We conclude $e^{-At}y(t) = x_0$ or $y(t) = e^{At}x_0 = x(t)$ for all $t \in \mathbb{R}$.

## Computation of $e^{At}$

If $A$ is diagonalizable we have seen how to compute $e^{At}$. This procedure cannot be applied to all matrices.

**Example.** The double integrator $\ddot{q} = u$ has the state-space description

$$\dot{x} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} x + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u = Ax + Bu.$$

$A$ is not diagonalizable. This is shown as follows:

- Since $\det(A - \lambda I) = \lambda^2$ the only eigenvalue is $\lambda_1 = 0$.
- Clearly $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ is a corresponding eigenvector since $e_1 \neq 0$ and $(A - \lambda_1 I)e_1 = 0$. Because $A - \lambda_1 I$ has a null-space of dimension one, there is no eigenvector that is linearly independent from $e_1$.
- Therefore the list from slide 13 is $e_1$ with $g = 1 < 2 = n$. Hence the theorem on that slide implies that $A$ is not diagonalizable.

## Computation of $e^{At}$

Now note that $(At)^2 = 0$. Hence we have

$$e^{At} = I + (At) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & t \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}.$$

All solutions of $\dot{x} = Ax$ are hence given by

$$x(t) = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} x_0 = \begin{pmatrix} x_{01} + tx_{02} \\ x_{02} \end{pmatrix}.$$

We observe that, next to $e^{0t}$, the components of the solution also involve the term $te^{0t}$. This illustrates the following general structural insight.

All elements of $e^{At}$ are linear combinations of the terms

$$e^{\lambda_1 t}, te^{\lambda_1 t}, \ldots, t^{n-1}e^{\lambda_1 t},\ \ldots,\ e^{\lambda_p t}, te^{\lambda_p t}, \ldots, t^{n-1}e^{\lambda_p t}$$

if $\lambda_1, \ldots, \lambda_p$ are the different eigenvalues of $A$.

# Jordan Form

The key tool is the always existing **Jordan canonical form** of a matrix.

For any $A \in \mathbb{R}^{n \times n}$ there exists a non-singular matrix $S$ such that

$$S^{-1}AS = J = \begin{pmatrix} J_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & J_g \end{pmatrix} \text{ with } J_l = \begin{pmatrix} \lambda_l & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & \lambda_l & 1 \\ 0 & \cdots & 0 & \lambda_l \end{pmatrix}.$$

$J_l$ are so-called **Jordan blocks**. Moreover:

- Up to permutation of the Jordan blocks the Jordan canonical form $\mathrm{diag}(J_1, \ldots, J_g)$ is **uniquely determined by** $A$.
- $\lambda_1, \ldots, \lambda_g$ are the (not necessarily different) eigenvalues of $A$.
- There are exactly $g$ linearly independent eigenvectors of $A$.
- $A$ is diagonalizable iff all Jordan blocks have dimension one.

## Example

Recall the example on slide 16. The command $[\mathtt{S}, \mathtt{J}] = \mathtt{jordan}(\mathtt{A})$ does indeed return $S^{-1}AS = J$ with

$$S = \begin{pmatrix} -3 & -1 & 1 & 5 & 3 \\ -2 & -1 & 1 & 4 & 3 \\ 0 & 0 & 1 & 1 & 1 \\ -1 & 0 & 1 & 2 & 2 \\ 2 & 1 & -1 & -3 & -2 \end{pmatrix} \quad \text{and} \quad J = \left(\begin{array}{cc|cc|c} -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \end{array}\right).$$

- $A$ has three Jordan blocks of dimension $2$, $2$ and $1$.
- The first, third, fifth column of $S$ are three linearly independent eigenvectors of $A$ for the eigenvalues $-1$, $1$, $1$ (see next slide).
- $A$ is not diagonalizable since it has Jordan blocks of dimension two.

**Warning.** Computing the Jordan canonical is numerically unreliable. Be very careful to thoughtlessly use the command `jordan` in Matlab!

## Example

What is the meaning of the columns of $S$? From $S^{-1}AS = J$ we clearly infer $S^{-1}(\lambda I - A)S = \lambda I - J$. By taking powers $S^{-1}(\lambda I - A)^2 S = (\lambda I - J)^2$ etc. This leads to the following insightful relation:

$$\boxed{(\lambda I - A)^k S = S(\lambda I - J)^k \quad \text{for} \quad k = 1, 2, 3, \ldots}$$

For the example matrix and for $\lambda = 1$ we obtain the relations

$$(\lambda I - A)S = S \begin{pmatrix} 2 & -1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad (\lambda I - A)^2 S = S \begin{pmatrix} 4 & -4 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

This shows that the third and fifth columns of $S$ are eigenvectors of $A$ with respect to $\lambda = 1$. The fourth column of $S$ is in the null-space of $(\lambda I - A)^2$ and is called a **generalized eigenvector** of $A$. In general $S$ is constructed from eigenvectors and generalized eigenvectors of $A$.

# About Jordan forms and generalized eigenvectors

True / ordinary eigenvectors:

$$(\lambda I - A)e = 0$$

Generalized eigenvectors:

$$(\lambda I - A)^k e = 0$$

number of generalized eigenvectors

=

number of ones above the main diagonal in the Jordan form

## About Jordan forms and generalized eigenvectors

The following are equivalent:

- number of Jordan blocks for $\lambda_i$
- geometric multiplicity of $\lambda_i$
- number of true (linearly independent) eigenvectors of $\lambda_i$
- $\dim\{\mathcal{N}(\lambda_i I - A)\}$, i.e., the nullity of $\lambda_i$
- $n - \text{rank}(\lambda_i I - A)$
- dimension of subspace spanned by the eigenvectors
- degeneracy of $(\lambda_i I - A)$

# About Jordan forms and generalized eigenvectors

More equivalences:

- size of the largest Jordan block for $\lambda_i$
- the index of $\lambda_i$
- smallest $k_i$ integer such that $\text{rank}(\lambda_i I - A)^{k_i} = n - m_i$, where $m_i$ is the algebraic multiplicity of $\lambda_i$
- length of the longest chain of eigenvectors - generalized eigenvectors for $\lambda_i$

# Example

By now, you should be able to tell everything about this matrix:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

## Computation of $e^{At}$

Once we have computed the Jordan form of $A$ we can determine $e^{At}$.

Indeed just by using the series expansion one easily verifies

$$e^{At} = Se^{Jt}S^{-1} = S\,\mathrm{diag}(e^{J_1 t}, \ldots, e^{J_g t})S^{-1}.$$

Then it remains to compute $e^{J_l t}$ if $J_l$ has dimension $d$:

$$e^{J_l t} = \begin{pmatrix} 1 & t & \frac{t^2}{2!} & \cdots & \frac{t^{d-2}}{(d-2)!} & \frac{t^{d-1}}{(d-1)!} \\ 0 & 1 & t & \cdots & \frac{t^{d-3}}{(d-3)!} & \frac{t^{d-2}}{(d-2)!} \\ \vdots & \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & t \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix} e^{\lambda_l t}.$$

This leads to an explicit formula for $e^{At}$. In Matlab use `expm(A*t)`. Observe that this proves the statement on 22. Also note how it reduces to the formula for diagonalizable $A$.

# Complex Eigenvalues

Partition the columns of $S$ and the rows of $T = S^{-1}$ according to $J$ as

$$S = \begin{pmatrix} C_1 & \cdots & C_g \end{pmatrix} \quad \text{and} \quad T = \begin{pmatrix} R_1 \\ \vdots \\ R_g \end{pmatrix}.$$

Then the next formula reveals how the modes of $A$ contribute to $e^{At}$:

$$e^{At} = C_1 e^{J_1 t} R_1 + C_2 e^{J_2 t} R_2 \cdots + C_g e^{J_g t} R_g$$

If $\lambda_k$ is real we can make sure that $C_k$ and $R_k$ are real.
If $\lambda_k$ is complex and $\lambda_k = \bar{\lambda}_l$, we can enforce $C_k = \bar{C}_l$ and $R_k = \bar{R}_l$.

For example if $\lambda_1 = \bar{\lambda}_2$ are complex and all other eigenvalues are real, we then obtain a representation which is analogous to that on slide 9:

$$\begin{aligned} e^{At} &= C_1 e^{J_1 t} R_1 + \bar{C}_1 e^{\bar{J}_1 t} \bar{R}_1 + C_3 e^{J_3 t} R_3 + \cdots + C_g e^{J_g t} R_g = \\ &= 2\mathrm{Re}[C_1 e^{J_1 t} R_1] + C_3 e^{J_3 t} R_3 + \cdots + C_g e^{J_g t} R_g. \end{aligned}$$

## Relation to Eigenvectors

Another way to read the formulas is

$$A \begin{pmatrix} C_1 & \cdots & C_g \end{pmatrix} = \begin{pmatrix} C_1 & \cdots & C_g \end{pmatrix} \mathrm{diag}(J_1, \ldots, J_g)$$

$$e^{At} \begin{pmatrix} C_1 & \cdots & C_g \end{pmatrix} = \begin{pmatrix} C_1 & \cdots & C_g \end{pmatrix} \mathrm{diag}(e^{J_1 t}, \ldots, e^{J_g t})$$

or $AC_k = C_k J_k$ and $e^{At}C_k = C_k e^{J_k t}$ for all $k = 1, \ldots, g.$

If $e_k$ denote the first columns of $C_k$ we infer from these two relations:

- $Ae_k = e_k \lambda_k = \lambda_k e_k$. Hence the collection $e_1, \ldots, e_g$ is a (largest) set of linearly independent eigenvectors of $A$ (as on slide 13).

- $e^{At} e_k = e_k e^{\lambda_k t}$. Hence with $\lambda_k = \sigma_k + i\omega_k$ we infer

$$e^{At}\mathrm{Re}(e_k) = \mathrm{Re}(e_k e^{\lambda_k t}) = e^{\sigma_k t}[\mathrm{Re}(e_k)\cos(\omega_k t) - \mathrm{Im}(e_k)\sin(\omega_k t)]$$
$$e^{At}\mathrm{Im}(e_k) = \mathrm{Im}(e_k e^{\lambda_k t}) = e^{\sigma_k t}[\mathrm{Im}(e_k)\cos(\omega_k t) + \mathrm{Re}(e_k)\sin(\omega_k t)].$$

## Example

Consider again the matrix on slide 24. We infer that

$$e^{At}\underbrace{\begin{pmatrix} -3 & -1 & 1 & 5 & 3 \\ -2 & -1 & 1 & 4 & 3 \\ 0 & 0 & 1 & 1 & 1 \\ -1 & 0 & 1 & 2 & 2 \\ 2 & 1 & -1 & -3 & -2 \end{pmatrix}}_{S} =$$

$$= \begin{pmatrix} -3 & -1 & 1 & 5 & 3 \\ -2 & -1 & 1 & 4 & 3 \\ 0 & 0 & 1 & 1 & 1 \\ -1 & 0 & 1 & 2 & 2 \\ 2 & 1 & -1 & -3 & -2 \end{pmatrix} \left(\begin{array}{cc|cc|c} e^{-t} & te^{-t} & 0 & 0 & 0 \\ 0 & e^{-t} & 0 & 0 & 0 \\ \hline 0 & 0 & e^{t} & te^{t} & 0 \\ 0 & 0 & 0 & e^{t} & 0 \\ \hline 0 & 0 & 0 & 0 & e^{t} \end{array}\right).$$

If $x_0$ is a linear combination of the first two columns of $S$ then $e^{At}x_0$ converges to zero for $t \to \infty$. This follows from

$$\lim_{t\to\infty} e^{-t} = 0 \quad \text{and} \quad \lim_{t\to\infty} te^{-t} = 0.$$

## Asymptotic Stability

The stability analysis of autonomous systems is related to examining the **asymptotic behavior** of $x(t)$ if $t \to \infty$. For linear systems the explicit representation of the solutions clearly allows to answer such questions. As an illustration we formulate the most important concept as follows.

The (equilibrium $0$ of the) linear system $\dot{x} = Ax$ is **asymptotically stable** if all solutions satisfy $\lim_{t\to\infty} x(t) = 0$.

Since convergence holds for all initial conditions, this is often called **global** asymptotic stability. The dynamic property of asymptotic stability can be verified by a simple algebraic test.

The system $\dot{x} = Ax$ is asymptotically stable if and only if all the eigenvalues of $A$ have a negative real part. Matrices $A$ with this property are called **Hurwitz**.

# Proof

1) Suppose that all eigenvalues of $A$ have negative real parts. If $\lambda = \sigma + i\omega$ is such an eigenvalue we have $\sigma < 0$. For $t \geq 0$ we infer

$$|t^k e^{\lambda t}| = |e^{\sigma t}[t^k e^{i\omega t}]| = e^{\sigma t} t^k \to 0 \ \text{ for } \ t \to \infty$$

since "negative exponentials converge faster than polynomials diverge". Since all elements of $e^{At}$ are linear combinations of such terms we infer

$$\lim_{t\to\infty} e^{At} = 0 \ \text{ and hence } \ \lim_{t\to\infty} e^{At} x_0 = 0 \ \text{ for all } \ x_0 \in \mathbb{R}^n.$$

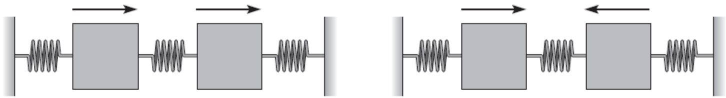
2) Conversely suppose that $\lambda = \sigma + i\omega$ is an eigenvalue of $A$ with $\sigma \geq 0$ with eigenvector $v \neq 0$. Then $e^{At}v = ve^{\lambda t}$ and hence

$$x(t) = e^{At}\mathsf{Re}(v) = e^{\sigma t}[\mathsf{Re}(v)\cos(\omega t) - \mathsf{Im}(v)\sin(\omega t)].$$

Since either $\mathsf{Re}(v) \neq 0$ or $\mathsf{Im}(v) \neq 0$ and since $\sigma \geq 0$, we infer $x(t) \not\to 0$ for $t \to \infty$. Hence we have found an initial condition $\mathsf{Re}(v)$ for which the solution does not converge to zero for $t \to \infty$.

# Example

Consider the mechanical system ([AM] pages 142-145) depicted in




and described by

$$\frac{d}{dt}\begin{pmatrix} q_1 \\ q_2 \\ \dot{q}_1 \\ \dot{q}_2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\frac{2k}{m} & \frac{k}{m} & -\frac{c}{m} & 0 \\ \frac{k}{m} & -\frac{2k}{m} & 0 & -\frac{c}{m} \end{pmatrix}\begin{pmatrix} q_1 \\ q_2 \\ \dot{q}_1 \\ \dot{q}_2 \end{pmatrix}.$$

The coordinate change $z_1 = \frac{1}{2}(q_1 + q_2)$, $z_2 = \dot{z}_1$, $z_3 = \frac{1}{2}(q_1 - q_2)$, $z_4 = \dot{z}_3$ leads to

$$\begin{pmatrix} \dot{z}_1 \\ \dot{z}_2 \\ \dot{z}_3 \\ \dot{z}_4 \end{pmatrix} = \left(\begin{array}{cc|cc} 0 & 1 & 0 & 0 \\ -\frac{k}{m} & -\frac{c}{m} & 0 & 0 \\ \hline 0 & 0 & 0 & 1 \\ 0 & 0 & -\frac{3k}{m} & -\frac{c}{m} \end{array}\right)\begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{pmatrix} = \underbrace{\begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}}_{A}\begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{pmatrix}.$$

## Example

The modes of the system are the eigenvalues of

$$A_1 = \begin{pmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{c}{m} \end{pmatrix} \quad \text{and} \quad A_2 = \begin{pmatrix} 0 & 1 \\ -\frac{3k}{m} & -\frac{c}{m} \end{pmatrix}.$$

If all constants are positive these are Hurwitz. Therefore the system on the previous slide is asymptotically stable.

A real $2 \times 2$ matrix $A$ is Hurwitz iff $\mathrm{trace}(A) < 0$ and $\det(A) > 0$.

From the definition of the matrix exponential we also infer

$$e^{At} = \begin{pmatrix} e^{A_1 t} & 0 \\ 0 & e^{A_2 t} \end{pmatrix}.$$

If $z_3(0) = 0$, $z_4(0) = 0$ then $z_3(t) = 0$, $z_4(t) = 0$ as well as

$$\begin{pmatrix} z_1(t) \\ z_2(t) \end{pmatrix} = e^{A_1 t} \begin{pmatrix} z_1(0) \\ z_2(0) \end{pmatrix}.$$

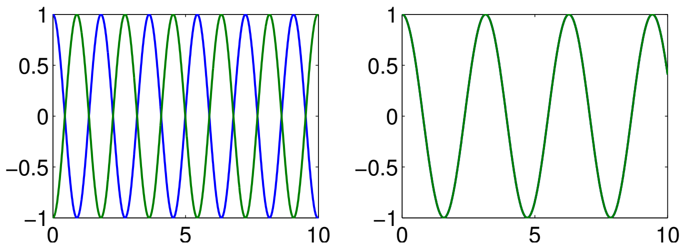
Since $q_1(t) = q_2(t)$ this is the movement in Mode 1 (see figure).

## Example

If there is **no damping** we have $c = 0$. With $\omega_1 = \sqrt{\frac{k}{m}}$ and $\omega_2 = \sqrt{\frac{3k}{m}}$ we then infer

$$e^{A_\nu t} = \begin{pmatrix} \cos(\omega_\nu t) & \frac{1}{\omega_\nu}\sin(\omega_\nu t) \\ -\omega_\nu \sin(\omega_\nu t) & \cos(\omega_\nu t) \end{pmatrix} \quad \text{for} \quad \nu = 1, 2.$$

All system motions are sustained oscillations. We can as well distinguish the two modes of movement as indicated in the figure:




Systems whose solutions do not decay to zero but only stay bounded for future times, as in this example, have their own interest.

## Lyapunov Stability

Often we are only interested in analyzing whether the solutions stay bounded for $t \to \infty$ and do not explode. This is related to what is called **Lyapunov stability** or neutral stability.

All solutions of the system $\dot{x} = Ax$ are bounded for $t \to \infty$ iff all eigenvalues of $A$ have a non-positive real part, and all Jordan blocks of eigenvalues with real part zero have dimension $1$.

- Lyapunov stability boils down to $e^{At}$ being bounded for $t \to \infty$. Hence solutions that start close to zero stay close to zero.
- Asymptotic stability just means $\lim_{t\to\infty} e^{At} = 0$. Note that the decay is actually **exponential**.

All other more refined stability questions (decay rates, frequency of oscillations) can be obtained by directly analyzing $e^{At}$.

# Nonlinear Systems

Stability analysis for nonlinear systems is substantially harder and we only touch upon some few aspects. Consider

$$\dot{x} = f(x) \;\text{ with }\; f : \mathcal{D} \to \mathbb{R}^n \;\text{ defined on a domain }\; \mathcal{D} \subset \mathbb{R}^n.$$

Suppose that $x_e \in \mathcal{D}$ is an equilibrium and hence satisfies $f(x_e) = 0$. Imprecise definitions of stability of $x_e$ in the sense of Lyapunov read as:

- $x_e$ is said to be (locally) **stable** if all solutions for which $x(0) \in \mathcal{D}$ is close to $x_e$ stay close to $x_e$ for all $t \geq 0$.
- $x_e$ is said to be (locally) **attractive** if all solutions for which $x(0) \in \mathcal{D}$ is close to $x_e$ satisfy $\lim\limits_{t\to\infty} x(t) = x_e$.
- $x_e$ is (locally) **asymptotically stable** if it is stable and attractive.
- $x_e$ is **globally asymptotically stable** if it is stable and $\lim\limits_{t\to\infty} x(t) = x_e$ holds for all initial conditions $x(0) \in \mathcal{D}$.

# Lyapunov Functions

A continuously differentiable function $V : \mathcal{D} \to \mathbb{R}$ is said to be a **Lyapunov function** candidate for the nonlinear system $\dot{x} = f(x)$ if

$$[\partial_x V(x)] \cdot f(x) \leq 0 \quad \text{for all} \quad x \in \mathcal{D}.$$

This is an algebraic property which is purely expressed in terms of the function $f$ and the gradient of $V$ without involving any system trajectory.

As the key feature, this property implies for any state-trajectory of the nonlinear system in $\mathcal{D}$, by the chain-rule, that

$$\frac{d}{dt}V(x(t)) = \partial_x V(x(t))\dot{x}(t) = \partial_x V(x(t)) f(x(t)) \leq 0.$$

Therefore along all state-trajectories of $\dot{x} = f(x)$ in $\mathcal{D}$, the function $V(x(t))$ is **monotonically non-increasing** for increasing times.

This feature turns $V$ into a "potential-function". Intuitively, state-trajectories should hence "converge to points in which $V$ is minimal".

## Lyapunov Stability Theorem

This is made precise in the following Lyapunov theorem.

Suppose $V(x)$ is a Lyapunov function candidate for $\dot{x} = f(x)$ with $f(x_e) = 0$.

1. If $V(x) > V(x_e)$ for all $x \in \mathcal{D} \setminus \{x_e\}$ then $x_e$ is stable.
2. If $V(x) > V(x_e)$ and $\partial_x V(x) f(x) < 0$ for all $x \in \mathcal{D} \setminus \{x_e\}$ then $x_e$ is asymptotically stable.

- The result is often formulated for $x_e = 0$ and $V(0) = 0$. Functions with $V(x) > 0$ for all $x \in \mathcal{D} \setminus \{0\}$ are called positive definite.
- Guaranteeing stability by searching for a Lyapunov function is the so-called "Direct Method of Lyapunov". There are no general recipes for doing this. Don't forget: The stability properties are local!
- If $\mathcal{D} = \mathbb{R}^n$ and if, in addition, $V$ in 2. is also radially unbounded ($V(x) \to \infty$ for $\|x\| \to \infty$) then $x_e$ is globally asymptotically stable.

# Example

Consider the unactuated mass-spring-damper system from Lecture 1:

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ -\frac{k}{m}x_1 - \frac{1}{m}c(x_2) \end{pmatrix} = f(x_1, x_2).$$

Suppose that $c(x_2) = 0$ exactly for $x_2 = 0$ such that $x_e = (0, 0)$ is the unique equilibrium. Motivated by the total energy define

$$V(x_1, x_2) = \frac{1}{2}kx_1^2 + \frac{1}{2}mx_2^2.$$

Note that $V(x) > V(0) = 0$ for all $x \neq 0$. Moreover

$$\partial_x V(x) f(x) = \begin{pmatrix} kx_1 & mx_2 \end{pmatrix} \begin{pmatrix} x_2 \\ -\frac{k}{m}x_1 - \frac{1}{m}c(x_2) \end{pmatrix} = -x_2 c(x_2).$$

Suppose that $c(\cdot)$ is an odd function such that $x_2 c(x_2) \geq 0$ for all $x_2 \in \mathbb{R}$. This implies that $x_e = 0$ is Lyapunov stable.

We are not able to conclude asymptotic stability from our result.

# Indirect Method of Lyapunov

Remember that the linearization of $\dot{x} = f(x)$ at $x_e$ is defined as

$$\dot{x}_\Delta = Ax_\Delta \quad \text{with} \quad A = \partial_x f(x_e).$$

Recall that we "hope" for $x(t) \approx x_e + x_\Delta(t)$: A solution of the linear system leads to a good approximation of that of the nonlinear system.

This hope is reality, at least for local asymptotic stability.

Suppose that the linearization is asymptotically stable ($A$ Hurwitz). Then $x_e$ is a locally asymptotically stable equilibrium of $\dot{x} = f(x)$.

Colloquially, asymptotic stability of the linearization also implies local asymptotic stability of the original non-linear system around the point of linearization.

This is a major motivation why to study linear systems in such depth!

## Example

Consider again the example on slide 43. The linearization at $x_e = 0$ is

$$\partial_x f(x) = \begin{pmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{1}{m}c'(x_2) \end{pmatrix} \quad \text{and thus} \quad A = \begin{pmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{1}{m}c'(0) \end{pmatrix}.$$

Recall that

$$\begin{pmatrix} 0 & 1 \\ -\alpha & -\beta \end{pmatrix} \quad \text{is Hurwitz iff} \quad \alpha > 0 \quad \text{and} \quad \beta > 0.$$

Hence the matrix $A$ of the linearization is Hurwitz if and only if $c'(0) > 0$.

$x_e = 0$ is locally asymptotically stable in case that $c'(0) > 0$.

- Note that $c'(0) > 0$ implies $x_2 c(x_2) > 0$ for $x_2 \neq 0$ close to zero.
- The very powerful Krasovski-Lasalle principle allows to actually show that $x_e = 0$ is globally asymptotically stable if $c(x_2)x_2 > 0$ for $x_2 \neq 0$.

## Variation-of-Constants-Formula

Let's come back to linear systems with inputs and outputs described as

$$\dot{x} = Ax + Bu \quad \text{and} \quad y = Cx + Du.$$

The matrix exponential allows to obtain a highlight result, an explicit solution formula for such systems.

For a given input function $u(t)$ and the initial condition $x(0) = x_0$, the unique system response is given by

$$x(t) = e^{At}x_0 + \int_0^t e^{A(t-\tau)}Bu(\tau)\,d\tau.$$

The output response hence is

$$y(t) = Cx(t) + Du(t) = Ce^{At}x_0 + \int_0^t [Ce^{A(t-\tau)}B]u(\tau)\,d\tau + Du(t).$$

Thus the solution is represented by a **convolution integral**.

# Proof

**Important Trick:** Search for solution $x(t) = e^{At}z(t)$ with suitable $z(t)$.

Differentiation implies

$$\dot{x}(t) = Ae^{At}z(t) + e^{At}\dot{z}(t) = Ax(t) + e^{At}\dot{z}(t).$$

This equals $Ax(t) + Bu(t)$ as desired if

$$e^{At}\dot{z}(t) = Bu(t) \quad \text{or} \quad \dot{z}(\tau) = e^{-A\tau}Bu(\tau).$$

Hence for any constant vector $c$ the function

$$z(t) = c + \int_0^t e^{-A\tau}Bu(\tau)\,d\tau$$

does the job. $c = x_0$ leads to satisfaction of $x(0) = x_0$. ∎

Note that this trick of **varying the constants** can often be directly applied for finding responses to explicitly given input signals.

## Example

Suppose the input is constant: $u(t) = u_e$ for all $t \geq 0$. Then

$$x(t) = e^{At}x_0 + \left(\int_0^t e^{A(t-\tau)}\,d\tau\right) Bu_e = e^{At}x_0 + \left(\int_0^t e^{A\rho}\,d\rho\right) Bu_e.$$

If $A$ is Hurwitz then it is invertible and

$$\int_0^t e^{A\rho}\,d\rho = \int_0^t \frac{d}{d\rho}e^{A\rho}A^{-1}\,d\rho = e^{At}A^{-1} - A^{-1}.$$

The state-response can hence be written with the matrix exponential as

$$x(t) = e^{At}[x_0 + A^{-1}Bu_e] - A^{-1}Bu_e.$$

For $t \to \infty$ we conclude $x(t) \to -A^{-1}Bu_e =: x_e$. Therefore the state converges to the unique solution of $Ax_e + Bu_e = 0$. Moreover

$$y(t) = Ce^{At}[x_0 + A^{-1}Bu_e] + [D - CA^{-1}B]u_e.$$

This nicely displays the **transient** and the **steady-state** response. The matrix $D - CA^{-1}B$ is called **steady-state gain** of the system.

## Superposition Principle

Clearly the matrix exponential $e^{At}$ determines the state-response both to non-zero initial conditions and to forcing inputs. The response depends **linearly** on both $x_0$ and $u(\cdot)$. This is called **superposition principle**.

If the system has multiple inputs we can partition

$$u(t) = \begin{pmatrix} u_1(t) \\ \vdots \\ u_m(t) \end{pmatrix}, \quad B = \begin{pmatrix} B_1 & \cdots & B_m \end{pmatrix}, \quad D = \begin{pmatrix} D_1 & \cdots & D_m \end{pmatrix}.$$

By the superposition principle the full output response is the sum of

$$Ce^{At}x_0 \quad \text{and} \quad \int_0^t Ce^{A(t-\tau)}B_k u_k(\tau)\,d\tau + D_k u_k(t), \quad k = 1, \ldots, m$$

and each of these contributions can be analyzed separately.

## Step and Impulse Response

Let $u_k(\cdot)$ be equal to the **step function** $s(t) = 1$ for $t \geq 0$. Then

$$\int_0^t Ce^{A(t-\tau)} B_k s(\tau)\, d\tau + D_k s_k(t) = \int_0^t Ce^{A\rho} B_k\, d\rho + D_k$$

is the **step response** of the system (in the $k$-th input).

If $u_k(\cdot)$ is equal to the impulse $\delta(\cdot)$ at zero then

$$\int_0^t Ce^{A(t-\tau)} B_k \delta(\tau)\, d\tau + D_k \delta(t) = Ce^{At} B_k + D_k \delta(t)$$

is the **impulse response** of the system (in the $k$-th input). Observe that it is the derivative of the step response.

The step- and impulse-responses of the system are given by

$$\int_0^t [Ce^{A\rho} B]\, d\rho + D \quad \text{and} \quad Ce^{At} B + D\delta(t)$$

and can be obtained by applying $m$ steps/impulses for each input.

# Step and Impulse Response

Steps and impulses should be viewed as test signals that allow to gather information about the dynamical behavior of a system.

Both responses can be obtained by exciting the system with $m$ well-defined input signals from initial condition zero. Knowledge of the impulse response

$$H(t) = Ce^{At}B + D\delta(t)$$

then allows to determine the **response for any other signal** by

$$\int_0^t H(t-\tau)u(\tau)\,d\tau.$$

This is a principle path in order to determine a system description from a finite number of (experimental) responses. You learn more about these issues in courses on system identification.

## Sinusoidal Inputs

For a complex $\lambda = \sigma + i\omega \in \mathbb{C}$ and $u_e \in \mathbb{R}^m$ consider the input

$$u(t) = u_e\, e^{\lambda t} = u_e\, e^{\sigma t}[\cos(\omega t) + i \sin(\omega t)].$$

Suppose $\lambda I - A$ is invertible. As on slide 48 we then have with $\rho = t - \tau$:

$$y(t) = C\left(e^{At}x_0 + \left[\int_0^t e^{A(t-\tau)}e^{\lambda\tau}\, d\tau\right] Bu_e\right) + D(u_e e^{\lambda t}) =$$

$$= C\left(e^{At}x_0 + e^{\lambda t}\left[\int_0^t e^{(A-\lambda I)\rho}\, d\rho\right] Bu_e\right) + D(u_e e^{\lambda t}) =$$

$$= C\left(e^{At}x_0 + e^{\lambda t}\left[e^{(A-\lambda I)t} - I\right](A - \lambda I)^{-1}Bu_e\right) + D(u_e e^{\lambda t}).$$

Reordering leads to the most important formula

$$y(t) = Ce^{At}\left[x_0 + (A - \lambda I)^{-1}Bu_e\right] + \left[C(\lambda I - A)^{-1}B + D\right](u_e e^{\lambda t})$$

which displays the transient and steady-state response. (Notation makes sense if $A$ is Hurwitz - then the transient decays to zero for $t \to \infty$.)

## Sinusoidal Inputs: Summary

For the exponentially weighted sinusoidal complex input signal

$$u(t) = u_e\, e^{\lambda t} = u_e\, e^{\sigma t}[\cos(\omega t) + i \sin(\omega t)] \quad (\lambda = \sigma + i\omega)$$

such that $\lambda I - A$ is invertible, the state-response is

$$x(t) = e^{At}\left[x_0 - (\lambda I - A)^{-1} B u_e\right] + \left[(\lambda I - A)^{-1} B\right](u_e e^{\lambda t})$$

and the output-response is

$$y(t) = Ce^{At}\left[x_0 - (\lambda I - A)^{-1} B u_e\right] + \left[C(\lambda I - A)^{-1} B + D\right](u_e e^{\lambda t}).$$

Since $A$, $B$, $C$, $D$ and $x_0$ are real, the responses to the inputs

$$v(t) = u_e\, e^{\sigma t} \cos(\omega t) \quad \text{and} \quad w(t) = u_e\, e^{\sigma t} \sin(\omega t)$$

are just obtained by taking the real and imaginary part.

For $\sigma = 0$ (pure sinusoidal inputs) the response is determined by the **frequency response** matrix $C(i\omega I - A)^{-1} B + D$.

## Solutions with Laplace Transform

If the signal $x(t)$ is defined for $t \geq 0$ its Laplace transform is

$$\hat{x}(s) = \int_0^\infty x(t)e^{-st}\,dt$$

for all those complex $s \in \mathbb{C}$ for which the integral is finite.

Taking the transforms of all system signals on slide 46 leads to

$$s\hat{x}(s) - x(0) = A\hat{x}(s) + B\hat{u}(s), \quad \hat{y}(s) = C\hat{x}(s) + D\hat{u}(s).$$

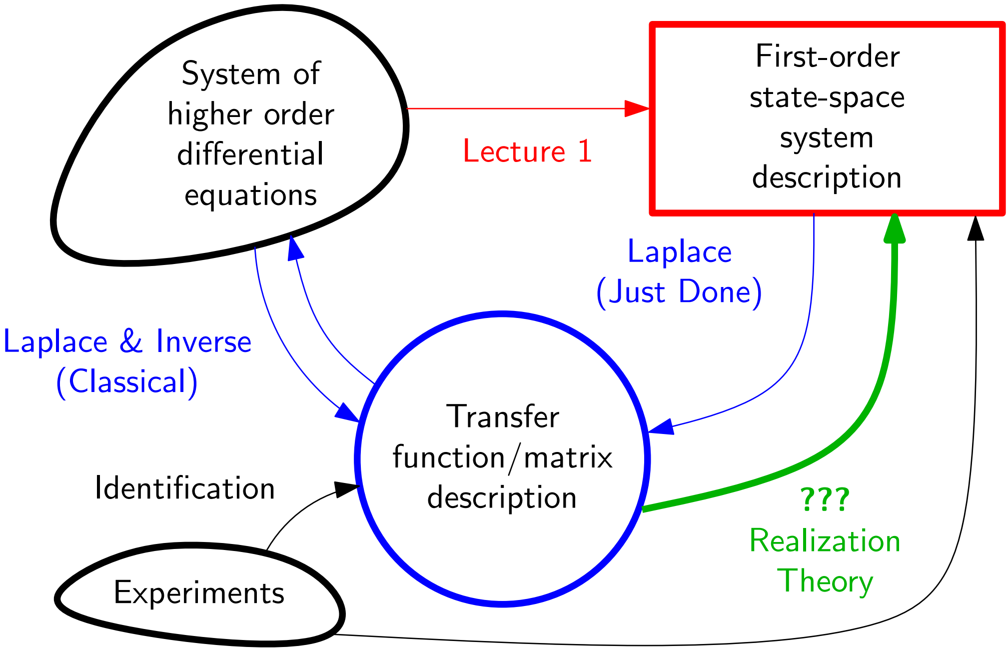
We can hence **algebraically** solve for $\hat{x}(s)$ to get with $x(0) = x_0$ that

$$\hat{x}(s) = (sI - A)^{-1}x_0 + (sI - A)^{-1}B\hat{u}(s),$$

$$\hat{y}(s) = C(sI - A)^{-1}x_0 + [C(sI - A)^{-1}B + D]\hat{u}(s).$$

This is the frequency-domain analogue of the time-domain formulas on slide 46. The inverse Laplace transform leads to $x(t)$ and $y(t)$.

# Relation to Classical Techniques and Identification



## Transfer Matrices

Given the system on slide 46 we have seen that the matrix

$$G(s) = C(sI - A)^{-1}B + D$$

depending the the complex variable $s \in \mathbb{C}$ plays a major role.

- For $s \in \mathbb{C}$ not being an eigenvalue of $A$ we can compute $(sI - A)^{-1}$.
- The elements of $(sI - A)^{-1}$ are **rational functions** of $s$, since

  $$(sI - A)^{-1} = \frac{1}{\chi(s)} \operatorname{adj}(sI - A) \quad \text{with} \quad \chi(s) = \det(sI - A)$$

  where $\operatorname{adj}$ is the algebraic adjoint. In fact each element can be written as $\frac{n_{ij}(s)}{\chi(s)}$ with a polynomial $n_{ij}(s)$ of degree strictly smaller than that of $\chi(s)$. Such rational functions are called **strictly proper**.
- The elements of $C(sI - A)^{-1}B + D$ are linear combinations of those of $(sI - A)^{-1}$ plus a constant matrix $D$.

## Transfer Matrices

$C(sI - A)^{-1}B + D$ is called the **transfer matrix** corresponding to the system with state-space description $\dot{x} = Ax + Bu$, $y = Cx + Du$.

All elements of $G(s) = C(sI - A)^{-1}B + D$ are rational functions of $s$ whose numerator degree is not larger than that of the denominator. Such rational functional are called **proper** and $G(s)$ is said to be a **proper rational matrix**. In Matlab $G(s)$ is computed by `tf`.

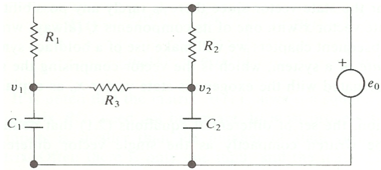
The **poles** of the transfer matrix $G(s)$ are determined as follows:

- Write each element of $G(s)$ in the form $\frac{n_{ij}(s)}{d_{ij}(s)}$ where the numerator and denominator polynomials **have no common zeros**.
- The union of all zeros of $d_{ij}(s)$ for all elements are the poles of $G(s)$.

$G(s)$ is said to be **stable** if all its poles have negative real parts.

# Example

Consider the electrical network from [F] p.21




and described by

$$\begin{pmatrix} \dot{v}_1 \\ \dot{v}_2 \end{pmatrix} = -\begin{pmatrix} \frac{1}{C_1R_1}+\frac{1}{C_1R_3} & -\frac{1}{C_1R_3} \\ -\frac{1}{C_2R_3} & \frac{1}{C_2R_2}+\frac{1}{C_2R_3} \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} + \begin{pmatrix} \frac{1}{R_1C_1} \\ \frac{1}{R_2C_2} \end{pmatrix} u.$$

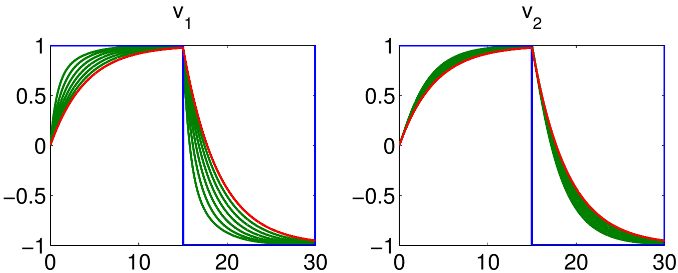
With $y = v$ and $C_1 = 1$, $C_2 = 2$, $R_1 = 1$, $R_2 = 2$, $R_3 = 3$ we have

$$G(s) = \begin{pmatrix} \frac{s+0.5}{s^2+1.75s+0.5} \\ \frac{0.25s+0.5}{s^2+1.75s+0.5} \end{pmatrix}.$$

The poles are the zeros of $s^2 + 1.75s + 0.5$ which are $-1.39,\ -0.36$.
$G(s)$ is stable. The steady-state gain is $G(0) = D - CA^{-1}B = (1\ \ 1)^T$.

# Example

Let's change $C_1$ in steps from $1$ to $4$ and consider the step responses from zero initial conditions:




For $C_1 = 4$ the transfer matrix is

$$G(s) = \begin{pmatrix} \frac{0.25s+0.125}{s^2+0.75s+0.125} \\ \frac{0.25s+0.125}{s^2+0.75s+0.125} \end{pmatrix} = \begin{pmatrix} \frac{0.25}{s+0.25} \\ \frac{0.25}{s+0.25} \end{pmatrix}.$$

We had two different poles for $C_1 = 1$ which reveal themselves in the different time-constants of the response. For $C_1 = 4$ only one pole $\{-0.25\}$ remains related to the red response. Why? We'll explain later!

# Working with Transfer Matrices: Summary

In view of slide 54 the transfer matrix is most useful in case that the **initial condition is zero**. For a given input the transfer matrix fully determines the output response.

In addition, if two systems have transfer matrices $G_1(s)$ and $G_2(s)$ then the parallel and series interconnections have transfer matrices

$$G_1(s) + G_2(s) \;\; \text{and} \;\; G_1(s)G_2(s)$$

with the **usual matrix operations**. Note that the dimensions must be compatible, and the **order** of the multiplication is important to be kept!

Moreover if $A$ is Hurwitz and $\lambda$ is not an eigenvalue of $A$ then

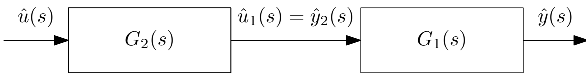
$$G(0)u_e, \quad G(i\omega)e^{i\omega t}u_e, \quad G(\lambda)e^{\lambda t}u_e$$

are the steady-state responses for constant, purely sinusoidal and exponentially weighted sinusoidal inputs in the direction $u_e$.

# Block-Diagrams

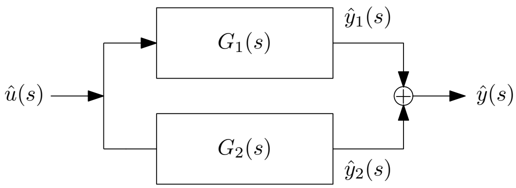
Parallel and series interconnections then translate into the block-diagrams as you are used to work with from classical control:

Series interconnection: $\hat{y}(s) = [G_1(s)G_2(s)]\hat{u}(s)$ is depicted as




Notice the order of the blocks!

Parallel interconnection: $\hat{y}(s) = [G_1(s) + G_2(s)]\hat{u}(s)$ is depicted as



## Realizations

A state-space system (uniquely) determines its transfer matrix by a simple computation. The converse question reads as follows.

Let us be given a $k \times m$-matrix $G(s)$ whose elements consist of proper rational functions. Can we compute matrices $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{k\times n}$ and $D \in \mathbb{R}^{k\times m}$ such that

$$G(s) = C(sI - A)^{-1}B + D\ ?$$

Let us introduce the following important terminology.

If $G(s) = C(sI - A)^{-1}B + D$ then $(A, B, C, D)$ is said to be a **realization** of $G(s)$.

The above formulated question is called the **realization problem**. In Matlab just use `ss`. The theory will be discussed in detail in Lecture 5.

## Coordinate Change

The matrices $(A, B, C, D)$ of a realization of $G(s)$ are **never unique**. One reason (there are many more!) is as follows.

A state-coordinate change modifies the describing matrices of a state-space system but it **does not change** the transfer matrix.

More precisely, for the state-space system

$$\dot{x} = Ax + Bu, \quad y = Cx + Du$$

perform the coordinate change $z = Tx$. This leads to

$$\dot{z} = \tilde{A}z + \tilde{B}u, \;\; y = \tilde{C}z + \tilde{D}u \;\; \text{with} \;\; \begin{pmatrix} \tilde{A} & \tilde{B} \\ \tilde{C} & \tilde{D} \end{pmatrix} = \begin{pmatrix} TAT^{-1} & TB \\ CT^{-1} & D \end{pmatrix}$$

while we have $\tilde{C}(sI - \tilde{A})^{-1}\tilde{B} + \tilde{D} = C(sI - A)^{-1}B + D.$

# Proofs

If we perform for $\dot{x} = Ax + Bu$, $y = Cx + Du$ the coordinate change $z = Tx$, we infer $x = T^{-1}z$ and we can thus conclude

$$\dot{z} = T\dot{x} = TAx + TBu = [TAT^{-1}]z + [TB]u \text{ and } y = [CT^{-1}]z + Du.$$

For $\tilde{A} = TAT^{-1}$, $\tilde{B} = TB$, $\tilde{C} = CT^{-1}$, $\tilde{D} = D$ this is nothing but

$$\dot{z} = \tilde{A}z + \tilde{B}u, \quad y = \tilde{C}z + \tilde{D}u.$$

Invariance of the transfer matrix is shown as follows:

$$\begin{aligned}
&\tilde{C}(sI - \tilde{A})^{-1}\tilde{B} + \tilde{D} = \\
&\qquad = [CT^{-1}]\left(sTT^{-1} - [TAT^{-1}]\right)^{-1}[TB] + D = \\
&\qquad = CT^{-1}T\left(sI - A\right)^{-1}T^{-1}TB + D = \\
&\qquad = C(sI - A)^{-1}B + D.
\end{aligned}$$

## Covered in Lecture 2

- Analysis of autonomous linear systems
  coordinate change, diagonalization, Jordan form, modes, all solutions
  matrix exponential
- Stability and Lyapunov functions
  Hurwitz matrices, Lyapunov functions, Lyapunov stability theorem, indirect method
- Analysis of LTI systems
  Variation-of-constants formula, computation of responses (steady-state, transient), Laplace
- Transfer matrices
  definition, working with transfer matrices, realizations, behavior under state coordinate-change