Ordinary differential equations and mathematical modelling MVE162/MMG511, spring 2023.

Lecture 1

Prerequisite knowledge for the course.

This relatively difficult course uses the whole scope of linear algebra and analysis that Chalmers students from Technical mathematics group and GU students from the group in mathematics learned during the first year. Students with different backgrounds might lack some of this material.

Before starting learning this course it is good to check notions and theorems that are supposed to be known during learning of the present course.

If you miss some of them, check Appendix 1 and Appendix 2 in the course book by Logemann and Ryan, where all necessary mathematical background is discussed in detail.

Some international students might also need to learn Matlab or use other programming tools to make computations in obligatory modeling projects.

Notions from linear algebra and analysis:

Vector space, normed vector space, norm of a matrix. Eigenvectors and eigenvalues of a matrix. Matrix diagonalization.

Cauchy sequence. Complete vector space (another name is Banach space). Open, closed and compact sets in $\mathbb{R}^n$. Continuous functions and their properties on compact sets. Uniform convergence for continuous functions.

Results from analysis:

Linear space $C(I)$ of continuous functions on a compact I is a complete vector space (Banach space). Example A.14, p. 272.

Bolzano-Weierstrass theorem. Theorem A.16, p. 273.

Weierstrass criterion for uniform convergence of functional series. Corollary A.23 , p. 277.

1 Introduction. Initial value problem, existence and uniqueness of solutions.

The main subject of the course is systems of differential equations in the form

$$x'(t) = f\,(t, x(t)) \tag{1}$$

classification and qualitative properties of their solutions. Here $f : J \times G \to \mathbb{R}^n$ is a vector valued function regular enough with respect to the time variable t and the space variable x. J is an interval, G is an open subset of $\mathbb{R}^n$. Equations where the function f is independent of t are called **autonomous**:

$$x'(t) = f(x(t))$$

Finding a function $x(t) : L \to \mathbb{R}^n$ satisfying the equation (1) for $t \in L$ on the interval $L \subset J$ together with the initial condition

$$x(\tau) = \xi \tag{2}$$

for $\tau \in L$ is called the **initial value problem (I.V.P.)**.

The curves $x(t)$ in G corresponding to solutions of (1), have the property that the vector field $f\,(t, x(t)) \in \mathbb{R}^n$ is tangent to the curve $x(t)$ at each time t and point $x(t) \in G$.

One can by integrating the left and rignt hand sides of (1), reformulate the I.V.P. (1),(2) in the form of the integral equation

$$x(t) = \xi + \int_{\tau}^{t} f\,(\sigma, x(\sigma))\,d\sigma \tag{3}$$

Continuous solutions $x(t)$ to the integral equation (3) can be interpreted as generalized solutions to (1),(2) in the case when $f\ (t, x)$ is only piecewise continuous with respect to t and therefore the integral in (3) does not have derivative in some isolated points. If f is continuous, then these two formulations are equivalent by the Newton-Leibnitz theorem (known to swedish students as the main theorem of calculus).

More general notions of solutions can be introduced in the case when $f\,(t, x(t))$ is integrable in the sense of Lebesque, but we do not consider such generalised solutions in this course.

2 Classification of ordinary differential equations and the plan of the course.

1. Equations where the right hand side is independent of time:

$$\begin{aligned} x'(t) &= f(x(t)) \\ f &= f(x), x \in G, \end{aligned}$$

are called **autonomous** as we mentioned before. General differential equations are with $f = f(t, x)$ are called **non-autonomous.**

Autonomous equations have a nice graphical interpretation. One can consider and also draw a picture of the vector field $f : G \to \mathbb{R}^n$. For every point $\xi \in G$ this vector field gives according to the differential equation, the velocity of a possible solution curve $x(t)$ going through the point ξ.

All solutions to an autonomous differential equation have the property that corresponding curves are tangent curves to the vector field $f : G \to \mathbb{R}^n$.

One often calls autonomous differential equations **continuous dynamical systems**.

2. General (**non-autonomous**) linear systems of differential equations in the form

$$x'(t) = A(t)x(t), \quad x(t) \in \mathbb{R}^n, \quad t \in J$$

with a matrix $A(t)$, $A(t) : J \to \mathbb{R}^{n\times n}$ that is a continuous matrix valued function of time t on the interval J. A particular class of non-autonomous linear systems is the class of **periodic linear systems** with periodic matrix $A(t + p) = A(t)$ with some period p.

3. We will also consider linear **non-homogeneous** systems of differential equations in the form

$$x'(t) = A(t)x(t) + g(t), \quad x(t), g(t) \in \mathbb{R}^n, \quad t \in J$$

with a given term $g(t)$ in the right hand side.

4. Linear autonomous systems of differential equations in the form

$$x'(t) = Ax(t), \quad x(t) \in \mathbb{R}^n, \quad t \in \mathbb{R}$$

with a constant matrix A.

The plan for the course is: to consider after some introductory examples and then all these types of equations in the reverse order, from simpler to more complicated: linear autonomous, linear non-autonomous, linear periodic, nonlinear autonomous. At the very end of the course we will consider the existence of solutions in the most general non-linear non-autonomous case. Many ideas will be introduced and exploated first on the example of linear autonomous ODEs. Later these ideas will be developed further and applied in more complicated situations. This way of studying pursues two goals: to have more material for exercises from the very beginning and to introduce several general mathematical ideas in a more "user friendly" way.

The course is divided into two large qualitatively different parts:

A) one - is devoted to linear equations and using and developing some advanced linear algebra, and

B) another one - is devoted to non-linear equations and using reasoning based on relatively advanced analysis.

3 Main types of problems posed for systems of ODEs

I) **Existence and uniquness** of solutions to I.V.P. Finding **maximal interval** of existence of solutions to I.V.P.

We give here two simple examples illustrating that solutions to a differential equation might exist not on any time interval (solutions can blow up - tend to infinity in finite time), and that solutions do not need to be unique (there can be two different solution curves going through one point (t, x))

II) One can for particular classes of equations pose the problem of finding a reasonable analytical description of all solutions to the above equation. Such an expression is called **general solution**.

III) Find particular types of solutions: **equilibrium points** $\eta \in \mathbb{R}^n$ of autonomous systems (points where $f(\eta) = 0$), **periodic solutions**, such that after some period $T > 0$ the solution comes back to the same point: $x(t) = x(t + T)$ for any starting time t.

IV) Find how solutions $x(t)$ behave in the vicinity of an equilibrium point η with $t \to \infty$: it is interesting if they stay close to η starting arbitrarily close to it, or solutions can go go out of η with time $t \to \infty$ for some initial points ξ situated arbitrarily close to η (we will call these properties for **stability** or **instability** of the equilibrium point η).

V) Find a geometric description of the set of all trajectories of solutions to an equation. By trajectory we mean here the curve $x(t)$, that the solution goes along, during the time $t \in I$ when it exists. In the case of autonomous systems of dimension 2 we will call such a picture *phase portrait.*

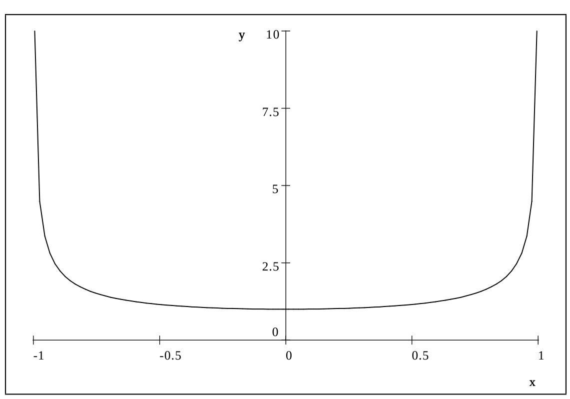
VI) Describe geometric properties of so called **limit sets**, or **"attractors"** of a solution: such a set that the solution $x(t)$ "approaches" infinitely close when $t \to \infty$.

Example of bounded maximal interval. (Ex. 1.2, p.14, L.R.) I.V.P.

$$x'(t) = t \cdot x^3; \quad x(0) = 1$$

. By separation of variables we arrive to a solution that exists only on a finite time interval $(-1, 1)$ called later **maximal interval** for this initial condition.

$$\begin{aligned}
\frac{dx}{x^3} &= t\,dt; \quad \int \frac{dx}{x^3} = \int t\,dt; \quad -\frac{1}{2x^2} = \frac{t^2}{2} + \frac{C}{2}; \quad -\frac{1}{x^2} = t^2 + C; \quad C = -1; \\
-x^2 &= \frac{1}{t^2 - 1}; \quad x^2 = \frac{1}{1 - t^2}; \quad x = \frac{1}{\sqrt{1 - t^2}}
\end{aligned}$$



Point out that for another initial conditions the maximal interval can be different.

Example of non-uniqueness. (Ex.1.1, p.13, L.R.) I.V.P.

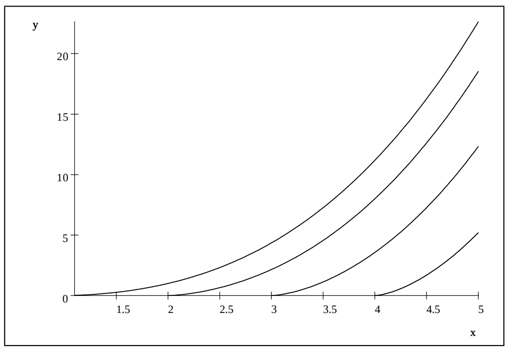
$$x'(t) = t \cdot x^{1/3}, t \in \mathbb{R}, \quad x(0) = 0.$$

Point out that the right hand side has infinite slope in x variable $\frac{d}{dx}(x^{1/3})$. We will say later, after giving corresponding definition, that this function is **not Lipschitz** with respect to x.

Constant solution $x(t) = 0$ exists. On the other hand for all $c > 0$ functions

$$x(t) = \frac{\left(t^2 - c^2\right)^{3/2}}{(3)^{3/2}}, \quad t \geq c$$

are also solutions to the equation. See the calculation below. By extending these solutions by zero to the left from $t = c$ we get a family of different solutions satisfying the same initial conditions $x(0) = 0$.



Calculation of solutions uses separation of variables.

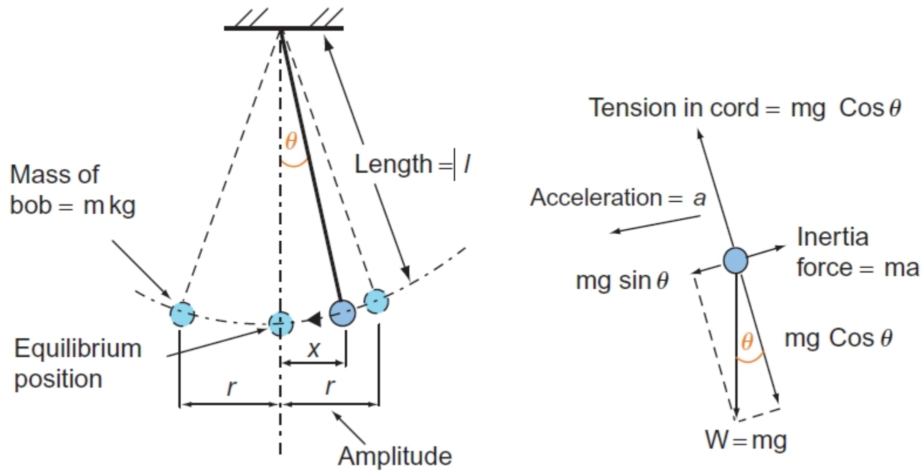
$$\begin{aligned} \frac{dx}{dt} &= tx^{1/3}; \qquad \frac{dx}{x^{1/3}} = tdt \\ \int \frac{dx}{x^{1/3}} &= \int tdt; \qquad \frac{3}{2}x^{\frac{2}{3}} = \frac{1}{2}\left(t^2 - c^2\right) \\ x^{2/3} &= \frac{t^2 - c^2}{3}; \qquad x = \frac{\left(t^2 - c^2\right)^{3/2}}{(3)^{3/2}} \end{aligned}$$

Here c is an arbitrary constant $c \leq t$. Check the solution:

$$\frac{d}{dt}x(t) = \frac{d}{dt}\left(\frac{\left(t^2 - c^2\right)^{3/2}}{(3)^{3/2}}\right) = \frac{1}{3}t\sqrt{3t^2 - 3c^2} = tx^{1/3}$$

Example of equilibrium points and periodic solutions

Pendulum is described by the Newton equation: $Force = m \cdot Acceleration$; $Acceleration = l \cdot \theta''(t)$,$Velocity = l \cdot \theta'(t)$.



(a) Simple pendulum

(b) Forces acting on bob

$$ml\theta''(t) = -\gamma l\theta'(t) - mg\sin(\theta(t)) = 0$$

Both for theoretical analysis and for numerical solution one always rewrites the second order equation as a system of two equations for $x_1(t) = \theta(t)$ and $x_2(t) = \theta'(t)$:

$$\begin{aligned} x_1'(t) &= x_2(t) \\ x_2'(t) &= -\frac{\gamma}{m}x_2(t) - \frac{g}{l}\sin(x_1(t)) \end{aligned}$$

We can rewrite it in general vector form as

$$x'(t) = f(x(t))$$

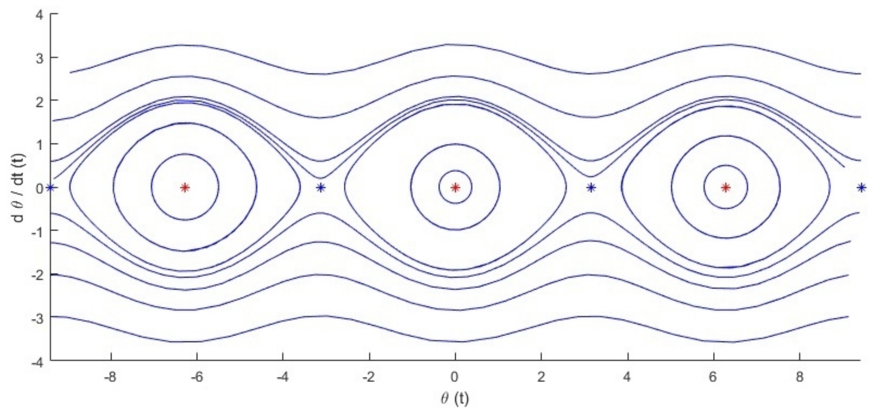
with

$$f(x) = \begin{bmatrix} x_2 \\ -\frac{\gamma}{m}x_2 - \frac{g}{l}\sin(x_1) \end{bmatrix}$$

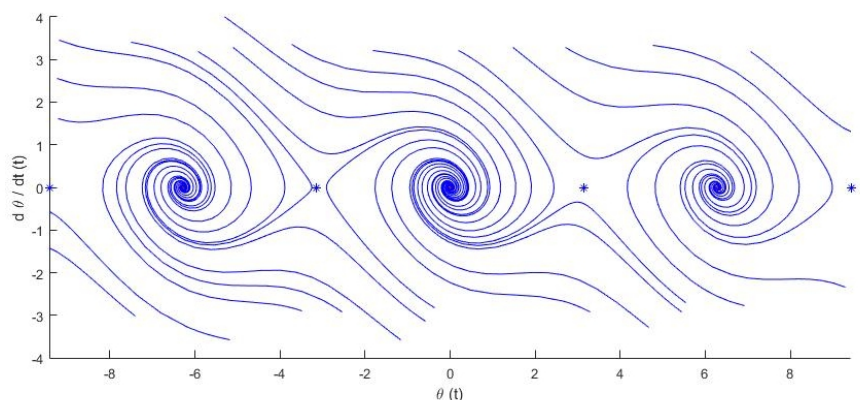
This non-linear system of equations cannot be solved analytically. We show below results of numerical solutions of this system in a form of a **phase portrait of the system.**

Phase portrait.

The picture of trajectories - curves $(x_1(t), x_2(t))$ corresponding different solutions to the equation for the pendulum in the **phase plane** of variables x_1 and x_2 looks as the following. Such pictures are called **phase portrait of the system.** We will draw many of them in this course, in particular in modelling projects.



Phase protrait in the case without friction: $\gamma = 0$



Phase portrait in the case with friction: $\gamma > 0$.

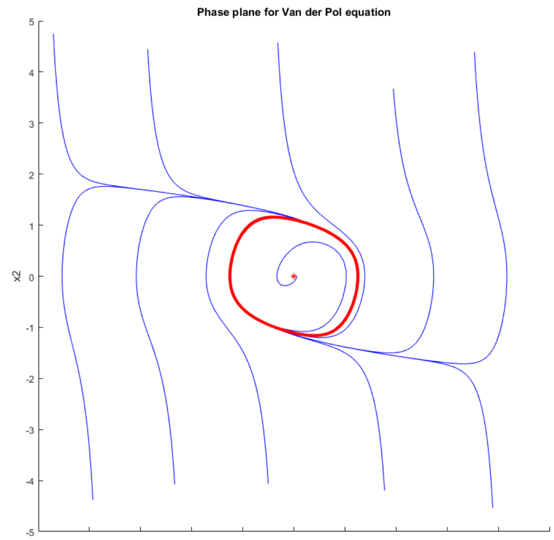
Points $\theta = 0 + 2\pi k$, $\theta' = 0$ and $\theta = \pi + 2\pi k$, $\theta' = 0$ on the first picture are equilibrium points. One can see closed orbits around equilibrium points $\theta = 2\pi k$, $\theta' = 0$, corresponding to periodic solutions. Points $\theta = \pi + 2\pi k$, $\theta' = 0$ correspond to the upper position of the pendulum that is a non-stable equilibrium point. Higher up and down when the angular velocity is large enough we observe non-bounded solutions corresponding to rotation of the pendulum around the pivot. Orbits for the pendulum without friction: $\gamma = 0$, can be described by a non-linear equation.

In the case with friction $\gamma > 0$ on the second picture one observes the same equilibrium points. But the phase portrait is completely different. Almost all trajectories tend to one of equilibrium points $\theta = 2\pi k$, $\theta' = 0$ when time goes to infinity. No closed orbits and no unbounded solutions are observed in this case.

Examples of attractors.

Van der Pol equation. (Example 1.1.1. p. 2 in Logemann/Ryan)

$$\begin{aligned} x'(t) &= f(x(t)) \\ f(x) &= \begin{bmatrix} x_2 \\ -x_1 + x_2(1 - (x_2)^2) \end{bmatrix} \end{aligned}$$



We see that the equilibrium point in the origin is unstable but all trajectories tend to a limit set or "attractor" that is a closed curve (depicted in red) that seems to be an orbit corresponding to a periodic solution.

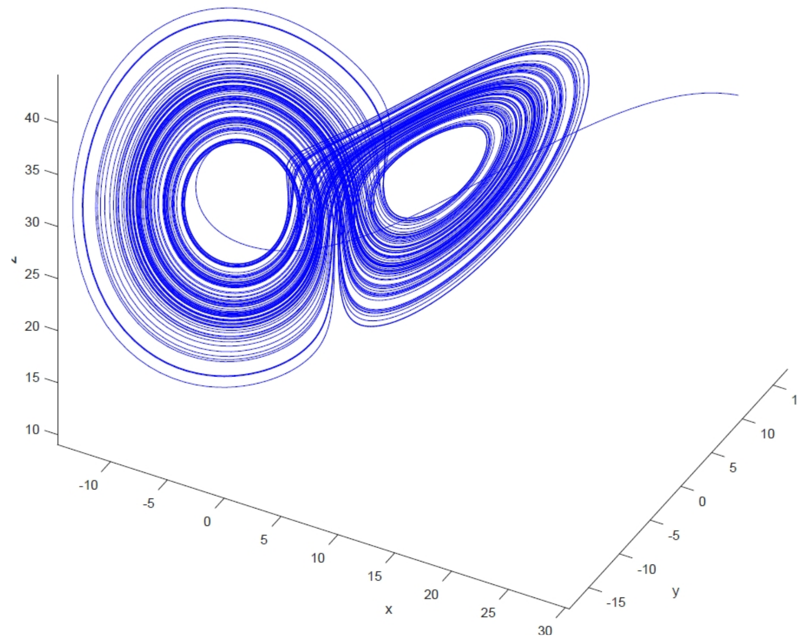
Lorenz's model for turbulence. Strange attractor.

For two dimensional systems only stationary points and closed orbits and some chains of stationary points connected with orbits are possible as "attractors". In dimension 3 much more complicated attractors are possible with a classical example being the Lorenz equation.

$$\begin{aligned} x' &= -\sigma(x-y) \\ y' &= rx - y - xz \\ z' &= xy - bz \end{aligned}$$

An orbit for $\sigma = 10$, $r = 28$, $b = 8/7$.

We can see that the trajectory tends to a set of very complicated structure.



4 Linear autonomous systems of ODEs

We will first consider general concepts in the course in the particular case for linear system of ODEs with constant matrix (linear autonomous systems).

$$x'(t) = Ax(t), \quad x(t) \in \mathbb{R}^n, \quad t \in \mathbb{R} \tag{4}$$

where A is a constant $n \times n$ matrix $A \in \mathbb{R}^{n\times n}$.

In particular we will find solutions to initial value problem (I.V.P.) with initial condition

$$x(\tau) = \xi, \tag{5}$$

We point out that all general results about linear systems of ODE are also valid in the case of solutions found in the complex vector space $x \in \mathbb{C}^n$, $\xi \in \mathbb{C}^n$ and for complex matrix $A \in \mathbb{C}^{n\times n}$. Some of the results are formulated in a more elegant form in the complex case or might be valid only in complex form.

Several general questions that we formulated above will be addressed for this type of systems.

The final goal in this particular case will be to give a detailed analytical description for the set of all solutions and to connect qualitative properties of solutions with specific properties of the matrix A, its eigenvalues and eigenvectors together with more subtle spectral properties such as subspaces of generalised eigenvectors that will be defined later.

4.1 The space of solutions for general non-autonomous linear systems

We make first two simple observations that are valid even for general non-autonomous linear systems with a matrix $A(t)$ that is not constant but is a continuous function of time on the interval J.

$$x'(t) = A(t)x(t), \quad x(t) \in \mathbb{R}^n, \quad t \in J \tag{6}$$

Lemma. The set of solutions $\mathcal{S}_{\text{hom}}$ to (4), and to (6) is a linear vector space.

Proof. $\mathcal{S}_{\text{hom}}$ includes the zero constant vector and is therefore not empty. By the linearity of the time derivative $x'(t)$ and of the matrix multiplication $A(t)x(t)$, for a pair of solutions $x(t)$ and $y(t)$ their sum $x(t) + y(t)$ and the product $Cx(t)$ with a constant C are also solutions to the same equation. Considering equations for $y(t)$ and $x(t)$

$$\begin{aligned} x'(t) &= A(t)x(t) \\ y'(t) &= A(t)y(t) \end{aligned}$$

together with the above arguments we derive the conclusion:

$$\begin{aligned} (x(t) + y(t))' &= A(t)(x(t) + y(t)) \\ (Cx(t))' &= A(t)(Cx(t)) \end{aligned}$$

■

4.2 Uniqueness of solutions to autonomous linear systems.

One shows the uniqueness of solutions to (4) by using a simple version of the Grönwall inequality that in general case will be considered later.

Grönwall inequality

Suppose that the I.V.P. (4),(5) for an autonomous linear system has a solution $x(t)$ on an interval I including τ. Consider the case when $\tau \leq t$.

We can write an equivalent integral equation for $x(t)$ for $t \in I$, $\tau \leq t$

$$x(t) = \xi + \int_\tau^t Ax(\sigma)d\sigma \tag{7}$$

We calculate the $\mathbb{R}^n$ norm of the left and right sides in the integral equation (7) and use the triangle inequality:

$$\|x(t)\| \leq \|\xi\| + \left\| \int_\tau^t Ax(\sigma)d\sigma \right\|$$

The triangle inequality for integrals:

$$\left\| \int_\tau^t x(\sigma) d\sigma \right\| \leq \int_\tau^t \|x(\sigma)\| \, d\sigma$$

and the definition of the matrix norm:

$$\|A\| \stackrel{\text{def}}{=} \sup_{\|x\| \neq 0} \left(\|Ax\| \, / \, \|x\| \right) = \sup_{\|x\|=1} \left(\|Ax\| \right)$$

imply that

$$\|x(t)\| \leq \|\xi\| + \int_\tau^t \|Ax(\sigma)\| \, d\sigma$$

and finally

$$\|x(t)\| \leq \|\xi\| + \int_\tau^t \|A\| \, \|x(\sigma)\| \, d\sigma$$

We will prove now that this integral inequality for $\|x(t)\|$ implies the famous **Grönwall inequality** for such kind of integral inequalities, giving an estimate for $\|x(t)\|$ in terms of the initial data $\|\xi\|$.

This is a standard argument that will be used within the course again later two more times for more complicated types of equations.

Introducing the notation $G(t) = \|\xi\| + \int_\tau^t \|A\| \, \|x(\sigma)\| \, d\sigma$ for te right hand side in the inequality, we conclude that $G(\tau) = \xi$, $\|x(t)\| \leq G(t)$, and

$$G'(t) = \|A\| \, \|x(t)\| \leq \|A\| \, G(t)$$

Multiplying the last inequality by the integrating factor $\exp(- \|A\| \, t)$ and referring to the chain rool for the derivative of $\exp(- \|A\| \, t)$, we arrive to

$$\begin{aligned}
G'(t) \exp(- \|A\| \, t) &\leq \|A\| \exp(- \|A\| \, t) G(t) \\
G'(t) \exp(- \|A\| \, t) - \|A\| \exp(- \|A\| \, t) G(t) &\leq 0 \\
G'(t) \exp(- \|A\| \, t) + G(t) \left(\exp(- \|A\| \, t) \right)' &\leq 0 \\
\left(G(t) \exp(- \|A\| \, t) \right)' &\leq 0
\end{aligned}$$

Integrating the left and the right hand side from τ to t we get the inequality

$$\begin{aligned}
G(t) \exp(- \|A\| \, (t)) - G(\tau) \exp(- \|A\| \, (\tau)) &\leq 0 \\
G(t) \exp(- \|A\| \, (t)) &\leq G(\tau) \exp(- \|A\| \, (\tau)) \\
G(t) &\leq \|\xi\| \exp(\|A\| \, (t - \tau))
\end{aligned}$$

The last relation implies the **Grönwall inequality** in this simple case:

$$\|x(t)\| \leq \|\xi\| \exp(\|A\| \, (t - \tau)) \tag{8}$$

that follows from the integral inequality:

$$\|x(t)\| \leq \|\xi\| + \int_{\tau}^{t} \|A\| \, \|x(\sigma)\| \, d\sigma$$

■**(Knowlege of this proof is required at the exam)**

Lemma. The solution to I.V.P. (4),(5) is unique.

$$x' = Ax, \quad x(\tau) = \xi$$

Proof. Suppose that there are two solutions $x(t)$ and $y(t)$ to the I.V.P. (4),(5) on a time interval including τ and both are equal to ξ at the initial time $t = \tau$. Consider the vector valued function $z(t) = x(t) - y(t)$ and the case when $\tau \leq t$. Then $z(t)$ is also a solution to the same equation (4) and satisfies the initial condition $z(\tau) = 0$.

The estimate (8) applied to $z(t)$ implies that

$$\|z(t)\| \leq 0 \exp(\|A\| \, (t - \tau)) = 0$$

$z(t) = 0$ and therefore the uniqueness of solution to I.V.P. (4),(5). The proof of the case $\tau \leq t$ is similar.■

4.3 Exponent of a matrix

Two ideas are used to construct analytical solutions to (4) :

1) One is to find a possibly simple basis $\{v_1(t), ..., v_N(t)\}$ to the solutions space $\mathcal{S}_{\text{hom}}$.

2) Another one is based on the observation that the matrix exponent

$$\exp\left(A\left(t-\tau\right)\right) = e^{A(t-\tau)} \overset{def}{=} I + A\left(t-\tau\right) + \frac{1}{2}A^2\left(t-\tau\right)^2 + ... + \frac{1}{k!}A^k\left(t-\tau\right)^k ... = \sum_{k=0}^{\infty}\frac{1}{k!}A^k\left(t-\tau\right)^k$$

gives an expression of the the unique solution to the I.V.P. (1), (1a) in the form:

$$x(t) = e^{A(t-\tau)}\xi$$

One can derive this property of the matrix exponent by considering the integral equation (7) for $x(t)$

$$x(t) = \xi + \int_{\tau}^{t} Ax(\sigma)d\sigma$$

equivalent to the I.V.P. (4),(5). We can try to solve this integral equation by iterations:

$$\begin{aligned}
x_{k+1}(t) &= \xi + \int_{\tau}^{t} Ax_k(\sigma)d\sigma && (9)\\
x_0 &= \xi\\
x_1(t) &= \xi + \int_{\tau}^{t} A\xi d\sigma = \xi + A\xi\int_{\tau}^{t} d\sigma = \left[I + A\left(t-\tau\right)\right]\xi && (10)\\
x_2(t) &= \xi + \int_{\tau}^{t} A\left[I + A\left(t-\tau\right)\right]\xi d\sigma = \left[I + A\left(t-\tau\right) + \frac{1}{2}A^2\left(t-\tau\right)^2\right]\xi && (11)\\
&\dots && (12)\\
x_k(t) &= \left[I + A\left(t-\tau\right) + \frac{1}{2}A^2\left(t-\tau\right)^2 + ... + \frac{1}{k!}A^k\left(t-\tau\right)^k\right]\xi
\end{aligned}$$

Iterations $x_k(t)$ converge uniformly on any finite time interval as $k \to \infty$ and the limit gives the series for $\exp(At)$ formulated above, times the initial data ξ.

The series for $\exp(At) = \sum_{k=0}^{\infty}\frac{1}{k!}A^k\left(t-\tau\right)^k$ converges uniformly on any finite time interval $[-T, T]$ including initial time point $\tau \in [-T, T]$ by the Weierstrass criterion. Most of you studied it before. We will remind it's formulation here. It will be used several times in the course.

Weierstrass criterion. Corollary A.23, p. 277 in L.R.

Let X be a normed vector space, Y be a complete normed vector space (Banach space) $K \subset X$ be compact,$\{f_n(x)\}_{n=1}^{\infty}$, $x \in K$ be a sequence of continuous functions $f_n : K \to Y$ and let $\{m_n\}_{n=1}^{\infty}$ a real sequence such that $\|f_n(x)\| \leq m_n$ for all $x \in K$ and all $n \in \mathbb{N}$, where $\|...\|$ is the norm in Y.

If $\sum_{n=1}^{\infty} m_n$ is convergent, then $\sum_{n=1}^{\infty} f_n(x)$ is unifomly convergent on K.$\square$

You studied this theorem in the case when $X = \mathbb{R}^N$, $Y = \mathbb{R}^M$. In our situation here K is a closed interval in $\mathbb{R}$ for example $[-T, T]$ in $\mathbb{R}$ and Y is a space of matrices $\mathbb{R}^{N\times N}$ (or $\mathbb{C}^{N\times N}$).

To prove that our series satisfies the Weierstrass criterion, we will apply the estimate for the norm of the product of two matrices:

$$\|AB\| \leq \|A\|$$

It implies that $\|A^2\| \leq \|A\| \, \|A\|$, $\|A^3\| \leq \|A\| \, \|A\| \, \|A\|$, ...and $\left\|A^k\right\| \leq \|A\|^k$ et.c.

Home exercise. Prove the inequality $\|AB\| \leq \|A\| \, \|B\|$ yourself!

Therefore the norm of each term in the series $\sum_{k=0}^{\infty} \frac{1}{k!} A^k (t-\tau)^k$ is estimated by a term from a convergent number series:

$$\begin{aligned}
\left\| \frac{1}{k!} A^k (t-\tau)^k \right\| &\leq \frac{1}{k!} \left\|A^k\right\| |t-\tau|^k \leq \\
\frac{1}{k!} \|A\|^k |t-\tau|^k &\leq \frac{1}{k!} \|A\|^k (2T)^k
\end{aligned}$$

for the exponential function $\exp(\|A\| (2T))$. We use here that $|t-\tau| \leq 2T$ for each $t \in [-T, T]$.

Application of the Weierstrass criterion to the series $\sum_{k=0}^{\infty} \frac{1}{k!} A^k (t-\tau)^k$ leads to the solution of the I.V.P. in the form

$$x(t) = e^{A(t-\tau)}\xi = \exp(A (t-\tau))\xi = \left(\sum_{k=0}^{\infty} \frac{1}{k!} (t-\tau)^k A^k \right) \xi$$

We make this conclusion by tending to the limit $k \to \infty$ in the integral equation (9) defining iterations. The expression under the integral in (9) converges uniformly and therefore the limit of the integral is a continuous function equal to the integral of this uniform limit. This solution is unique by the Lemma we proved before.

The last formula together with the Lemma about uniqueness of solutions imply the following Corollary about existence and uniquness of solutions to linear autonomous ODEs.

Corollary 2.9 in L.&R.

The function $x(t) = \exp(A(t-\tau))\xi$ is the unique solution to the I.V.P. (4),(5).

This theoretical expression for unique solutions to I.V.P. dispite of it's elegansce has a huge disadvantage that the series $\exp(At) = \sum_{k=0}^{\infty} \frac{1}{k!} (t-\tau)^k A^k$ is not possible to calculate analytically in a simple way.

We will try instead to find a basis of the vector space $\mathcal{S}_{\text{hom}}$ of all solutions to (1).

4.4 The dimension of the space $\mathcal{S}_{\text{hom}}$ of solutions to an autonomous linear system of ODEs

Theorem. (Proposition 2.7, p.30, L.R. in the case of non-autonomous systems).

Let $b_1, ..., b_N$ be a basis in $\mathbb{R}^N$(or $\mathbb{C}^N$). Then the functions $y_j : \mathbb{R} \to \mathbb{R}^N$(or $\mathbb{C}^N$) defined as solutions to the I.V.P. (4),(5)

$$x'(t) = Ax(t), \quad A \in \mathbb{R}^{N \times N}(\mathbb{C}^{N \times N})$$

with $y_j(\tau) = b_j$, $j = 1, ...N$, by $y_j(t) = \exp(A(t-\tau))b_j$, form a basis for the space $\mathcal{S}_{\text{hom}}$ of solutions to (4). The dimension of the vector space $\mathcal{S}_{\text{hom}}$ of solutions to (4) is equal to N - the dimension of the system (4).

The idea of the proof. This property is a consequence of the linearity of the system and the uniqueness of solutions to the system and is independent of detailed properties of the matrices $A(t)$ and A in (4) and (6).

Proof. Consider a linear combination of $y_j(t)$ equal to zero for some time $\sigma \in \mathbb{R}$: $l(\sigma) = \sum_{j=1}^{N} \alpha_j y_j(\sigma) = 0$. Observe that the trivial constant zero solution $0(t)$ coinsides with l at this time point.

But by the uniqueness of solutions to (4) it implies that $l(t)$ at arbitrary time must coinside with the trivial zero solution for all times t and in particular at time $t = \tau$. Therefore $l(\tau) = \sum_{j=1}^{N} \alpha_j b_j = 0$ (point out that $y_j(\tau) = b_j$). It implies that all coefficients $\alpha_j = 0$ because $b_1, ..., b_N$ are linearly independent vectors in $\mathbb{R}^N$(or $\mathbb{C}^N$). It implies that $y_1(t), ..., y_N(t)$ are linearly independent for all $t \in \mathbb{R}$ by definition. Arbitrary initial data $x(\tau) = \xi$ in $\mathbb{R}^N$(or $\mathbb{C}^N$) can be represented as a linear combination of basis vectors $b_1, ..., b_N$: $\xi = \sum_{j=1}^{N} C_j b_j$. The construction of $y_1(t), ..., y_N(t)$ shows that an arbitrary solution to (4) can be represented as linear combination of $y_1(t), ..., y_N(t)$.

$$\begin{aligned} x(t) &= \exp(A(t-\tau))\xi = \exp(A(t-\tau)) \sum_{j=1}^{N} C_j b_j = \\ &= \sum_{j=1}^{N} C_j [\underbrace{\exp(A(t-\tau)) b_j}_{=y_j(t)}] = \sum_{j=1}^{N} C_j y_j(t) \end{aligned}$$

Therefore $\{y_1(t), ..., y_N(t)\}$ is the basis in the space of solutions $\mathcal{S}_{\text{hom}}$ and therefore $\mathcal{S}_{\text{hom}}$ has dimension N.$\blacksquare$

(Knowlege of this proof is required at the exam)

By taking $\xi = e_1, ..., e_N$ we observe that each column in the matrix $\exp(A\,(t-\tau))$ is a solution to the equation (4). We have just shown in the theorem before that these columns are linearly independent and build a basis in the space of solutions.

Lecture 2
Summary of the material introduced in Lecture 1

1. Initial value problem (I.V.P.) for an ordinary differential equation (ODE).

$$x'(t) = f(t, x(t)),\ \ x(\tau) = \xi$$

2. Types of differential equations: autonomous, non-autonomous, linear, linear non-homogeneous, non-linear.

3. Questions of interest about ODEs.

a) Existence and uniqueness of solutions. Examples.

b) General solutions (an analytic expression for all solutions)

c) Finding specific solutions: equilibrium: $f(x_0) = 0$, periodic.

d) Stability of equilibrium points (do all solutions stay close to an equilibrium point?)

e) Attractors of solutions (sets that solutions tend to when $t \to \infty$,can be equilibrium points or periodic orbits)

4. Autonomous linear ODEs ($x' = Ax$ with constant matrix A)

5. Exponent of a matrix as a tool for calculating solutions to I.V.P.

6. Uniqueness of solutions and the proof with Grönwalls inequality. (**required at the exam**)

7. The space of solutions to $x' = Ax$, it's dimension and a construction of a basis.(**required at the exam**)

4.5 Properties of the matrix exponent.

We collect in the following Lemma some (may be partially known) properties of the matrix exponent.

For a complex matrix M the notation M^* means transpose and complex conjugate matrix (called also Hermitian transpose)

Lemma (Lemma 2.10 , p. 34 in L.&R.) Let P and Q be matrices in $\mathbb{R}^{N\times N}$ or $\mathbb{C}^{N\times N}$

(1) For a diagonal matrix $P = diag(\lambda_1, \lambda_2, ..., \lambda_n)$

$$\exp(P) = diag(\exp(\lambda_1), ..., \exp(\lambda_n))$$

(2) $\exp(P^*) = (\exp(P))^*$

(3) for all $t \in \mathbb{R}$,

$$\frac{d}{dt}\exp(At) = A\exp(At) = \exp(At)A$$

(4) If P and Q are two **commuting matrices** $PQ = QP$, then $\exp(P)Q = Q\exp(P)$ and

$$\exp(P + Q) = \exp(P)\exp(Q) = \exp(Q)\exp(P)$$

(5) $\exp(-P)\exp(P) = \exp(P)\exp(-P) = I$ or $\exp(-P) = (\exp(P))^{-1}$

Proof

Proofs of (1),(2) are left as exercises. We proof first (4) by direct calculation.

$$
\begin{aligned}
(P+Q)^k &= \sum_{m=0}^{k} \binom{k}{m} P^m Q^{k-m} \quad (for \quad commuting \quad matrices) \\
& \binom{k}{m} \overset{\text{notation}}{=} \left(\frac{k!}{m!\,(k-m)!}\right) \\
e^{P+Q} &= \sum_{k=0}^{\infty} \frac{1}{k!} (P+Q)^k = \\
&= \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{m=0}^{k} \left(\frac{k!}{m!\,(k-m)!}\right) P^m Q^{k-m} == \sum_{k=0}^{\infty} \sum_{m=0}^{k} \left(\frac{1}{m!\,(k-m)!}\right) P^m Q^{k-m} = \\
&= \sum_{m=0}^{\infty} \sum_{\substack{p=0 \\ (k=p+m)}}^{\infty} \frac{P^m}{m!} \frac{Q^p}{p!} = \\
&= \left(\sum_{m=0}^{\infty} \frac{1}{m!} P^m\right) \left(\sum_{p=0}^{\infty} \frac{1}{p!} Q^p\right) = e^P e^Q
\end{aligned}
$$

(3) Can be proved in three different ways.

It follows from the definition of $\exp(At)$ by elementwise differentiation of the corresponding uniformly converging series.

It follows also from the observation above that each column in $\exp(At)$ with index k is a solution to the system of equations $x' = Ax$ with initial data $x(0) = e_k$.

A straightforward proof can be given by the definition of derivative and using the relation (4). We use the formula $\exp(P+Q) = \exp(P)\exp(Q)$ for commuting matrices, the fact that At and As commute for any t and s and the Taylor formula applied to for $\exp(Ah) - I$ for small h:

$$
\begin{aligned}
\exp(A\,(t+h)) - \exp(At) &= (\exp(Ah) - I)\exp(At) = \\
& \left(Ah + O(h^2)\right)\exp(At)
\end{aligned}
$$

Therefore

$$
\begin{aligned}
\frac{d}{dt}(\exp(At)) &= \lim_{h\to 0} \frac{(\exp(A\,(t+h)) - \exp(At))}{h} = \\
\lim_{h\to 0} \frac{(Ah + O(h^2))\exp(At)}{h} &= A\exp(At)
\end{aligned}
$$

■

4.6 Analytic solutions. Case when a basis of eigenvectors exists.

We know that the unique solution to the initial value problem (I.V.P.) $x'(t) = A\,x(t)$, $x(\tau) = \xi$ can be expressed as

$$x(t) = \exp(A(t-\tau))\xi$$

or $x(t) = \exp(At)\xi$ in the case when the initial time $\tau = 0$. But this beautiful expression does not give an explicit formula for the solution because the matrix exponent $\exp(At)$ is defined as an infinite series.

An idea that leads to an explicit analytical solution is to use the theorem about the basis in the space of solutions. We can try to find a basis $\{y_1(t), ..., y_N(t)\}$ to the solution space $\mathcal{S}_{\text{hom}}$ by finding a particular basis $\{v_1, ..., v_N\}$ in $\mathbb{C}^N$ or $\mathbb{R}^N$ such that the matrix exponent $\exp(At)$ acts on the elements of this basis in a particularly simple way, so that all solutions $y_k(t) = \exp(A\,(t-\tau))v_k$ can be calculated explicitely. We will consider mainly the case $\tau = 0$ for autonomous systems.

The simplest example that illustrates this idea is given by eigenvectors to A. These are vectors $v \neq 0$ such that

$$Av = \lambda v$$

for some number λ. Numbers λ are called eigenvalues of A. Eigenvalues must be roots of the characteristic polynomial

$$p(\lambda) = \det(A - \lambda I)$$

because rewriting the definition of an eigenvector we arrive to a homogeneous system of linear equations with matrix $(A - \lambda I)$

$$(A - \lambda I)\,v = 0$$

with $v \neq 0$. Using the definition $Av = \lambda v$ for the eigenvalue and the eigenvector k times we conclude that $A^k v = \lambda^k v$. Substituting this formula into the expression $e^{At}v = \sum_{k=0}^{\infty} \frac{1}{k!} t^k A^k v$ we conclude that (!!!)

$$e^{At}v = \sum_{k=0}^{\infty} \frac{1}{k!} t^k \lambda^k v = e^{\lambda t} v.$$

Important new idea.

Another more general idea leads to the same formula, but has an advantage that it can be applied in more complicated situations. We use here that the eigenvector v corresponding to the eigenvalue λ makes all powers $(A - \lambda I)^k\, v = 0$ except $k = 0$:

$$\begin{aligned} e^{At}v &= \exp\left(\lambda t I + (At - \lambda t I)\right) v = \exp\left(\lambda t I\right) \exp\left(\left(A - \lambda I\right) t\right) v = \\ &= \left(e^{\lambda t} I\right) \sum_{k=0}^{\infty} \frac{1}{k!} t^k \left(A - \lambda I\right)^k v = e^{\lambda t} v. \end{aligned} \tag{13}$$

This observation leads to a simple conclusion that if the matrix A has N linearly independent eigenvectors $\{v_k\}_{k=1}^N$, then any solution to (4) with initial data $\xi = \sum_{k=1}^N C_k v_k$ can be expressed as a linear combination in the form

$$x(t) = \sum_{k=1}^{N} C_k \left(e^{\lambda_k t} v_k\right)$$

with vector functions $y_k(t) = e^{\lambda_k t} v_k$, $k = 1, ..., N$, building a basis for the space of solutions to the equation $x' = Ax$ (4).

It is the case when a diagonalization of the matrix A exists.

Check the diagonalization property and it's use for solution of ODEs in the course of linear algebra!

We point out that λ and v can be a complex eigenvalue and a complex eigenvector here. In the case when all these eigenhvalues are real, this basis will be real. In the case if a real matrix A has some complex eigenvalues, they appear as pairs of complex conjugate eigenvalues and corresponding eigenvectors, that still can be used to build a real basis for solutions. We will demonstrate it on a couple of examples later.

Example 1. Consider system $x' = Ax$ with matrix $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. The matrix A has characteristic polynomial $p(\lambda) = \lambda^2 - 1$ and two eigenvalues $\lambda_1 = 1$ and $\lambda_2 = -1$.

Corresponding eigenvectors satisfy homogeneous systems $(A - \lambda_1)\, v_1 = 0$ with matrix $(A - \lambda_1 I) = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$ and $(A - \lambda_1 I)\, v_2 = 0$ with matrix $(A - \lambda_2 I) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$.

Eigenvectors are $v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $v_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and are linearly independent (in particular it follows from the fact that eigenvalues are different). Solutions $y_1(t) = e^t v_1$ and $y_2(t) = e^{-t} v_2$ are linearly independent.

Arbitrary real solution to the system of ODEs has the form

$$x(t) = C_1 y_1(t) + C_2 y_2(t) = C_1 e^t \begin{bmatrix} 1 \\ 1 \end{bmatrix} + C_2 e^{-t} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

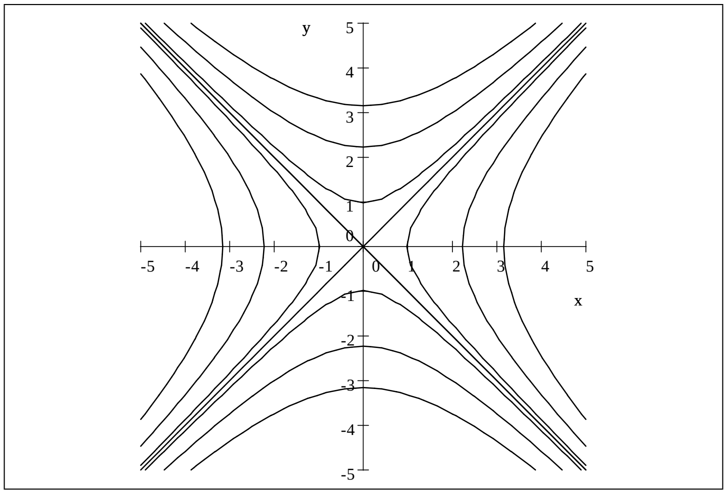
with arbitrary coefficients C_1 and C_2.Corresponding phase portrait includes a particular solutions tending to infinity along the vector $v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, a solutions tending to the origin along the vector $v_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and other solutions filling the rest of the plain having orbits in the form of hyperbolas. One can observe it by integrating the differential equation

$$\begin{aligned} x_1' &= x_2 \\ x_2' &= x_1 \end{aligned}$$

$$\frac{dx_2}{dx_1} = \frac{x_1}{x_2}; \quad x_2 dx_2 = x_1 dx_1$$

with separable variables that follows from the system and concluding that

$$x_1^2 - x_2^2 = Const$$



Similar phase portraits will be observed in the arbitrary case when the 2×2 real non-degenerate matrix A ($\det A \neq 0$) has real eigenvalues with different signs but the picture will be rotated and might be less symmetric depending on the directions of the eigenvectors v_1 and v_2 (here they are orthogonal). One can still draw trajectories along eigenvectors and then sketch other trajectories according to the directions of trajectories along eigenvectors.

5 Generalised eigenvectors and eigenspaces.

It is easy to give examples of matrices that cannot be diagonalized. For linear autonomous systems with such matrices the expression of arbitrary solutions in terms of linearly independent eigenvectors is impossible because we just do not have N linearly independent ones.

Example 3.

$$\left\{ \begin{array}{c} x_1' = -x_1 \\ x_2' = x_1 - x_2 \end{array} \right. \quad \text{or} \quad x'(t) = Ax \text{ with } A = \begin{bmatrix} -1 & 0 \\ 1 & -1 \end{bmatrix}, \text{ the characteristic polynomial is } p(\lambda) = (\lambda+1)^2.$$

$$p(\lambda) = \lambda^2 - \lambda Tr\,(A) + \det(A) \text{ in dimension 2}$$

Matrix A has an eigenvalue $\lambda = -1$ with algebraic multiplicity $m(\lambda) = 2$. There is only one linearly independent eigenvector $v = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ satisfying the equation $(A - \lambda I)v = 0$.

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

The function $x(t) = e^{-t}v$ is a solution to the system. One likes to find a basis of solutions to the space $\mathcal{S}_{\text{hom}}$ of all solutions. We need another linearly independent solution for that. Observe that

$$x_1(t) = C_1 e^{-t}$$

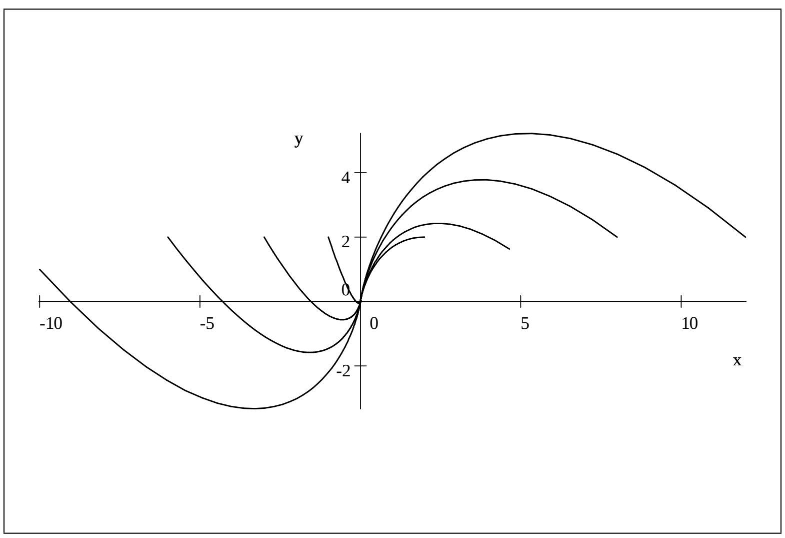
is the solution to the first equation, substitute it into the second equation and solve it explicitely with respect to $x_2(t)$:

$$\begin{aligned}
x_2'(t) &= -x_2(t) + C_1e^{-t} \\
e^t x_2'(t) + e^t x_2(t) &= C_1 \\
\left(e^t x_2(t)\right)' &= C_1 \\
e^t x_2(t) &= C_2 + C_1 t \\
x_2 &= C_2e^{-t} + C_1 t\, e^{-t}
\end{aligned}$$

Therefore the general solution to this particular system has the form

$$\begin{aligned}
x(t) &= \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = C_1e^{-t}\begin{bmatrix} 1 \\ t \end{bmatrix} + C_2e^{-t}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \\
& C_1e^{-t}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) + C_2e^{-t}\begin{bmatrix} 0 \\ 1 \end{bmatrix} \\
&= C_1e^{-t}\left(v^{(1)} + tv\right) + C_2e^{-t}v
\end{aligned}$$

where $v^{(1)} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. The phase portrait looks as:



In this particular example we could find an explicit solution using the fact that the matrix A is trianglular. This idea cannot be generalized to the arbitrary case but can be used for linear system with variable coefficients and triangular matrix.

We point out that the initial value for the derived solution

$$x(t) = C_1e^{-t}\left(v^{(1)} + tv\right) + C_2e^{-t}v \text{ is } x(0) = \xi = C_1\begin{bmatrix} 1 \\ 0 \end{bmatrix} + C_2\begin{bmatrix} 0 \\ 1 \end{bmatrix} = C_1v^{(1)} + C_2v.$$

Vector $v^{(1)} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is linearly independent of the eigenvector $v = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, and applying the lemma about the basis of the solution space, we conclude that $e^{-t}v$ and $e^{-t}\left(v^{(1)}+tv\right)$ are linearly independent for all $t \in R^N$ and build a basis for the space of solutions to the system.

Observe that $v^{(1)}$ has a remarkable property that $(A-\lambda I)\,v^{(1)} = v$ as $\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and therefore $(A-\lambda I)^2\,v^{(1)} = (A-\lambda I)\,v \overset{def}{=} 0$. Such vectors are called **generalised eigenvectors** to A corresponding to the eigenvalue λ.

We point out that the initial data in this explicit solution are represented as a linear combination of an eigenvector and a generalised eigenvector: $x(0)=C_1v^{(1)} + C_2 v$.

A more general idea!!!

We observe also that the general solution we have got could be derived by applying the same idea as in the formula $e^{At} = \exp(\lambda t I)\exp((A-\lambda I)\,t)$: (13) before, but applied to the generalised eigenvector $v^{(1)}$:

$$\begin{aligned}
\exp(At)v^{(1)} &= \exp\left(\lambda t I + (At - \lambda t I)\right)v^{(1)} = \exp(\lambda t I)\exp\left((A-\lambda I)\,t\right)v^{(1)} \\
e^{\lambda t}\sum_{k=0}^{\infty}\frac{1}{k!}t^k (A-\lambda I)^k v^{(1)} &= e^{\lambda t}\left(v^{(1)} + t\,(A-\lambda I)\,v^{(1)}\right) = \\
&= e^{\lambda t}\left(v^{(1)} + tv\right) \\
(A-\lambda I)^k v^{(1)} &= 0, \quad k \geq 2
\end{aligned}$$

This reasoning again gives the second basis vector in the space of solutions, that we have got before by the trick with separation of variables, and gives a clue what might be a general way to explicit solution to the linear system of ODEs $x' = Ax$ with an arbitrary constant matrix A.■

Definition of generalised eigenvectors.

A non-zero vector $z \in \mathbb{C}^N$(or $\mathbb{R}^N$) is called a generalised eigenvector to the matrix $A \in \mathbb{C}^{N\times N}$ corresponding to the eigenvalue λ with the algebraic multiplicity $m(\lambda)$ if $(A-\lambda I)^{m(\lambda)}\,z = 0$.□

If $(A-\lambda I)^r\,z = 0$ and $(A-\lambda I)^{r-1}\,z \neq 0$ for some $0 < r < m(\lambda)$ we say that z is a generalised eigenvector of rank (or height) r to the matrix A.□

An eigenvector u is a generalised eigenvector of rank 1 because $(A-\lambda I)\,u = 0$.

Notation.

The set $\ker\left((A-\lambda I)^{m(\lambda)}\right)$ - (kernel or nullspace) of all generalized eigenvectors of an eigenvalue λ is denoted by $E(\lambda)$ in the course book. $E(\lambda)$ is a subspace in $\mathbb{C}^N$(*or* $\mathbb{R}^N$).

Proposition on A - invariance of $E(\lambda)$.

$E(\lambda)$ is A- invariant, namely if $z \in E(\lambda)$, then $Az \in E(\lambda)$.

Proof. We check it by taking $z \in E(\lambda)$ such that $(A-\lambda I)^{m(\lambda)}\,z = 0$ and calculating $(A-\lambda I)^{m(\lambda)}\,(Az) = A\left((A-\lambda I)^{m(\lambda)}\,z\right) = 0$, the last equality is valid because A and $(A-\lambda I)^{m(\lambda)}$ commute.■

Proposition on $\exp(At)$ - invariance of $E(\lambda)$.

$E(\lambda)$ is invariant under the action of $\exp(At)$, namely if $z \in E(\lambda)$, then $\exp(At)z \in E(\lambda)$.

Proof. Consider the expression for the $\exp(At)z$ as a series

$$\exp(At)z = \sum_{k=0}^{\infty} \frac{1}{k!} t^k A^k z = \lim_{m\to\infty} \left. \sum_{k=0}^{m} \underbrace{\frac{1}{k!} t^k A^k z}_{\in E(\lambda)} \right\} \in E(\lambda)$$

All terms $A^k z$ in the sum belong to $E(\lambda)$. One can see it by repeating the argument in the previous proposition.

The expression for $\exp(At)z$ is therefore a limit of linear combinations of elements from the finite dimensional generalized eigenspace $E(\lambda)$ that is a closed and complete set. Therefore $\exp(At)z$ must belong to $E(\lambda)$.■

A remarkable property of generalised eigenvectors z is that the series for the matrix exponent $\exp(At)$ applied to z namely $\exp(At)z$ can be expressed in such a way that it would include only a finite number of terms and can be calculated analytically.

Notation.

$\sigma(A)$ is the set of all eigenvalues of the matrix A, or spectrum of the matrix A.

Theorem (2.11, Part 1), p. 35 in the course book) Let $A \in \mathbb{C}^{N\times N}$. For an eigenvalue $\lambda \in \sigma(A)$ with algebraic multiplicity $m(\lambda)$ denote the subspace of its associated generalised eigenvectors by $E(\lambda) = \ker(A - \lambda I)^{m(\lambda)}$ and for $z \in \mathbb{C}^N$denote by $x_z(t) = \exp(At)z$ - the solution of I.V.P. with initial data $x_z(0) = z$. Then for any $\lambda \in \sigma(A)$ and any $z \in E(\lambda)$ a generalised eigenvector

$$\exp(At)\, z = e^{\lambda t} \sum_{k=0}^{m(\lambda)-1} \frac{t^k}{k!} (A - \lambda I)^k z$$

Proof.

We show it by the following direct calculation:

$$\begin{aligned} x_z(t) &= \exp(At)\, z = \exp(t\lambda I) \exp((A - \lambda I)\, t)\; z = \qquad (14) \\ \left(e^{\lambda t} I\right) \sum_{k=0}^{\infty} \frac{t^k}{k!} (A - \lambda I)^k z &= e^{\lambda t} \sum_{k=0}^{m(\lambda)-1} \frac{t^k}{k!} (A - \lambda I)^k z \end{aligned}$$

because powers $(A - \lambda I)^k z = 0$ - terminate on $z \in E(\lambda)$ for all $k \geq m(\lambda)$ by the definition of generalised eigenvectors.

We also use at the first step of calculations the property (4) from the Lemma about matrix exponents: $\exp(P + Q) = \exp(P) \exp(Q)$ for commuting matrices P and Q. ■

5.1 Analytic solutions. General case using a basis of generalized eigenvectors.

The next theorem gives a theoretical background for a method of constructing analytic solutions to (4) ($x'(t) = Ax(t)$), by representing arbitrary initial data $x(0) = \xi$ using a basis of generalised eigenvectors to

A in $\mathbb{C}^N$. We are going to consider initial conditions for autonomous systems only at the point $\tau = 0$, because all other solutions are derived from such ones just by a shift in time, because the right hand side in the equation does not depend on time explicitely and if $x(t)$ is a solution, then $x(t+\tau)$ is also a solution.

Definition The sum $V_1 + V_2 + ... + V_s$ of subspaces V_1,V_2 ...V_s in a vector space is a set of vectors in the form $v_1 + v_2 + ... + v_s$ with vectors $v_j \in V_j$, $j = 1, ..., s$. $\square$

Definition Direct sum $V_1 \oplus V_2 \oplus ... \oplus V_s$ of subspaces V_1,V_2 ...V_s is a usual sum $V_1 + V_2 + ... + V_s$ of these subspaces with a *special additional property* that any vector in $v \in V_1 \oplus V_2 \oplus ... \oplus V_s$ is represented only in a **unique way** as a sum $v = v_1 + v_2 + ... + v_s$ of vectors $v_j \in V_j$, $j = 1, ..., s$. $\square$

It makes in this case any set of vectors $v_j \in V_j$, $j = 1, ..., s$ belonging to different V_j linearly independent.

Subspaces V_j, $j = 1, ..., s$ have only one common point - zero.

Theorem (generalized eigenspace decomposition theorem A.8, p. 268 in the course book, without proof)

Let $A \in \mathbb{C}^{N\times N}$ and $\lambda_1, ..., \lambda_s$ be all distinct eigenvalues of A with multiplicities m_j, $\sum_{j=1}^{s} m_j = N$. Then $\mathbb{C}^N$ can be represented as a direct sum of generalised eigenspaces $E(\lambda_j) = \ker(A - \lambda_j I)^{m_j}$ to A having dimensions m_j:

$$\dim\left(\ker(A - \lambda_j I)^{m_j}\right) = m_j$$

$$\mathbb{C}^N = \ker(A - \lambda_1 I)^{m_1} \oplus ... \oplus \ker(A - \lambda_s I)^{m_s} \tag{15}$$

$\square$

If the matrix A is real having real eigenvalues, then the result will be valid for $\mathbb{R}^N$.

The formula (14) together with the decomposition of $\mathbb{C}^N$ into direct the sum of generalised eigenspaces gives a recipe for a finite analytic representation of solutions to I.V.P. to $x' = Ax$ (4) and a representation of te general solution to (4).

Theorem (2.11, part 2, p. 35 in the course book) *Let $z \in E(\lambda)$ be a generalized eigenvector corresponding to the eigenvalue λ. Denote by $x_z(t) = \exp(At)z$ - the solution of I.V.P. with $x_z(0) = z$.*

Let $B(\lambda_j)$ be a basis in $E(\lambda_j)$ having dimesion m_j, and denote $\mathcal{B} = \cup_{j=1}^{s} B(\lambda_j)$ - the union of all bases of generalized eigenspaces $E(\lambda_j)$ for all distinct eigenvalues $\lambda_j \in \sigma(A)$.

The set of functions $\{x_z : z \in \mathcal{B}\}$ is a basis of the solution space S_{hom} of (4).

Proof. By the generalized eigenspace decomposition theorem $\mathbb{C}^N = \ker(A-\lambda_1)^{m_1} \oplus ... \oplus \ker(A-\lambda_s)^{m_s}$ and therefore all subspaces $E(\lambda_j) = \ker(A - \lambda_j)^{m_j}$ making them linearly independent. The total number of these basis vectors is $\sum_{j=1}^{s} m_j = N$ that is equal to the dimension of $\mathbb{C}^N$. Therefore $\mathcal{B}$ is a basis in $\mathbb{C}^N$.

From the theorem on the dimension of the solution space $\mathcal{S}_{\text{hom}}$ of a linear system it follows that solutions with initial data taken from the basis $\mathcal{B}$ build a basis in the solution space $\mathcal{S}_{\text{hom}}$ of (4).

$\blacksquare$

Lecture 3
Summary of Lecture 2.

1. Generalized eigenvector z to a matrix A corresponding to an eigenvalue λ is a vector satifying the relation

$$(A - \lambda I)^{m(\lambda)} z = 0$$

where $m(\lambda)$ is the algebraic multiplicity of λ: $p(w) = (w - \lambda)^{m(\lambda)}(...)$.

2. Generalised eigenspace $E(\lambda)$ is a subspace of all gen. eigenvectors corresponding to λ.

3. A remarcable observation that series for $\exp(At)z$ include only finite number of terms up to te order $m(\lambda) - 1$ in t.

4. There is a basis of generalized eigenvectors in $\mathbb{C}^N(\mathbb{R}^N)$

5. We will find general solution to $x' = Ax$ with help of such a basis.

Practical calculation of solutions to autonomous linear systems of ODEs

We continue with a description of how this theorem can be used for practical calculation of solutions to I.V.P.

Let the matrix A have s distinct eigenvalues $\lambda_1, ..., \lambda_s$ with corresponding generalised eigenspaces $E(\lambda_j)$. Represent the arbitrary initial data $x(0) = \xi$ for the solution $x(t)$ in a *unique way* as a sum of its components from different generalised eigenspaces:

$$\xi = \sum_{j=1}^{s} x^{0,j}, \quad x^{0,j} \in E(\lambda_j)$$

We remind here that $\mathbb{C}^N = E(\lambda_1) \oplus E(\lambda_2) \oplus ... \oplus E(\lambda_s)$ and it implies that any vector $\xi \in \mathbb{C}^N$ is represented i such a way in a unique way. Here $x^{0,j} \in E(\lambda_j)$ - are components of ξ in the generalized eigenspaces $E(\lambda_j)$ $= \ker(A - \lambda_j)^{m_j}$ of the matrix A. These subspaces intersect only in the origin and are invariant with respect to A and $\exp(At)$. It implies that for the solution $x_z(t)$ with initial data $z \in E(\lambda_j)$, we have $x_z(t) = \exp(At)z \in E(\lambda_j)$ for all $t \in \mathbb{R}$.

Let m_j be the algebraic multiplicity of the eigenvalue λ_j. We apply the formula (14) to this representation and derive the following expression for solutions for arbitrary initial data as a finite sum (instead of series):

$$x(t) = e^{At}\xi = e^{At}\left(\sum_{j=1}^{s} x^{0,j}\right) = \sum_{j=1}^{s} e^{At}x^{0,j} = \tag{16}$$

$$\sum_{j=1}^{s}\left(e^{\lambda_j t}\left[\sum_{k=0}^{m_j-1}(A - \lambda_j I)^k \frac{t^k}{k!}\right]x^{0,j}\right) \tag{17}$$

Series expressing $\exp(At)x^{0,j}$ terminates on each of the generalised eigenspaces $E(\lambda_j)$.

The last formula still needs specification to derive to an explicit solution. General solution can be written explicitely by finding a basis of of eigenvectors v_j and generalized eigenvectors for each generalised eigenspace $E(\lambda_j)$ and expressing all components $x^{0,j}$ of ξ in the generalized eigenspaces $E(\lambda_j)$ in the form

$$x^{0,j} = ...C_p v_j + C_{p+1}v_j^{(1)} + C_{p+2}v_j^{(2)}... \tag{18}$$

including all linearly independent eigenvectors v_j corresponding to λ_j *(it might exist several eigenvectors v_j corresponding to one λ_j*) and enough many linearly independent generalized eigenvectors $v_j^{(1)}$,..., $v_j^{(l)}$.

We consider first the equation for usual eigenvectors

$$(A - \lambda_j I)v = 0$$

solve it using Gauss elimination and find the number of free variables that gives the number of linearly independent eigenvectors. Then we look for linearly independent generalized eigenvectors to build up the whole basis for $E(\lambda_j)$.

We will start with examples illustrating this idea in some simple cases.

Example 4. Matrix 3x3 with two linearly independent eigenvectors.

Consider a system of equations $x' = Ax$ with matrix $A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ It is easy to see that $\lambda = 1$ is the only eigenvalue with algebraic multyplicity 3. The characteristic polynomial is $p(\lambda) = (1-\lambda)^3$.

The eigenvectors satisfy the equation $(A - I)\,v = 0$: $A - I = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. This equation insludes just one scalar equation for components of the vector $v = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$ just as: $y_2 + y_3 = 0$. It has two linearly independent solutions that can be chosen as $v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ and $v_2 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$. The eigenspace is a plane through the origin orthogonal to the vector $\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$

We like to find a generalised eigenvector linearly independent of v_1 and v_2. We take the eigenvector v_1 and solve the equation

$$(A - \lambda I)v_1^{(1)} = v_1.$$

because if it is valid, then

$$(A - \lambda I)^2 v_1^{(1)} = (A - \lambda I)(A - \lambda I)v_1^{(1)} = (A - \lambda I)v_1 = 0$$

We denote it by two indexes to point out that it belongs to a chain with base on v_1. Denoting $v_1^{(1)} = [y_1, y_2, y_3]^T$ we consider the system

$$\begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

It gives a solution $y_3 = 1,\ y_2 = 0,\ y_1 = 0.\ v_1^{(1)} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. We point out that if we try to find a chain of generalised eigenvectors starting from the eigenvector v_2, it leads to a system $(A - I)v_2^{(1)} = v_2$

$$\begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

that has no solutions (the second and the third equation is never satified).

If we try to extend the chain of generalised eigenvectors with one more: $v_1^{(2)}$by solving the system $(A - I)v_1^{(2)} = v_1^{(1)}$

$$\begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

we find that it has no solutions (in fact we know that there cannot be more linearly independent generalised eigenvectors because we have already found 3 of them).

Vectors v_1, v_2 and $v_1^{(1)}$ form a basis in $\mathbb{R}^3$.

We can write general solution to the system of ODE with matrix A using the general formula (16) and expressing the initial data as a linear combination of eigenvectros v_1 and v_2 and the generalised eigenvector $v_1^{(1)}$:

$$\begin{aligned} x(t) &= e^{\lambda t} \left[\sum_{k=0}^{2} (A - \lambda I)^k \frac{t^k}{k!} \right] \left(C_1 v_1 + C_2 v_2 + C_3 v_1^{(1)} \right) \\ \xi &= C_1 v_1 + C_2 v_2 + C_3 v_1^{(1)} \end{aligned}$$

$m(\lambda) = 3$. It is why we put upper bound in the sum equal to $m(\lambda) - 1 = 2$.

The expression above simplifies (using that by the construction $(A - \lambda I)v_1^{(1)} = v_1$ and therefore $(A - \lambda I)^2 v_1^{(1)} = (A - \lambda I)v_1 = 0$. to

$$\begin{aligned} x(t) &= \exp(At)\xi = C_1 e^t v_1 + C_2 e^t v_2 + C_3 e^t \left[I + (A - I)\, t \right] v_1^{(1)} \\ &= \quad C_1 e^t v_1 + C_2 e^t v_2 + C_3 e^t v_1^{(1)} + C_3 t e^t v_1 \end{aligned}$$

Example 5. Matrix 3x3 with one linearly independent eigenvector.

Consider a system of equations $x' = Ax$ with matrix $A = \begin{bmatrix} -1 & -1 & 0 \\ 0 & -1 & -2 \\ 0 & 0 & -1 \end{bmatrix}$ It is easy to see that $\lambda = -1$ is the only eigenvalue with multiplicity 3: $m(\lambda) = 3$.

Eigenvectors satisfy the equation

$$(A - \lambda I)\, v = 0$$

$A - \lambda I = A + I = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{bmatrix}$. It has one linearly independent solution that can be chosen as $v = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$.

We will build a **chain of generalised eigenvectors** starting with this eigenvector. Solve the equation $(A - \lambda I)^2 v^{(1)} = 0$ as before we take instead the equation $(A - \lambda I) v^{(1)} = v$ that would give us a solution to the first equation.

$$(A + I)v = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

The first equation in the system implies that $y_2 = -1$, and we are free to choose $y_1 = 0$ and $y_3 = 0$. $v^{(1)} = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}$.

The next generalised eigenvector $v^{(2)}$ such that $(A - \lambda I)^{m(\lambda)} v^{(2)} = (A - \lambda I)^3 v^{(2)} = 0$ in the chain must satisfy the equation

$$(A - \lambda I) v^{(2)} = v^{(1)}$$

$$(A + I)v^{(2)} = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}$$

$y_3 = 1/2$, $y_2 = 0$, $y_1 = 0$. $v^{(2)} = \begin{bmatrix} 0 \\ 0 \\ 1/2 \end{bmatrix}$.Express initial data ξ as $\xi = C_1 v + C_2 v^{(1)} + C_3 v^{(2)}$.

$$
\begin{aligned}
x(t) &= \exp(At)\xi = e^{\lambda t}\left[\sum_{k=0}^{2}(A-\lambda I)^k\frac{t^k}{k!}\right]\left(C_1 v + C_2 v^{(1)} + C_3 v^{(2)}\right) = \\
&\quad C_1 e^{\lambda t} v + C_2 e^{\lambda t} v^{(1)} + C_2 t e^{\lambda t}(A-\lambda I) v^{(1)} + \underbrace{C_2\left(\frac{t^2}{2}\right) e^{\lambda t}(A-\lambda I)^2 v^{(1)}}_{=0} \\
&\quad + C_3 e^{\lambda t} v^{(2)} + C_3 t e^{\lambda t}\underbrace{(A-\lambda I) v^{(2)}}_{=v^{(1)}} + C_3\left(\frac{t^2}{2}\right) e^{\lambda t}\underbrace{(A-\lambda I)^2 v^{(2)}}_{=v}
\end{aligned}
$$

$$
\begin{aligned}
x(t) &= C_1 e^{\lambda t} v + C_2 e^{\lambda t} v^{(1)} + C_2 t e^{\lambda t} v \\
&\quad + C_3 e^{\lambda t} v^{(2)} + C_3 t e^{\lambda t} v^{(1)} + C_3\left(\frac{t^2}{2}\right) e^{\lambda t} v
\end{aligned}
$$

$$
\begin{aligned}
x(t) &= C_1 e^{-t}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + C_2 e^{-t}\begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} + C_2 t e^{-t}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \\
&\quad + C_3 e^{-t}\begin{bmatrix} 0 \\ 0 \\ 1/2 \end{bmatrix} + C_3 t e^{-t}\begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} + C_3\left(\frac{t^2}{2}\right) e^{-t}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}
\end{aligned}
$$

$$
x(t) = \begin{bmatrix} C_1 e^{-t} + tC_2 e^{-t} + \frac{1}{2}t^2 C_3 e^{-t} \\ -C_2 e^{-t} - tC_3 e^{-t} \\ \frac{1}{2}C_3 e^{-t} \end{bmatrix}
$$

■

5.2 Chains of generalised eigenvectors

A practical method for calculating a basis of linearly independent generalized eigenvectors in the general case is an extension of the approach that we used in the last examples.

1) We find a basis of the eigenspace to λ consisting of $r(\lambda)$ linearly independent eigenvectors satisfying the equation $(A-\lambda I)\,u_0 = 0$.Their number $r(\lambda)$ is called **geometric multiplicity** of λ and $r(\lambda) \leq m(\lambda)$.

2) Then for each eigenvector $u_1 \neq 0$ from this basis we find a vector $u_2 \neq 0$ satisfying the equation $(A-\lambda I)\,u_2 = u_1$, and continue this calculation, building a **chain of generalised eigenvectors** $u_1..., u_l$ together with u_1 satisfying equations.

$$(A-\lambda I)\,u_k = u_{k-1} \tag{19}$$

up to the index $k = l$ when $(A-\lambda I)^l\,u_l = 0$ and there will be no solution to the next equation. The largest possible number l is $(m(\lambda) - r(\lambda) + 1)$, but it can be smaller if $r(\lambda) > 1$when there are several linearly independent eigenvectors to λ.

Claim.

Point out that depending on the range of the operator with matrix $(A-\lambda I)$ (column space of the matrix $(A-\lambda I)$) one might need to be careful choosing non-unique (!) eigenvectors u_0 and generalised eigenvectors u_k in the equations (19) so that they would belong to the column space of the matrix $(A-\lambda I)$ (if possible!) to guarantee that the next equations (19) have a solution. Check the solution to the Exercise 864 in the file with exercises, where these observations are important.

Alternatively one can start this algorithm from above, solving first the equation for a generalized eigenvector u_l of rank l and then can apply equations (19) to calculate generalized eigenvectors of lower rank that belong to corresponding chain of generalized eigenvectors. The last vector in this calculation will be an eigenvector.

$$\begin{aligned}
(A-\lambda I)^l\,u_l &= 0; \quad (A-\lambda I)^{l-1}\,u_l = u_1 \neq 0 \\
u_{l-1} &= (A-\lambda I)\,u_l \\
u_{l-2} &= (A-\lambda I)\,u_{l-1} \\
&\dots \\
u_1 &= (A-\lambda I)\,u_2 \\
0 &= (A-\lambda I)\,u_1
\end{aligned}$$

Theorem. The chain of generalised eigenvectors $\{u_k\}_{k=1}^{l}$constructed in *(19)* is linearly independent. □

Proof. It can prove it using the definition of linear independence and analysing the linear combination

$$\sum_{k=1}^{l} \alpha_k u_k = 0.$$

The definition of the chain of eigevalues gives

$$u_k = (A - \lambda I)^{l-k} \, u_l, \quad k = 1, \ldots l - 1$$

Substituting it into the condition for linear independence we observe

$$\sum_{k=1}^{l} \alpha_k \, (A - \lambda I)^{l-k} \, u_l = 0$$

We like to prove that all α_k are zero. We will use the fact that $(A - \lambda I)^m \, u_k = 0$ for all $m \geq l$. We see it by the following calculation:

$$(A - \lambda I)^m \, u_k = (A - \lambda I)^{m-l} \, (A - \lambda I)^l \, u_k = (A - \lambda I)^{m-l} \, \underbrace{(A - \lambda I) \, u_1}_{=0} = 0$$

Applying $(A - \lambda I)^{l-1}$ to the expression $\sum_{k=1}^{l} \alpha_k \, (A - \lambda I)^{l-k} \, u_l = 0$ we get

$$\sum_{k=1}^{l} \alpha_k \, (A - \lambda I)^{2l-k-1} \, u_l = 0$$

Since $(A - \lambda I)^{2l-k-1} \, u_l = 0$ for $k \leq l - 1$, the last equation simplifies and we get

$$\alpha_l \, (A - \lambda I)^{l-1} \, u_l = \alpha_l u_1 = 0$$

that implies $\alpha_l = 0$ eftersom $u_1 \neq 0$. The linear combination above reduces to

$$\sum_{k=1}^{l-1} \alpha_k \, (A - \lambda I)^{l-k} \, u_l = 0$$

Repeating similar argument and multiplying this expression by $(A - \lambda I)^{l-2}$ we get now that

$$\sum_{k=1}^{l-1} \alpha_k \, (A - \lambda I)^{2l-k-1} \, u_l = \alpha_{l-1} \, (A - \lambda I)^{l-1} \, u_l = \alpha_{l-1} u_1 = 0$$

and $\alpha_{l-1} = 0$, because $(A - \lambda I)^{2l-k-2} \, u_l = 0$ for $k \leq l - 2$. Repeating the same argument we arrive to that all $\alpha_k = 0$ $k = 1, \ldots l$.

■

Theorem. A set of generalised eigenvectors $\{u_k\}_{k=1}^{l}$corresponding to different p chains of generalized eigenvectors as in *(19)* corresponding to the same eigenvalue λ is linearly independent if and only if eigenvectors in the bottom of corresponding chains of generalised eigenvectors are linearly independent.$\square$

Proof. Similar as the proof to the previous theorem.

In the case when all eigenvalues $\lambda_1, ..., \lambda_s$ to a real matrix $A \in \mathbb{R}^{N\times N}$ are real, the generalized eigenvectors will be also real and therefore

$$\mathbb{R}^N = \ker(A - \lambda_j)^{m_1} \oplus ... \oplus \ker(A - \lambda_j)^{m_s} = E(\lambda_1) \oplus ... \oplus E(\lambda_s)$$

In this case chains of eigenvectors and generalized eigenvectors built by the procedure as above gives a basis in $\mathbb{R}^N$.

To find a basis in the generalized eigenspace $E(\lambda_j)$ one can carry out calculations in different ways. For example

1) Find all linearly independent eigenvectors to the eigenvalue λ_j that are linearly independent solutions to the equation $(A - \lambda_j I)\, v = 0$ and collect them in a set denoted by $\mathcal{E}$.

2) Then find all linearly independent solutions to $(A - \lambda_j I)^2\, v^{(1)} = 0$ (that are not eigenvectors) and adding them $\mathcal{E}$.

3) Next one finds solutions to $(A - \lambda_j I)^3\, v^{(2)} = 0$ linearly independent from those in $\mathcal{E}$ and collecting them also in $\mathcal{E}$ e.t.c. Continuing in this way one finishes when the total number of the derived linearly independent generalised eigenvectors will be equal to m_j - the algebraic multiplicity of the eigenvalue λ_j.

Alternative more systematic approach to this problem is to calculate such a basis as a chain of generalised eigenvectors corresponding to each of linearly independent eigenvector as it is was suggested in examples before:

$$\begin{aligned}
(A - \lambda_j I)\, v_j &= 0, \\
(A - \lambda_j I)\, v_j^{(1)} &= v_j \\
(A - \lambda_j I)\, v_j^{(2)} &= v_j^{(1)} \\
& \quad e.t.c. \\
(A - \lambda_j I)\, v_j^{(l)} &= v_j^{(l-1)}
\end{aligned}$$

This approach has also an advantage that using chains of generalised eigenvectors as a basis leads to a particularly simple representation of the system of equations (4) with matrix A in so called Jordan canonical form, that we will learn later.

Substituting the expression (18) for arbitrary initial data ξ in to the general formula above and calculating all matrix $(A - \lambda_j I)$ powers and matrix-vector, multiplications we get a general solution with a set of arbitrary coefficients $C_1, ..., C_N$.

Keep in mind that $(A - \lambda_j I)\, v_j = 0$ and $(A - \lambda_j I)^2\, v_j^{(1)} = 0$ e.t.c., so many terms in the general expression for the solution can be zeroes.

Initial value problems. To solve an I.V.P. one needs to express a particular initial data ξ in terms of the basis of generalized eigenvectors soving a linear system of equations for coefficients $C_1, ..., C_N$ in (18) like for example $\xi = C_1 v + C_2 v^{(1)} + C_3 v^{(2)}$. We solve a linear system of equations for $C_1, ..., C_N$.

Look for exercises in a separate file Exercises_3.pdf with exercises on linear autonomous systems of ODE. Check modulus with lecture notes in Canvas.

Lecture 4 (mainly exercises considered in a separate file) Summary of the theory on autonomous linear ODEs given in the first week of the course.

1. Initial value problem

$$\begin{aligned} x'(t) &= Ax(t), \quad x(0) = \xi \\ A &\in \mathbb{R}^{N\times N} \quad (A \in \mathbb{C}^{N\times N}) \\ x &: \mathbb{R} \to \mathbb{R}^N \quad (x : \mathbb{R} \to \mathbb{C}^N) \end{aligned}$$

2. Existence of solutions and representation of the solution to an I.V.P. by a matrix exponent:

$$x(t) = \exp(At)\xi = \left(\sum_{k=0}^{\infty} \frac{1}{k!} A^k (t-\tau)^k\right)\xi$$

Uniqueness of solutions to I.V.P. based on Grönwall's inequality.

3. The dimension of the solution space ($= N$). If $\{b_k\}_k^N$ is a basis in $\mathbb{R}^N$ ($\mathbb{C}^N$), then functions $\{\exp(At)b_k\}_k^N$ form a basis of the solution space to the equation $x'(t) = Ax(t)$.

4. Generalized eigenvectors v and generalized eigenspaces $E(\lambda)$ for an eigenvalue λ to the matrix A, having algebraic multiplicity $m(\lambda)$

$$(A - \lambda I)^{m(\lambda)} v = 0$$

$\dim(E(\lambda)) = m(\lambda)$.

5. Decomposition of $\mathbb{C}^N$ into a direct sum of generalized eigenspaces of all distinct eigenvalues $\lambda_1, ...\lambda_s$ of a matrix A.

$$\mathbb{C}^N = E(\lambda_1) \oplus ... \oplus E(\lambda_s)$$

$\mathbb{R}^N = E(\lambda_1) \oplus ... \oplus E(\lambda_s)$ if all eigenvalues and the matrix A are real.

6. An important idea:

$$\exp(At) = \exp(t\lambda I)\exp((A - \lambda I)t) = e^{\lambda t}\sum_{k=0}^{\infty} \frac{t^k}{k!}(A - \lambda I)^k$$

implies that for a generalized eigenvector z to te eigenalue λ with algebraic multiplicity $m(\lambda)$ the $\exp(At)\,z$ can be expressed explicitely.

$$
\begin{aligned}
x_z(t) &= \exp(At)\,z = \exp(t\lambda I)\exp((A-\lambda I)\,t)\;z = \\
\left(e^{\lambda t} I\right)\sum_{k=0}^{\infty}\frac{t^k}{k!}\,(A-\lambda I)^k\,z &= e^{\lambda t}\sum_{k=0}^{m(\lambda)-1}\frac{t^k}{k!}\,(A-\lambda I)^k\,z
\end{aligned}
\tag{20}
$$

7. Together with the decomposition theorem it gives a way to find a basis of the solution space for the equation $x' = Ax$ by finding a basis for each of $E(\lambda_j)$, $j = 1, ..., s$. We do it using chains of generalised eigenvectors corresponding to each linearly independent eigenvector v_j of the eigenalue λ_j:

$$
\begin{aligned}
(A-\lambda_j I)\,v_j &= 0, \\
(A-\lambda_j I)\,v_j^{(1)} &= v_j \\
(A-\lambda_j I)\,v_j^{(2)} &= v_j^{(1)} \\
& \quad e.t.c. \\
(A-\lambda_j I)\,v_j^{l} &= v_j^{l-1}
\end{aligned}
$$

We continue with examples collected in the separate file Exercises-Lecture4.pdf .

5.3 Real solutions for systems with real matrices having complex eigenvalues.

We will consider an example of a system in plane with real matrix having two simple, conjugate complex eigenvalues (no more because of the small dimension). The idea of solution was to build a complex solution corresponding to one of these eigenvalues and use it's real and imaginary part at two linearly independent solutions to construct a general solution.

The same idea works in the general case when a real matrix might have conjugate complex eigenvalues (might be multiple in higher dimensions).

We build a basis of eigenvectors and generalized eigenvectors for invariant generalized eigenspaces corresponding to distinct conjugate complex eigenvalues. One can start with one of these eigenvalues and then can just choose the basis for the second one as a complex conjugate . Then we construct arbitrary complex solutions in the invariant generalized eigenspace corresponding to the first of these conjugate eigenvalues. Real and imaginary parts of these solutions are linearly independent and build a basis of solutions in the corresponding real invariant subspace.

Example 2. Real matrix with complex eigenvalues.

$x' = Ax$ with $A = \begin{bmatrix} 3 & -2 \\ 4 & -1 \end{bmatrix}$, find a general real solution to the system. In this case we find first a general complex solution and then construct a general real solution based on it.

Solution. $A = \begin{bmatrix} 3 & -2 \\ 4 & -1 \end{bmatrix}$, characteristic polynomial: $\lambda^2 - 2\lambda + 5 = 0$;

Hint. We point out here that in the case of 2×2 matrices the characteristic polynomial always has a simple representation

$$p(\lambda) = \lambda^2 - \lambda \mathrm{tr}(A) + \det(A)$$

where $\mathrm{tr}(A)$ is the sum of diagonal elements in A called trace, and $\det(A)$ is determinant. Here $\mathrm{tr}(A)$ $=\lambda_1 + \lambda_2$; $\det A = \lambda_1\lambda_2$.$\blacksquare$

Eigenvalues are: $\lambda_1 = 1 - 2i$, and $\lambda_2 = 1 + 2i$ - complex conjugate

They are complex conjugate:

$$\begin{aligned} \overline{\lambda_1} &= \lambda_2 \\ p(\lambda) &= (\lambda - \lambda_1)(\lambda - \lambda_2) \end{aligned}$$

because the characteriscic polynomial has real coefficients.

Eigenvectors satisfy the equations $(A - \lambda I)v_1 = \begin{bmatrix} 2+2i & -2 \\ 4 & -2+2i \end{bmatrix} v_1 = 0$ and

$\begin{bmatrix} 2-2i & -2 \\ 4 & -2-2i \end{bmatrix} v_2 = 0.$

These eigenvectors muct be also complex conjugate. We see it by considering the equations for v_1 that is

$(A - \lambda_1 I)v_1 = 0$ and its formal complex conjugate $(A - \overline{\lambda_1} I)\overline{v_1} = 0$ that is satisfied because the conjugate of the real matrix A is the matrix A itself. Therefore $\overline{v_1}$ is the eigenvector corresponding to the eigenvalue $\lambda_2 = \overline{\lambda_1}$. We point out that this argument is independent of this particular example and would be valid for any real matrix with complex eigenvalues.

The first and the second equation in each of these systems are equivalent because rows are linearly dependent (homogeneous system has non-trivial solutions and the determinant of the matrix $A - \lambda I$ is zero).

We solve the first equation in the first system by choosing the first component in the complex vector v_1 equal to 1. It implies that the second component denoted here by z satisfies the equation $2 + 2i - 2z = 0$ and therefore $z = 1 + i$. The second eigenvector is just the complex conjugate of the first one.

$v_1 = \left\{ \begin{bmatrix} 1 \\ 1+i \end{bmatrix} \right\} \leftrightarrow \lambda_1 = 1 - 2i$, and $v_2 = \overline{v_1} = \left\{ \begin{bmatrix} 1 \\ 1-i \end{bmatrix} \right\} \leftrightarrow \lambda_2 = 1 + 2i.$

They are linear independent as eigenvectors corresponding to different eigenvalues.

One complex solution is $x^*(t) = e^{\lambda_1 t} v_1 = e^{(1-2i)t} \begin{bmatrix} 1 \\ 1+i \end{bmatrix}$, another one is $y^*(t) = e^{\lambda_2 t} v_2 = e^{(1+2i)t} \begin{bmatrix} 1 \\ 1-i \end{bmatrix}$

$x^*(t)$ and $y^*(t)$ are linearly independent at any time as corresponding to linearly independent initial vectors v_1 and v_2 (according to the theorem before) and build a basis of complex solutions to the system. Therefore the matrix $[x^*(t), y^*(t)]$ has determinant $\det([x^*(t), y^*(t)]) \neq 0$.

Two linearly independent real solutions can be chosen as real and imaginary parts of $x^*(t)$ (or $y^*(t)$): $\mathrm{Re}\,[x^*(t)] = \frac{1}{2}(x^*(t) + y^*(t))$ and $\mathrm{Im}\,[x^*(t)] = \frac{1}{2i}(x^*(t) - y^*(t))$ that are linearly independent because the the matrix $T = \frac{1}{2}\begin{bmatrix} 1 & 1/i \\ 1 & -1/i \end{bmatrix}$ of the transformation

$$
\begin{aligned}
[x^*(t), y^*(t)]\,T &= \begin{bmatrix} x_1^* & y_1^* \\ x_2^* & y_2^* \end{bmatrix} \begin{bmatrix} 1/2 & 1/(2i) \\ 1/2 & -1/(2i) \end{bmatrix} \\
\begin{bmatrix} \frac{1}{2}x_1^* + \frac{1}{2}y_1^* & \frac{1}{2i}x_1^* - \frac{1}{2i}y_1^* \\ \frac{1}{2}x_2^* + \frac{1}{2}y_2^* & \frac{1}{2i}x_2^* - \frac{1}{2i}y_2^* \end{bmatrix} &= [\mathrm{Re}\,[x^*(t)]\,, \mathrm{Im}\,[x^*(t)]]
\end{aligned}
$$

is invertible: $\det T = -\frac{1}{2i} \neq 0$ and therefore, by the property of the determinant for the product of matrices,

$$\det\,[x^*(t), y^*(t)]\det(T) = \det\left([\mathrm{Re}\,[x^*(t)]\,, \mathrm{Im}\,[x^*(t)]]\right) \neq 0$$

and $\mathrm{Re}\,[x^*(t)]$ and $\mathrm{Im}\,[x^*(t)]$ are linearly independent.

Therefore real valued vector functions $\mathrm{Re}\,[x^*(t)]$ and $\mathrm{Im}\,[x^*(t)]$ can be used as a basis for representing the general real solution to the system:

$$x(t) = C_1\,\mathrm{Re}\,[x^*(t)] + C_2\,\mathrm{Im}\,[x^*(t)]\,.$$

We express $x^*(t)$ with help of Euler formulas and separate real and imaginary parts

$$
\begin{aligned}
&x^*(t) = e^{\lambda_1 t}v_1 = e^{(1-2i)t}\begin{bmatrix} 1 \\ 1+i \end{bmatrix} = e^t\left(\cos 2t - i\sin 2t\right)\begin{bmatrix} 1 \\ 1+i \end{bmatrix} = \\
&e^t\begin{bmatrix} \cos 2t - i\sin 2t \\ (1+i)\cos 2t + (1-i)\sin 2t \end{bmatrix} = e^t\begin{bmatrix} \cos 2t - i\sin 2t \\ \cos 2t + \sin 2t + i\left(\cos 2t - \sin 2t\right) \end{bmatrix} = \\
&e^t\begin{bmatrix} \cos 2t \\ \cos 2t + \sin 2t \end{bmatrix} - i\,e^t\begin{bmatrix} \sin 2t \\ (\sin 2t - \cos 2t) \end{bmatrix}
\end{aligned}
$$

The answer follows as a linear combination of real and imaginary parts: $x(t) = C_1\,\mathrm{Re}\,[x^*(t)] + C_2\,\mathrm{Im}\,[x^*(t)]\,.$

Answer: $x(t) = C_1 e^t\begin{bmatrix} \cos 2t \\ \cos 2t + \sin 2t \end{bmatrix} + C_2 e^t\begin{bmatrix} \sin 2t \\ \sin 2t - \cos 2t \end{bmatrix}.$

We will transform this expression to clarify its geometric meaning and the shape of orbits in the phase plane. We observe first that if we drop exponents e^t, in the expression for $x(t)$ and consider the expression $x(t)e^{-t} = C_1\begin{bmatrix} \cos 2t \\ \cos 2t + \sin 2t \end{bmatrix} + C_2\begin{bmatrix} \sin 2t \\ \sin 2t - \cos 2t \end{bmatrix}$,we will observe that it represents a movement along ellipses in the plane.

We use an elementary trick that makes that any linear combination of $\sin(\gamma)$ and $\cos(\gamma)$ is $C\sin(\gamma + \beta)$

or $C\cos(\gamma-\beta)$ with some constants C, β.

$$
\begin{aligned}
x_1(t)e^{-t} &= C_1\cos(2t)+C_2\sin(2t) = \\
&\sqrt{C_1^2+C_2^2}\left(\left(\frac{C_1}{\sqrt{C_1^2+C_2^2}}\right)\cos 2t+\left(\frac{C_2}{\sqrt{C_1^2+C_2^2}}\right)\sin 2t\right) \\
&= \sqrt{C_1^2+C_2^2}\,(\cos(\theta)\cos 2t+\sin(\theta)\sin 2t) \\
&= \sqrt{C_1^2+C_2^2}\cos(2t-\theta) \\
\theta &= \arccos\left(\left(\frac{C_1}{\sqrt{C_1^2+C_2^2}}\right)\right)
\end{aligned}
$$

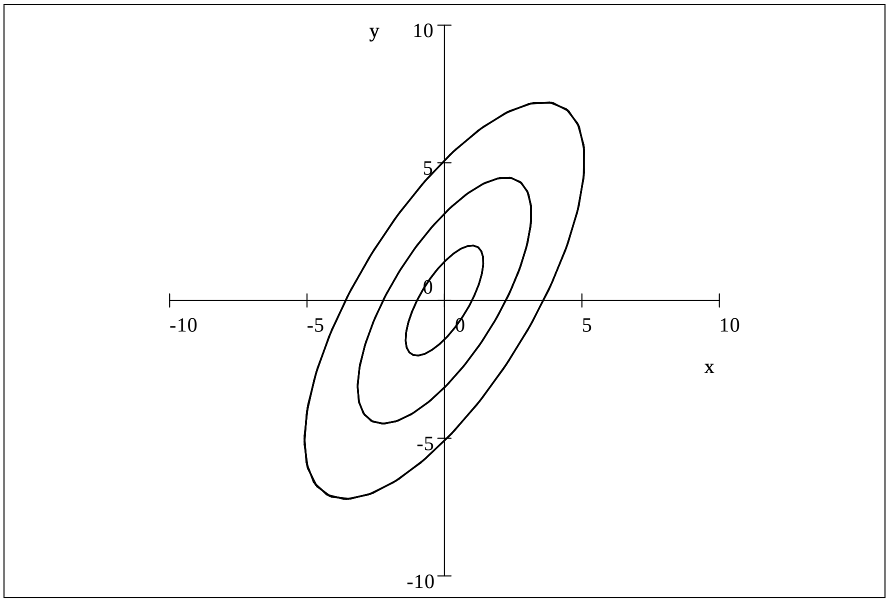
Similarly

$$
\begin{aligned}
[x_2(t)-x_1(t)]\,e^{-t} &= C_1\sin(2t)-C_2\cos(2t) = \\
&\sqrt{C_1^2+C_2^2}\left(\left(\frac{C_1}{\sqrt{C_1^2+C_2^2}}\right)\sin 2t-\frac{C_2}{\sqrt{C_1^2+C_2^2}}\cos 2t\right) \\
&= \sqrt{C_1^2+C_2^2}\,(\cos(\theta)\sin 2t-\sin(\theta)\cos 2t) \\
&= \sqrt{C_1^2+C_2^2}\sin(2t-\theta)
\end{aligned}
$$

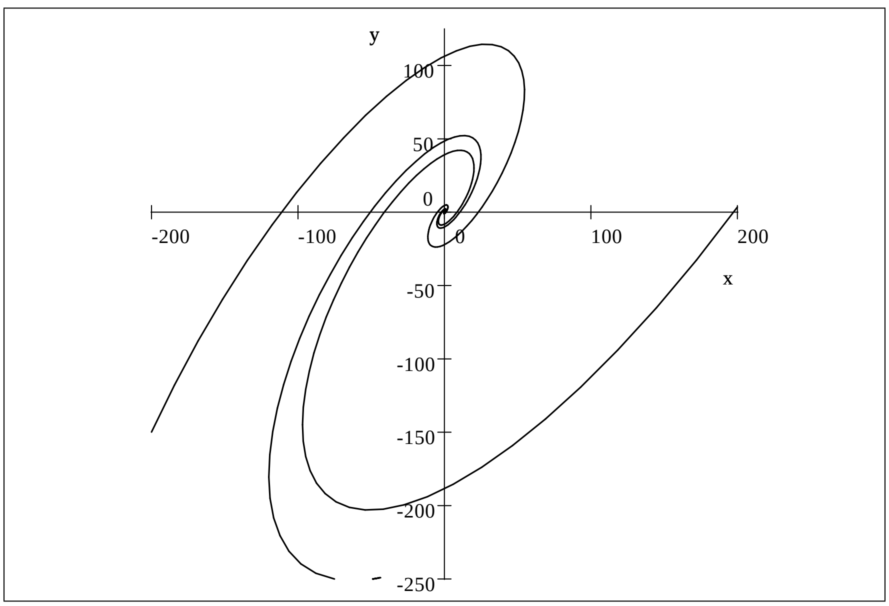
Finally we arrive to a parametric expression for a periodic movement along ellipses with size depending on C_1 and C_2.

$$
\begin{aligned}
x(t)e^{-t} &= C_1\begin{bmatrix}\cos 2t\\ \cos 2t+\sin 2t\end{bmatrix}+C_2\begin{bmatrix}\sin 2t\\ \sin 2t-\cos 2t\end{bmatrix} \\
&= \sqrt{C_1^2+C_2^2}\begin{bmatrix}\cos(2t-\theta)\\ \cos(2t-\theta)+\sin(2t-\theta)\end{bmatrix} \\
&= \sqrt{C_1^2+C_2^2}\begin{bmatrix}\cos(2t-\theta)\\ \sqrt{2}\,[\sin(\pi/4)\cos(2t-\theta)+\cos(\pi/4)\sin(2t-\theta)]\end{bmatrix} \\
&= \sqrt{C_1^2+C_2^2}\begin{bmatrix}\cos(2t-\theta)\\ \sqrt{2}\,[\sin(2t-\theta+\pi/4)]\end{bmatrix}
\end{aligned}
$$

illustrated in the next picture:



This movement is modulated in our solution $x(t)$ by the exponential term e^t giving orbits as spirals going to infinity out of the origin that is an unstable equilibrium point for this system.



Example. It is good to consider here the solution to the exercise 858.

Ideas about solutions to systems with complex eigenvalues demonstrated in exercises can in the general situation be expressed by the following Theorem.

Theorem 2.14. p. 38 on real solutions to autonomous systems with real matrix and complex eigenvalues (without proof)

Let $A \in R^{N\times N.}$ for λ an eigenvalue, let $m(\lambda)$ be the algebraic multiplicity of λ, $E(\lambda) = \ker(A - \lambda I)^{m(\lambda)}$ denote it's generalised eigenspace. Let $B(\lambda)$ be a basis in $E(\lambda)$ chosen to be real for real λ.

For all $z \in \mathbb{C}^N$, we denote x_z, $y_z : R \to R^N$ real solutions to the equation $x' = Ax$ as

$$x_z = \exp(At)\,\mathrm{Re}\, z, \quad y_z = \exp(At)\,\mathrm{Im}\, z$$

Then

1) Let B_0 (respectively B_+) denote the union of all $B(\lambda)$ for all real eigenvalues λ to A (correspondingly for all λ with $\operatorname{Im}\lambda > 0$) The set of real functions given by

$$\{x_z, \quad z \in B_0 \cup B_+\} \cup \{y_z : \quad z \in B_+\}$$

forms a basis of the solutuion space to $x' = Ax$.

2) If λ is a real eigenvalue to A, then for every generalized eigenvector $z \in E(\lambda)$, the solution x_z is expressed as

$$x_z(t) = e^{\lambda t} \sum_{k=0}^{m(\lambda)-1} \frac{t^k}{k!} (A - \lambda I)^k z$$

3) If $\lambda = \alpha + i\beta$ with $\beta \neq 0$, is an eigenvalue of A,then for every generalized eigenvector $z \in E(\lambda)$, solutions $x_z = \exp(At)\operatorname{Re} z$ and $y_z = \exp(At)\operatorname{Im} z$ with initial data $\operatorname{Re} z$ and $\operatorname{Im} z$ are expressed as

$$\begin{aligned}
x_z(t) &= e^{\lambda t} \sum_{k=0}^{m(\lambda)-1} \frac{t^k}{k!} (A - \lambda I)^k \operatorname{Re} z = e^{\alpha t} \sum_{k=0}^{m(\lambda)-1} \frac{t^k}{k!} \left[\cos(\beta t) \operatorname{Re}\left((A - \lambda I)^k z\right) - \sin(\beta t) \operatorname{Im}\left((A - \lambda I)^k z\right)\right] \\
y_z(t) &= e^{\lambda t} \sum_{k=0}^{m(\lambda)-1} \frac{t^k}{k!} (A - \lambda I)^k \operatorname{Im} z = e^{\alpha t} \sum_{k=0}^{m(\lambda)-1} \frac{t^k}{k!} \left[\cos(\beta t) \operatorname{Re}\left((A - \lambda I)^k z\right) + \sin(\beta t) \operatorname{Im}\left((A - \lambda I)^k z\right)\right]
\end{aligned}$$

□

The theorem shows the how $m(\lambda)$ real linearly independent solutions can be obtained for a real matrix A with complex eigenvalues λ. The part 1) of the theorem shows that such solutions build a real basis of the solution space for $x' = Ax$ with a real matrix.

Lecture 5.

5.4 Jordan canonical form of a matrix. Functions of matrices.

5.5 Change of variables. Properties of similar matrices. Block matrices.

We tried in previous lectures to find a basis $\left\{v_1, v_1^{(1)}, ...\right\}$ in $\mathbb{C}^N$ or in $\mathbb{R}^N$ such that expressing initial data ξ in I.V.P.

$$x'(t) = Ax(t), \quad x(0) = \xi$$

in terms of this basis led to a particularly simple expression of the solution as an explicit linear combination including polynomials of matrices $t(A - \lambda_i I)$ acting on basis vectors. We can interpret these results by introducing a linear change of variables

$$x = Vy; \quad y = V^{-1}x$$

with matrix V of this transformation having columns consisting of N linearly independent basis vectors.

In terms of the new variable y the system of ODEs has the form:

$$x'(t) = A\left(Vy\right), \quad x(0) = \xi$$

Multiply by V^{-1} left and right hand sides:

$$V^{-1}x'(t) = y'(t) = \left(V^{-1}AV\right)y, \quad y(0) = V^{-1}\xi$$

$$y'(t) = V^{-1}\left(A\left(Vy\right)\right)$$

In the case when the matrix A has N linearly independent eigenvectros the matrix $V^{-1}AV = D$ is diagonal with eigenvalues $\{\lambda_1 ..., \lambda_j, ...\}$ of the matrix A standing on the diagonal of D - $m(\lambda_j)$ times equal to the algebraic multiplicity of λ_j. The number $r(\lambda_j)$ of linearly independent eigenvectors belonging to λ_j is called geometric multiplicity of λ_j and is equal to $m(\lambda_j)$ in this case.

Definition. Matrices A and $V^{-1}AV$ are called similar.

They have several characteristics the same: determinant, and characteristic polynomials. It is a simple consequence of properties of determinants of products of matrices.

Prove it as an exercise using: $\det(AB) = \det(A)\det(B)$; $\det(B^{-1}) = (\det B)^{-1}$ if $\det B \neq 0$.

Using the associative property of matrix multiplication we arrive to the property

Theorem. If matrices A and B are similar through $B = V^{-1}AV$, $A = VBV^{-1}$ then

$$\begin{aligned}
B^k &= V^{-1}(A^k)V; \\
\exp(B) &= V^{-1}(\exp A)V \\
A^k &= V(B^k)V^{-1} \\
\exp(A) &= V(\exp B)V^{-1}
\end{aligned}$$

Prove it as an exercise.

Corollary. If the matrix A is diagonalisable, then $\exp(A) = V\exp(D)V^{-1}$ where V matrix of linearly independent eigenvectors and the matrix D is diagonal matrix of eigenvalues λ_j and $\exp(D)$ is a diagonal matrix with $\exp(\lambda_j)$ on the diagonal. In this case the system in new variables $y(t) = V^{-1}x(t)$ consists of independent differential equations $y_j'(t) = \lambda_j y_j(t)$ for he components $y_j(t)$ of $y(t)$ that have simple solutions $y_j(t) = C_j e^{\lambda_j t}$

Definition. Block - diagonal matrices

Block-diagonal matrices are square matrices that have a number of square blocks $\mathbb{B}_1$,... along diagonal and other terms all zero. For example:

$$B = \begin{bmatrix} \mathbb{B}_1 & \mathbb{O} & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \mathbb{B}_2 & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \mathbb{B}_3 & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \mathbb{O} & \mathbb{B}_4 \end{bmatrix}$$

$$B^2 = \begin{bmatrix} \mathbb{B}_1 & \mathbb{O} & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \mathbb{B}_2 & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \mathbb{B}_3 & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \mathbb{O} & \mathbb{B}_4 \end{bmatrix} \begin{bmatrix} \mathbb{B}_1 & \mathbb{O} & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \mathbb{B}_2 & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \mathbb{B}_3 & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \mathbb{O} & \mathbb{B}_4 \end{bmatrix} = \begin{bmatrix} \mathbb{B}_1^2 & \mathbb{O} & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \mathbb{B}_2^2 & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \mathbb{B}_3^2 & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \mathbb{O} & \mathbb{B}_4^2 \end{bmatrix}$$

These matrices have a property that their powers lead to block diagonal matrices of the same structure with powers of original blocks on the diagonal:

$$B^k = \begin{bmatrix} (\mathbb{B}_1)^k & \mathbb{O} & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & (\mathbb{B}_2)^k & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & (\mathbb{B}_3)^k & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \mathbb{O} & (\mathbb{B}_4)^k \end{bmatrix}$$

This simple observation leads immediately to the formula for the exponent of a block diagonal matrix.

$$\exp(B) = \sum_{k=0}^{\infty} \frac{1}{k!} B^k = \begin{bmatrix} \exp(\mathbb{B}_1) & \mathbb{O} & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \exp(\mathbb{B}_2) & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \exp(\mathbb{B}_3) & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \mathbb{O} & \exp(\mathbb{B}_4) \end{bmatrix}$$

In fact the same relation would be valid even for an arbitrary analytical function f with power series $f(z) = \sum_{k=0}^{\infty} a_k z^k$, converging in the whole $\mathbb{C}$:

$$f(B) = \begin{bmatrix} f(\mathbb{B}_1) & \mathbb{O} & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & f(\mathbb{B}_2) & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & f(\mathbb{B}_3) & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \mathbb{O} & f(\mathbb{B}_4) \end{bmatrix}$$

Claim. Let the space $\mathbb{C}^N$ or $\mathbb{R}^N$ be represented as a direct sum of subspaces $V_1, ..., V_s$, **invariant**

under the action of operator Ax:

$$\mathbb{C}^N = V_1 \oplus V_2 \oplus ... \oplus V_s$$

It means that for all vectors $z \in V_k$ it is valid that $Az \in V_k$ for $k = 1, ...s$.

Then there is a basis $\{u_1, ..., u_N\}$ in $\mathbb{C}^N$, correspondingly $\mathbb{R}^N$ such that the operator Ax in this basis has matrix B similar to $A : B = U^{-1}AU$, or

$$UB = AU$$

that is *block diagonal,* with blocks of size equal to dimensions of subspaces $V_1, ..., V_s$ and matrix U that has columns u_1,...,u_N.

The basis $\{u_1, ..., u_N\}$ is easy to choose as a union of bases for each invariant subspace V_j. It is easy to observe that this construction leads to a block diagonal matrix for the operator Ax because columns with index j in the matrix B are equal to $U^{-1}Au_j$ that are coordinates of vectors Au_j in terms of the basis $\{u_1, ..., u_N\}$ and belong to the same invariant subspace as u_j.

We illustrate this fact on a simple example with two invariant subspaces.

Consider a decomposition of the space $\mathbb{C}^N$ into *the direct sum* of two subspaces V and W , $\dim V = m$, $\dim W = p$, $m + p = N$ invariant with respect to the operator defined by the multiplication Ax. Choose base vectors in each of these subspaces: $\{u_1, ..., u_m\}$ and $\{w_1, ..., w_p\}$. They constitute a basis $\{u_1, ..., u_m, w_1, ..., w_p\}$ for the whole space $\mathbb{C}^N$.

Introduce a matrix $T = [u_1, ..., u_m, w_1, ..., w_p]$ with basis vectors of the whole $\mathbb{C}^N$ collected according to the invariant subspace they belong to.

Represent a vector x in terms of this basis: $x = Ty$ where

$$y = [y_1, ..., y_m, y_{m+1}, ..., y_{p+m}]$$

is a vector of coordinates of x in the basis consisting of columns in T. The operator Ax acting on the vector x is expressed in terms of these coordinates y as

$$Ax = ATy$$

We express now the image of this operation also in terms of
the basis $\{u_1, ..., u_m, w_1, ..., w_p\}$:

$$T\left(T^{-1}Ax\right) = ATy$$

Here $(T^{-1}Ax)$ gives coordinates of the vector Ax in terms of the basis $\{u_1, ..., u_m, w_1, ..., w_p\}$ that are columns in the matrix T. It implies that

$$T^{-1}Ax = \left(T^{-1}AT\right)y$$

So the matrix $(T^{-1}AT)$ is a standard matrix of the original mapping Ax in terms of the basis $\{u_1, ..., u_m, w_1, ..., w_p\}$

associated with invariant subspaces V and W. Now observe that taking vector of y - coordinates with only components $y_1, ..., y_m$ non-zero we get vectors that belong to the invariant subspace V, namely vectors having only y - coordinates $1, ..., m$ non-zero. It means that first m columns in $(T^{-1}AT)$ must have elements $m+1, ...m+p$ equal to zero because A maps V into itself.

$$T^{-1}AT = \begin{bmatrix} \mathbb{B}_1 & \mathbb{O} \\ \mathbb{O} & \mathbb{B}_2 \end{bmatrix}$$

If we choose y coordinates with only components $y_{m+1}, ...y_{m+p}$ non-zero, we get a vector that belongs to the subspace W, namely vectors that have only coordinates $m+1$, ..., $m+p$ non-zero. It means that last p columnst in $(T^{-1}AT)$ must have elements $1, ...m$ equal to zero because A maps W into itself. It means finaly that $(T^{-1}AT)$ has a block diagonal structure with blocks of size $m \times m$ and $p \times p$ corresponding to the invariant subspaces V and W.

5.6 Jordan canonical form of a matrix and it's functions.

We will observe now that a basis of generalised eigenvectors

$$\mathbb{C}^N = E(\lambda_1) \oplus E(\lambda_2) \oplus ... \oplus E(\lambda_s)$$

build with help of chains of generalised eigenvectors as we discussed before, leads to a particular "canonical" matrix J similar to the matrix A by the transformation

$$V^{-1}AV = J$$

or $A = VJV^{-1}$ with the matrix

$$V = \left[...v, v^{(1)}, ..., v^{(r-1)}...\right]$$

where columns are generalised eigenvectors from different chains of generalised eigenvectors corresponding to linearly independent eigenvectors put in the same order as in (21).

Consider first an $m \times m$ matrix A in $\mathbb{C}^{m\times m}$ that has **only one eigenvalue** λ from the characteristic polynomial $p(z) = (z - \lambda)^m$, of multiplicity m and **only one** linearly independent eigenvector v.Corresponding chain of generalised eigenvectors $\left\{v, v^{(1)}, ..., v^{(m-1)}\right\}$ has rank m equal to the dimension of the space and satisfies equations:

$$\begin{aligned}
(A - \lambda I)\, v &= 0, \\
(A - \lambda I)\, v^{(1)} &= v \\
(A - \lambda I)\, v^{(2)} &= v^{(1)} \\
& \quad e.t.c. \\
(A - \lambda I)\, v^{(m-1)} &= v^{(m-2)}
\end{aligned} \tag{21}$$

$(A - \lambda I)^m\, v^{(m-1)} = 0$; $(A - \lambda I)^{m-1}\, v^{(m-1)} = v \neq 0$.

We rewrite this chain of equations as

$$\begin{aligned}
Av &= \lambda v \\
Av^{(1)} &= \lambda v^{(1)} + v \\
Av^{(2)} &= \lambda v^{(2)} + v^{(1)} \\
& \quad e.t.c. \\
Av^{(m-1)} &= \lambda v^{(m-1)} + v^{(m-2)}
\end{aligned}$$

Using the definition of the matrix product and the matrix V defined as

$$V = \left[v, v^{(1)}, ..., v^{(m-1)}\right]$$

we observe that

$$\begin{aligned} AV &= \left[Av, Av^{(1)}, ..., Av^{(m-1)}\right] \\ VD &= \left[\lambda v, \lambda v^{(1)}, ..., \lambda v^{(m-1)}\right] \\ V\mathcal{N} &= \left[0, v, v^{(1)}, ..., v^{(m-2)}\right] \end{aligned}$$

and that vector equations for the chain of generalised eigenvectors are equivalent to the matrix equation

$$AV = VD + V\mathcal{N} = V(D + \mathcal{N})$$

where D is the diagonal matrix with the eigenvalue λ on the diagonal and the matrix $\mathcal{N}$ has all elements zero except elements over the diagonal that are equal to one:

$$D = \begin{bmatrix} \lambda & 0 & 0 & ... & 0 & 0 \\ 0 & \lambda & 0 & ... & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & ... & 0 & 0 \\ 0 & 0 & 0 & ... & \lambda & 0 \\ 0 & 0 & 0 & ... & 0 & \lambda \end{bmatrix}; \quad \mathcal{N} = \begin{bmatrix} 0 & 1 & 0 & ... & 0 & 0 \\ 0 & 0 & 1 & ... & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & ... & 1 & 0 \\ 0 & 0 & 0 & ... & 0 & 1 \\ 0 & 0 & 0 & ... & 0 & 0 \end{bmatrix};$$

Shifting property of the right multiplication by the matrix $\mathcal{N}$.

The specific structure of $\mathcal{N}$ makes that the product $B\mathcal{N}$ of an arbitrary square matrix B by the matrix $\mathcal{N}$ from the right is a matrix where each column k is a column $k-1$ from the matrix B shifted one step to the right, except the first one that consists of zeroes. It follows from the definition of the matrix product and the observation that elements from the column k in the matrix B in the product $B\mathcal{N}$ meet exactly one non zero element 1 in the column $k+1$ in the matrix $\mathcal{N}$:

$$B = \begin{bmatrix} B_{11} & B_{12} & B_{13} & ... & B_{1(m-1)} & B_{1m} \\ B_{21} & B_{22} & B_{23} & ... & B_{2(m-1)} & B_{2m} \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & ... & \vdots & \vdots \\ B_{(m-1)1} & B_{(m-1)2} & B_{(m-1)3} & ... & B_{(m-1)(m-1)} & B_{(m-1)m} \\ B_{m1} & B_{m2} & B_{m3} & ... & B_{m(m-1)} & B_{mm} \end{bmatrix}; \quad \mathcal{N} = \begin{bmatrix} 0 & 1 & 0 & ... & 0 & 0 \\ 0 & 0 & 1 & ... & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & ... & 1 & 0 \\ 0 & 0 & 0 & ... & 0 & 1 \\ 0 & 0 & 0 & ... & 0 & 0 \end{bmatrix};$$

We observe this transformation in equations for the chain of generalized eigenvectors with the matrix V instead of an arbitrary matrix B.

Observe also that $\mathcal{N}^m = 0$, were m is the size of $\mathcal{N}$.

Therefore

$$
\begin{aligned}
AV &= V(D+\mathcal{N}) \\
V^{-1}AV &= (D+\mathcal{N}) = J
\end{aligned}
$$

Definition of the Jordan block. The matrix $J = D + \mathcal{N}$

$$
J = \begin{bmatrix}
\lambda & 1 & 0 & \dots & 0 & 0 \\
0 & \lambda & 1 & \dots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \dots & 1 & 0 \\
0 & 0 & 0 & \dots & \lambda & 1 \\
0 & 0 & 0 & \dots & 0 & \lambda
\end{bmatrix}
$$

is called **Jordans block.** Here D is a diagonal matrix with the eigenvalue λ on the diagonal and the matrix $\mathcal{N}$ defined above, consists of zeroes except for the diagonal above the main one consisting of ones.

We have proven the following theorem.

Theorem (special case of Theorem A.9 , p. 268) Let $m \times m$ matrix A have one eigenvalue of multiplicity m (characteristic polynomial $p(z) = (z - \lambda)^m$) and only one linearly independent eigenvector v. Then the matrix A is similar to the Jordans block J with the similarity relations:

$$
\begin{aligned}
A &= VJV^{-1} \\
J &= V^{-1}AV
\end{aligned}
$$

where the matrix V has columns $V = \left[v, v^{(1)}, \dots, v^{(m-1)}\right]$ that are elements from the chain of generalized eigenvectors built as solutions to the equations (21).

The "shifting" property of the matrix $\mathcal{N}$ implies that $\mathcal{N}^2$ consists of zeroes except the second diagonal over the main one filled by 1, $\mathcal{N}^3$ consists of zeroes except the third diagonal over the main one filled by 1, and finally $\mathcal{N}^m = 0$.

Definition A matrix with such property that for some integer r we have $\mathcal{N}^r = 0$ is called nilpotent.

Corollary

$$
\exp(J) = \exp(D+\mathcal{N}) = \exp(D)\exp(\mathcal{N}) = e^{\lambda} \sum_{k=0}^{m-1} \frac{1}{k!} (\mathcal{N})^k \tag{22}
$$

$$
\exp(J) = e^{\lambda} \begin{bmatrix}
1 & 1 & 1/2 & \dots & \frac{1}{(m-2)!} & \frac{1}{(m-1)!} \\
0 & 1 & 1 & \dots & \frac{1}{(m-3)!} & \frac{1}{(m-2)!} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \dots & 1 & 1/2 \\
0 & 0 & 0 & \dots & 1 & 1 \\
0 & 0 & 0 & \dots & 0 & 1
\end{bmatrix}
$$

because $\exp(J) = \exp(\lambda I + \mathcal{N}) = \exp(\lambda I)\exp(\mathcal{N}) = e^{\lambda}\sum_{k=0}^{m-1}\frac{1}{k!}(\mathcal{N})^k$ and each term with index k in the sum is a matrix with k -th diagonal over the main one, filled by $\frac{1}{k!}$ ■

Similarly

$$\exp(Jt) = e^{\lambda t}\sum_{k=0}^{m-1}\frac{t^k}{k!}(\mathcal{N})^k \tag{23}$$

$$\exp(Jt) = e^{\lambda t}\begin{bmatrix} 1 & t & t^2/2 & \dots & \frac{t^{m-2}}{(m-2)!} & \frac{t^{m-1}}{(m-1)!} \\ 0 & 1 & t & \dots & \frac{t^{m-3}}{(m-3)!} & \frac{t^{m-2}}{(m-2)!} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & t & t^2/2 \\ 0 & 0 & 0 & \dots & 1 & t \\ 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix}$$

By properties of similar matrices we arrive to the

Corollary. See proof of the spectral theorem 2.19 on page 60-61 in Logemann Ryan.

For an $m \times m$ matrix A having just one eigenvalue of multiplicity m and only one linearly independent eigenvector v it follows the following expression for $\exp(At)$:

$$\exp(At) = V\exp(Jt)V^{-1} = V\left(e^{\lambda t}\sum_{k=0}^{m-1}\frac{t^k}{k!}(\mathcal{N})^k\right)V^{-1}$$

Remark.

If instead of the exponential function we like to calculate an arbitrary analytical function that has converging in $\mathbb{C}$ Maclorain series

$$f(z) = \sum_{k=0}^{\infty}\frac{f^{(k)}(0)}{k!}z^k$$

then the same reasoning and the Maclorain series for the function f lead to an expression for the matrix function $f(J)$

$$f(J) = \sum_{k=0}^{m-1}\frac{f^{(k)}(\lambda)}{k!}(\mathcal{N})^k \tag{24}$$

Theorem A.9 , on Jordan canonical form of matrix p. 268 in Logemann Ryan.

Let $A \in \mathbb{C}^{N\times N}$,. There is an invertible matrix $T \in \mathbb{C}^{N\times N}$ and an integer $k \in \mathbb{N}$ such that

$$J = T^{-1}AT$$

has the block diagonal structure

$$\mathbb{J} = \begin{bmatrix} J_1 & \mathbb{O} & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & J_2 & \mathbb{O} & \mathbb{O} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{O} & \mathbb{O} & \mathbb{O} & J_k \end{bmatrix}$$

where J_j has dimension $r_j \times r_j$ and is a Jordan block.

$$J_j = \begin{bmatrix} \lambda_j & 1 & 0 & ... & 0 & 0 \\ 0 & \lambda_j & 1 & ... & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & ... & 1 & 0 \\ 0 & 0 & 0 & ... & \lambda_j & 1 \\ 0 & 0 & 0 & ... & 0 & \lambda_j \end{bmatrix}$$

Furthermore, $\sum_{j=0}^{k} r_j = N$ and if $r_j = 1$ them $J_j = \lambda$ for some eigenvalue $\lambda \in \sigma(A)$. Every eigenvalue λ occurs at least at one block; the same λ can occur in more than one block. The number of blocks with the same eigenvalue λ on the diagonal is equal to the number of linearly independent eigenvectors corresponding to this eigenvalue λ (it's geometric multiplicity $g(\lambda)$).

Lecture 6
Summary of the main new material in Lecture 5.

1) The *block diagonal structure of standard matrix* for linear operators Ax having *invariant subspaces* V_1,...,V_s that decompose the whole space $\mathbb{R}^N$ into a direct sum: $\mathbb{R}^N(\mathbb{C}^N) = V_1 \oplus ... \oplus V_s$.

$$\forall x \in V_i \quad \Longrightarrow \quad Ax \in V_i$$

2) The standard matrix of the operator Ax having just one eigenvalue and just one linearly independent eigenvector v has a particularly simple structure in terms of the basis consistig of the chain of generalised eigenvectors associated with this eigenvector: Jordan block:

$$J_j = \begin{bmatrix} \lambda_j & 1 & 0 & ... & 0 & 0 \\ 0 & \lambda_j & 1 & ... & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & ... & 1 & 0 \\ 0 & 0 & 0 & ... & \lambda_j & 1 \\ 0 & 0 & 0 & ... & 0 & \lambda_j \end{bmatrix}$$

3) 1) and 2) together with the theorem about generalized eigenspaces for a matix $A \in \mathbb{C}^{N \times N}$ imply the theorem;

$$\mathbb{J} = T^{-1}AT$$

has the block diagonal structure

$$\mathbb{J} = \begin{bmatrix} J_1 & \mathbb{O} & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & J_2 & \mathbb{O} & \mathbb{O} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{O} & \mathbb{O} & \mathbb{O} & J_k \end{bmatrix}$$

where columns in the matrix T are chains of generalized eigenvectors corresponding to linearly indepenent eigenvectors.

4) Explicit formula for the exponent of a Jordan block and Jordan matrix:

$$\exp(Jt) = e^{\lambda t} \begin{bmatrix} 1 & t & t^2/2 & \dots & \frac{t^{m-2}}{(m-2)!} & \frac{t^{m-1}}{(m-1)!} \\ 0 & 1 & t & \dots & \frac{t^{m-3}}{(m-3)!} & \frac{t^{m-2}}{(m-2)!} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & t & t^2/2 \\ 0 & 0 & 0 & \dots & 1 & t \\ 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix}$$

Specification of detailes for Theorem A.9 with a sketch of the proof.

1) Our considerations about chains of generalised eigenvectors and the **special case of Theorem A.9** considered above imply that the matrix T in the general theorem A.9 on Jordan canonical form can be chosen in such a way that it's columns are elements from chains of generalised eigenvectors built on the maximal number of linearly independent eigenvectors to the matrix A.

2) The matrix $J = T^{-1}AT$ describing how operator Ax acts in terms of the basis of columns in the matrix T, has a block diagonal structure with one block corresponding to each linearly independent eigenvector. It follows from the fact that generalised eigenspaces are invariant with respect to the transformation A and from the fact that linear envelopes of the chains of generalised eigenvectors are linearly independent of each other and are also invariant with respect to A.

3) Each block corresponding to a particular eigenvector is a Jordan block with corresponding eigenvalue on diagonal, because of the special case of Theorem A.9 considered above. The size of a particular Jordan block in the Jordan canonical form depends on the length of the corresponding chain of generalised eigenvectors, that is the smallest integer r such that the equations $(A-\lambda I)^r v^{(r)} = 0$ and $(A-\lambda I)^{r-1} v^{(r)} \neq 0$ are satisfied.

4) It follows from the structure of the canonical Jordan form that the algebraic multiplicity $m(\lambda)$ of an eigenvalue λ is equal to the sum of sizes r_j of Jordan blocks corresponding to λ and coinsides with the dimension of it's generalised eigenspace $E(\lambda) = \ker\left((A-\lambda)^{m(\lambda)}\right)$.

Definition. An eigenvalue is called semisimple if it's generalised eigenspace consists only of eigenvectors and its algebraic multiplicity is equal to its geometric multiplicity: $m(\lambda) = r(\lambda)$. In this case corresponding the Jordan blocks will all have size 1×1.

Jordan blocks in the Jordan canonical form are unique but can be combined in various orders. The position of Jordan blocks within a canonical Jordan form depends on positions of the chains of generalised eigenvectors in the transformation matrix T and is not unique in this sense.

Example of calculating the Jordan canonical form of a matrix.

(Try to solve yourself exercises from the file with exercises on linear autonomous systems, where all answers and some solutions are given)

Consider matrix $C = \begin{bmatrix} 1 & -1 & -2 & 3 \\ 0 & 0 & -2 & 3 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix}$, Find its canonical Jordan's form and corresponding basis.

Find first the characteristic polynomial.

$$\det(C - \lambda I) = \det \begin{bmatrix} 1-\lambda & -1 & -2 & 3 \\ 0 & -\lambda & -2 & 3 \\ 0 & 1 & 1-\lambda & -1 \\ 0 & 0 & -1 & 2-\lambda \end{bmatrix} = (1-\lambda)\det \begin{bmatrix} -\lambda & -2 & 3 \\ 1 & 1-\lambda & -1 \\ 0 & -1 & 2-\lambda \end{bmatrix} =$$

$$(1-\lambda)(-\lambda)\det \begin{bmatrix} 1-\lambda & -1 \\ -1 & 2-\lambda \end{bmatrix} - (1-\lambda)\det \begin{bmatrix} -2 & 3 \\ -1 & 2-\lambda \end{bmatrix} =$$

$$(1-\lambda)(-\lambda)\left(\lambda^2 - 3\lambda + 1\right) - (1-\lambda)(2\lambda - 1) = (1-\lambda)\left(3\lambda^2 - \lambda - \lambda^3\right) + (1-\lambda)(1-2\lambda) =$$

$$(1-\lambda)\left(3\lambda^2 - 3\lambda - \lambda^3 + 1\right) = (1-\lambda)(1-\lambda)^3 = (1-\lambda)^4.$$

Matrix C has one eigenvalue $\lambda = 1$ with multiplicity 4. Consider the equation for eigenvectors $(C-I)x = 0$ with matrix

$$(C-\lambda I) = \begin{bmatrix} 0 & -1 & -2 & 3 \\ 0 & -1 & -2 & 3 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & -1 & 1 \end{bmatrix} \text{ Gauss elimination gives} \Longrightarrow \begin{bmatrix} 0 & -1 & -2 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 2 \\ 0 & 0 & -1 & 1 \end{bmatrix} \Longrightarrow \begin{bmatrix} 0 & -1 & -2 & 3 \\ 0 & 0 & -2 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

with *two free variables*: x_1 and x_4. Therefore the dimension of the eigenspace is 2. There are two linearly independent eigenvectors that can be chosen as

$v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$ and $v_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$. Each of these eigenvectors might generate a chain of generalised eigenvectors.

We check the equation $(C - \lambda I\ )v_1^{(1)} = v_1$ with extended matrix $\begin{bmatrix} 0 & -1 & -2 & 3 & 1 \\ 0 & -1 & -2 & 3 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & -1 & 1 & 0 \end{bmatrix}$ and carry out the same Gauss elimination as before: $\Longrightarrow \begin{bmatrix} 0 & -1 & -2 & 3 & 1 \\ 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & -2 & 2 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix}$. The second equation is not compatible and the system has no solution.

For the second eigenvector v_2 we solve similar system $(C - \lambda I\ )v_2^{(1)} = v_2$ with the extended matrix

$$\begin{bmatrix} 0 & -1 & -2 & 3 & 1 \\ 0 & -1 & -2 & 3 & 1 \\ 0 & 1 & 0 & -1 & 1 \\ 0 & 0 & -1 & 1 & 1 \end{bmatrix}$$

Gauss elimination implies the echelon matrix $\begin{bmatrix} 0 & -1 & -2 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 2 & 2 \\ 0 & 0 & -1 & 1 & 1 \end{bmatrix} \Longrightarrow \begin{bmatrix} 0 & -1 & -2 & 3 & 1 \\ 0 & 0 & -1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ that has a two-dimensional set of solutions. We choose one as $v_2^{(1)} = \begin{bmatrix} 1 \\ 1 \\ -1 \\ 0 \end{bmatrix}$ and build up the chain of generalized eigenvectors by solving one more equation $(C - \lambda I\ )v_2^{(2)} = v_2^{(1)}$ with the extended matrix $\begin{bmatrix} 0 & -1 & -2 & 3 & 1 \\ 0 & -1 & -2 & 3 & 1 \\ 0 & 1 & 0 & -1 & -1 \\ 0 & 0 & -1 & 1 & 0 \end{bmatrix} \Longrightarrow \begin{bmatrix} 0 & -1 & -2 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 2 & 0 \\ 0 & 0 & -1 & 1 & 0 \end{bmatrix} \Longrightarrow \begin{bmatrix} 0 & -1 & -2 & 3 & 1 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ leading to a generalized eigenvector (not unique)

$v_2^{(2)} = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}$. Finally we conclude that the Jordan canonic form of the matrix C in the basis v_1, v_2, $v_2^{(1)}$, $v_2^{(2)}$ is $J = T^{-1}CT = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$, with transformation matrix $T = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$, inverse:

$T^{-1} = \begin{bmatrix} 1 & 1 & 2 & -4 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 1 \\ 0 & -1 & -1 & 2 \end{bmatrix}$;

If we like to solve $x' = Cx$, with initial condition $x(0) = \xi$, we apply general formulas to

$$\xi = C_1 v_1 + C_2 v_2 + C_3 v_2^{(1)} + C_4 v_2^{(2)}$$

and find C_1, C_2, C_3, C_4 such that

We could use the general formula for solutions to the ODE as

$$x(t) = T\exp(J)T^{-1}\xi$$

$$\exp(Jt) = e^{\lambda t}\begin{bmatrix} 1 & t & t^2/2 & ... & \frac{t^{m-2}}{(m-2)!} & \frac{t^{m-1}}{(m-1)!} \\ 0 & 1 & t & ... & \frac{t^{m-3}}{(m-3)!} & \frac{t^{m-2}}{(m-2)!} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & ... & t & t^2/2 \\ 0 & 0 & 0 & ... & 1 & t \\ 0 & 0 & 0 & ... & 0 & 1 \end{bmatrix}$$

$$\exp\left(Ct\right) = \sum_{k=0}^{\infty} \frac{t^k}{k!} C^k = \begin{bmatrix} \exp\left(tJ_1\right) & \mathbb{O} & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \exp\left(tJ_2\right) & \mathbb{O} & \mathbb{O} \\ ... & ... & ... & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \mathbb{O} & \exp\left(tJ_r\right) \end{bmatrix}$$

6 Theorem about conditions for the exponential decay and for the boundedness of the norm $\|\exp(At)\|$ (Corollary 2.13)

Theorem.

Let $A \in \mathbb{C}^{N\times N}$ be a complex matrix (the real case $A \in \mathbb{R}^{N\times N}$ is included!). Let $\mu_A = \max\{\operatorname{Re}\lambda : \lambda \in \sigma(A)\}$ where $\sigma(A)$ is the set of all eigenvalues to A. μ_A is the maximal real part of all egenvalues to A.

Then three following statements are valid.

1. $\|\exp(At)\|$ decays exponentially if and only if $\mu_A < 0$. (It means that there are $M_\beta > 0$ and $\beta > 0$ such that $\|\exp(At)\| \le M_\beta e^{-\beta t}$)

2. $\lim_{t\to\infty} \|\exp(At)\xi\| = 0$ for every $\xi \in \mathbb{C}^N$ (it means that all solutions to the ODE $x' = Ax$ tend to zero) if and only if $\mu_A < 0$.

3. if $\mu_A = 0$ then $\sup_{t\ge 0} \|\exp(At)\| < \infty$ if and only if all purely imaginary eigenvalues and zero eigenvalues are semisimple meaning that $m(\lambda) = g(\lambda)$.

Remark. One can prove this theorem in two slightly different but essentially equivalent ways.

1) Using the similarity of the matrix A and it's Jordan matrix J

$$J = T^{-1}AT; \quad A = TJT^{-1}$$

corresponding expression of $\exp(At)$ in terms of $\exp(Jt)$ that is known explicitely:

$$\exp(At) = T\exp(Jt)T^{-1}$$

2) Using the expression for general solution to a linear autonomous system in terms of eigenvectors and generalized eigenvectors to A :

$$x(t) = \exp(At)x_0 = \sum_{j=1}^{s}\left(\left[\sum_{k=0}^{m_j-1}(A-\lambda_j I)^k \frac{t^k}{k!}\right] x^{0,j}e^{\lambda_j t}\right)$$

for solutions with initial data $x_0 = \sum_{j=1}^{s} x^{0,j}$ with $x^{0,j} \in E(\lambda_j)$ - components of x_0 in the generalized eigenspaces $E(\lambda_j) = \ker(A-\lambda_j)^{m_j}$ of the matrix A, where λ_j, $j = 1..., s$ are all distinct eigenvalues to A with algebraic multiplicities m_j.

The first method is shorter and more explicit.

In the course book the second method is used for proving Theorem 2.12 that is formulated in a slightly unfriendly style.

The Corollary 2.13 is almost equivalent and can be proven in exactly the same way as Theorem 2.12 but a bit simpler.

We give here a proof based on the expression $\exp(At) = T\exp(Jt)T^{-1}$ using Jordans canonical matrix J.

Proof.

Preliminary technical observations in the proof.

We point out that any matrix $A \in \mathbb{C}^{N\times N}$ can be represented with help of its Jordan matrix J as $A = TJT^{-1}$ where T is an invertible matrix with columns that are linearly independent eigenvectors and generalized eigenvectors to A ordered as in chains of generalised eigenvectors. The Jordan matrix J is a block diagonal matrix

$$J = \begin{bmatrix} J_1 & \mathbb{O} & \ldots & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & J_2 & \ldots & \mathbb{O} & \mathbb{O} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \mathbb{O} & \mathbb{O} & \ldots & J_{p-1} & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \ldots & \mathbb{O} & J_p \end{bmatrix}$$

where the number of blocks p is equal to the number of linearly independent eigenvectors to A. The symbol $\mathbb{O}$ denotes zero block.

Each Jordan block J_k has the structure as the following:

$$J_k = \begin{bmatrix} \lambda_i & 1 & 0 & 0 & 0 \\ 0 & \lambda_i & 1 & 0 & 0 \\ 0 & 0 & \lambda_i & 1 & 0 \\ 0 & 0 & 0 & \lambda_i & 1 \\ 0 & 0 & 0 & 0 & \lambda_i \end{bmatrix}$$

with possibly some blocks of size 1×1 being just one number λ_i . The sum of sizes of blocks is equal to N.

We use the expression

$$\exp(At) = T\exp(Jt)T^{-1}$$

that reduces analysis of the boundedness and limits of the norm $\|\exp(At)\|$ to the similar analysis for the matrix $\exp(Jt)$ because for two matrices A and B the estimate $\|AB\| \leq \|A\|\,\|B\|$ and therefore

$$\|\exp(At)\| \leq \|T\|\,\|T^{-1}\|\,\|\exp(Jt)\|$$

Similarly

$$T^{-1}\exp(At)\,T = \exp(Jt)$$

and

$$\begin{aligned} \|\exp(Jt)\| &\leq \|T\|\,\|T^{-1}\|\,\|\exp(At)\| \\ \|T\|^{-1}\,\|T^{-1}\|^{-1}\,\|\exp(Jt)\| &\leq \|\exp(At)\| \end{aligned}$$

For $\exp(Jt)$ we have the following explicit expression in terms of eigenvalues and their algebraic and geometric multiplicities:

$$\exp(Jt) = \begin{bmatrix} \exp(J_1 t) & \mathbb{O} & \mathbb{O} & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \exp(J_2 t) & \mathbb{O} & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \dots & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \mathbb{O} & \exp(J_{p-1} t) & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \mathbb{O} & \mathbb{O} & \exp(J_p t) \end{bmatrix} \tag{25}$$

where for example the block of size 5×5 looks as

$$\exp(J_k t) = \exp(\lambda_i t) \begin{bmatrix} 1 & t & \frac{t^2}{2!} & \frac{t^3}{3!} & \frac{t^4}{4!} \\ 0 & 1 & t & \frac{t^2}{2!} & \frac{t^3}{3!} \\ 0 & 0 & 1 & t & \frac{t^2}{2!} \\ 0 & 0 & 0 & 1 & t \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{26}$$

For a block of the size 1×1 we will get $\exp(J_k t) = \exp(\lambda_i t)$. If an eigenvalue λ_i is semisimple, that means it has the number of linearly independent eigenvectors (geometric multiplicity) $r(\lambda_i)$ equal to the algebraic multiplicity $m(\lambda_i)$ of λ_i. In this case all blocks corresponding to this eigenvalue and corresponding blocks in the exponent $\exp(Jt)$ all have size 1×1 and have this form $\exp(J_k t) = \exp(\lambda_i t)$.

Matrices $N \times N$ build a finite dimentional linear space with dimension $N \times N$. All norms in a finite dimensional space are equivalent. It means that for any two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ in the space of matrices, there are constants C_1, $C_2 > 0$ such that for any matrix A

$$C_1 \|A\|_1 \leq \|A\|_2 \leq C_2 \|A\|_1$$

It is easy to observe that the expression $\max_{i,j=1\dots N} |A_{ij}| = \|A\|_{\max}$ is a norm in the space of matrices and therefore can be used instead of the standart eucledian norm. There are constants B_1 and $B_2 > 0$ such that

$$B_1 \|A\|_{\max} \leq \|A\| \leq B_2 \|A\|_{\max}$$

I makes that to show the boundedness of the matrix norm $\|\exp(Jt)\|$ for $\exp(Jt)$, it is enough to show boundedness of all elements in $\exp(Jt)$. Similarly, to show that $\|\exp(Jt)\| \to 0$ when $t \to \infty$ it is enough to show that all elements in $\exp(Jt)$ go to zero when $t \to \infty$

To prove the statements in the theorem we need just to check how elements in the explicit expressions (26) for blocks in $\exp(Jt)$ *see (25), behave depending on the maximum of the real part of eigenvalues:* $\max\{\operatorname{Re}\lambda : \lambda \in \sigma(A)\}$ *and check situations when blocks of size* 1×1 *not including powers* t^p *can appear.*

- We observe in *(26)* that all elements in $\exp(Jt)$ have the form: $\exp(\lambda_i t)$ or $C\exp(\lambda_i t)t^p$ with some constants $C > 0$ and some $p > 0$ with possibly similar λ_i in different blocks.
- Absolute values of the elements in $\exp(Jt)$ have the form: $\exp((\operatorname{Re}\lambda_i)\, t)$ or $C\exp((\operatorname{Re}\lambda_i)\, t)t^p$ where all $\operatorname{Re}\lambda_i \leq \mu_A$. because $|\exp(i \operatorname{Im}\lambda_j)| = 1$ according to the Euler formula.

Proof of sufficiency of conditions in the theorem

We prove first sufficiency of the conditions in the statement **1.** for the formulated conclusions.

1. If $\mu_A < 0$ then maximum of absolute values of all elements $[\exp(Jt)]_{ij}$ in $\exp(Jt)$ satisfy the inequality

$$\max_{i,j} \left| [\exp(Jt)]_{ij} \right| \leq M \exp\left[(\mu_A + \delta)t\right] \underset{t\to\infty}{\longrightarrow} 0$$

and tends to zero exponentially for some constant $M > 0$ and δ so small that $-\beta = \mu_A + \delta < 0$. It follows because

$$\begin{aligned} \exp(\operatorname{Re}\lambda_i t)t^p &\leq \exp(\mu_A t)t^p = \exp\left[(\mu_A + \delta - \delta)\, t\right] t^p \\ &= \exp\left[(\mu_A + \delta)\, t\right] \underbrace{\left(t^p \exp\left[-\delta t\right]\right)}_{\leq M} \leq M \exp\left[-\beta t\right] \end{aligned}$$

Therefore $\|\exp(Jt)\| \leq M_\beta \exp\left[-\beta t\right] \underset{t\to\infty}{\longrightarrow} 0$ with another constant M_β and therefore $\|\exp(At)\| \leq (\|T\|\, \|T^{-1}\|\, M_\beta) \exp\left[-\beta t\right]$ decays exponentially.

Now we prove the sufficiency of the conditions in the statement **2.** for the formulated conclusion.

2. The definition of the matrix norm implies immediately that if $\mu_A < 0$ then by the result for the matrix norm $\|\exp(At)\|$ that $\lim_{t\to\infty} \|\exp(At)\xi\| \leq \|\xi\| \lim_{t\to\infty} \|\exp(At)\| = 0$ for every $\xi \in \mathbb{C}^N$.

Now we prove the sufficiency and necessity of the conditions in the statement **3.** for the uniform boundedness of the transition matrix $\exp(At)$: $\sup_{t\geq 0} \|\exp(At)\| < \infty$.

3. If $\mu_A = 0$ and then there are purely imaginary or zero eigenvalues λ. Then elements in the blocks of $\exp(Jt)$ corresponding to purely imaginary or zero eigenvalues will have the form $\exp(i \operatorname{Im}\lambda_i t)$ or $C\exp(i\operatorname{Im}\lambda_i t)t^p$. The absolute values of these elements will be 1 or Ct^p because $|\exp(i\operatorname{Im}\lambda_i t)| = 1$. Therefore absolute values of these elements will be bounded if and only if corresponding blocks are of size 1×1 and therefore elements Ct^p with powers of t are not present. This situation takes place if and only if purely imaginary and zero eigenvalues are **semisimple** (have geometric and algebraic multiplicities equal: $m(\lambda) = g(\lambda)$). Elements in $\exp(Jt)$ in the blocks corresponding to eigenvalues with negative real parts will be exponentially decreasing by the arguments in the proof of statement

Finish of the proof of sufficiency of conditions in the theorem.

■

Proof of necessity of conditions in the theorem

Finally we prove necessity of the conditions in the theorem

First for the Statement **1.** We observe that if $\mu_A = 0$ then referring to the analysis in **3.** absolute values of the elements corresponding to purely imaginary or zero λ_i in $\exp(Jt)$ are be bounded in the case if the conditions in **3.** are satified, or otherwise they have the form Ct^p and go to infinity when $t \to \infty$. Therefore the norm $\|\exp(At)\|$ does not decay exponentially in this case. If $\mu_A > 0$ the matrix $\exp(Jt)$ will include terms that are exponentially rising and the norm $\|\exp(At)\|$ can not decay exponentially in this case.

The nessecity of the conditions in the Statement **2** follows from the behaviour of the elements in $\exp(Jt)$ considered before or from the formula for general solution to the linear autonomous system.

Nessecity of the conditions in the Statement **3:**

The condition $\mu_A \geq 0$ means that there are eigenvalues λ with real part $\operatorname{Re}\lambda$ positive or zero. In the first case choosing vector ξ equal to a generalized eigenvector or an eigenvector corresponding to λ_i with $\operatorname{Re}\lambda_i > 0$ we get a solution $\exp(At)\xi$ represented as a sum with terms including exponents $\exp(\lambda_i t)$ such that $|\exp(\lambda_i t)| = |\exp(\operatorname{Re}\lambda_i t)| \to \infty$. In the second case there are eigenvalues $\lambda_i = i\operatorname{Im}\lambda_i$. Choosing ξ equal to one of corresponding generalized eigenvectors we obtain a solution $\exp(At)\xi$ represented as a sum including terms with constant absolute value or an absolute value that rises as some power t^p with $t \to \infty$. It implies the necessity of conditions in **2.** for having $\lim_{t\to\infty} \exp(At)\xi = 0$ for every $\xi \in \mathbb{C}^N$.■

The proof of the Corollary 2.13 in the book uses the explicit expression of solutions that we discussed at the beginning of this chapter of lecture notes and is a bit more complicated.

$$x\prime(t) = Ax(t) \quad \Longrightarrow x(t) = \exp(At)\xi$$

7

Lecture 7
Examples of phase portraits for linear autonomous ODEs in plane and calculations of matrix exponents.

8 Definition of stable equilibrium points.

Definition. A point $x_* \in G$ is called an equilibrium point to the equation $x' = f(x)$ if $f(x_*) = 0$.

The corresponding solution $x(t) \equiv x_*$ is called an equilibrium solution.

Definition. (5.1, p. 169, L.R.)

The equilibrium point x_* is said to be stable if, for any $\varepsilon > 0$, there is $\delta > 0$ such that, for any **maximal solution** $x : I \to G$ to the I.V.P.

$$\begin{aligned} x' &= f(x) \\ x(0) &= \xi \end{aligned}$$

such that $0 \in I$ and $\|x(0) - x_*\| \leq \delta$ we have $\|x(t) - x_*\| \leq \varepsilon$ for any $t \in I \cap \mathbb{R}_+$ for all "future times".

Below a picture is given in the case $x_* = 0$.

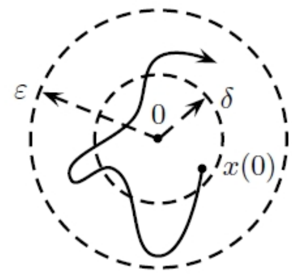


Figure 5.1 Stable equilibrium

Definition. (5.14, p. 182, L.R.)

The equilibrium point x_* of (**??**) is said to be *attractive* if there is $\delta > 0$ such that for every $\xi \in G$ with $\|\xi - x_*\| \leq \delta$ the following properties hold: the solution $x(t) = \varphi(t, \xi)$ to I.V.P. with $x(0) = \xi$ exists on $\mathbb{R}_+$and $\varphi(t, \xi) \to x_*$ as $t \to \infty$.

Definition. We say that the equilibrium x_* is **asymptotically stable** if it is both stable and attractive.

In the analysis of stability we will always choose a system of coordinates so that the origin coinsides with the equilibrium point. In the course book this agreement is applied even in the definition of stability.

Definition. The equilibrium point x_* is said to be *unstable* if it is not stable. It means that there is a $\varepsilon_0 > 0$, such that for any $\delta > 0$ there is point $x(0) : \|x(0) - x_*\| \leq \delta$ such that for some $t_0 \in I$ we have $\|x(t_0) - x_*\| > \varepsilon_0$.(a formal negation to the definition of stability).

9 Classification of phase portraits of autonomous linear systems in the plane.

Characteristic polynomial for a 2×2 matrix A is

$$p(\lambda) = \lambda^2 - \lambda \mathrm{Tr} A + \det A$$

Eigenvalues are:

$$\lambda_{1,2} = \frac{\mathrm{Tr} A}{2} \pm \sqrt{\frac{(\mathrm{Tr} A)^2}{4} - \det A}$$

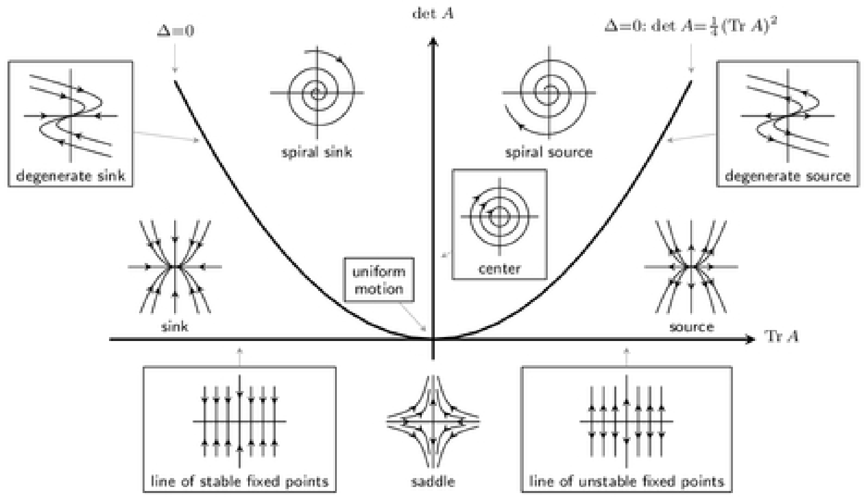
The line $\det A = \frac{(\mathrm{Tr} A)^2}{4}$ separates points in the plane $(\mathrm{Tr} A, \det A)$ corresponding to real and complex eigenvalues of the matrix A.

For $\mathrm{Tr} A, \det A$ in the first and second quadrants in the plane $(\mathrm{Tr} A, \det A)$ both $\operatorname{Re} \lambda_{1,2}$ are correspondingly positive and negative.

In the half plane where $\det A < 0$ eigenvalues $\lambda_{1,2}$ are real but have different signs.

These observations imply the following classification of phase portraits for linear autonomous systems in plane.

Poincaré Diagram: Classification of Phase Portraits in the $(\det A, \mathrm{Tr}\, A)$-plane



A classification of phase portraits for non-degenerate linear autonomous systems in plane in terms of the determinant and the trace of the matrix A.

Stable (unstable) nodes when eigenvalues λ_1, λ_2 are real, different, negative (positive). $\det(A) < \frac{1}{4}(tr(A))^2$; $\det(A) > 0$; $tr(A) < 0$, ($tr(A) > 0$).

Saddle (always unstable) when eigenvalues λ_1, λ_2 are real, with different signs. $\det(A) < 0$.

Stable (unstable) focus - spiral when λ_1, λ_2 are complex, with negative (positive) real parts. $\det(A) > \frac{1}{4}(tr(A))^2 \neq 0$, $tr(A) < 0$ ($tr(A) > 0$).

Stable (unstable) improper - degenerate node when eigenvalue λ_1 is real negative (positive) with multiplicity 2 having only one linearly independent eigenvector. $\det(A) = \frac{1}{4}(tr(A))^2$, $tr(A) < 0$ ($tr(A) > 0$).

Center (stable but not asymptotically stable) when λ_1, λ_2 are complex purely imaginary. $tr(A) = 0$; $\det(A) > 0$

Stable (unstable) star, when eigenvalue λ_1 is real negative (positive) with multiplicity 2 as for improper node, but having two linearly independent eigenvectors (diagonal matrix A)

$\square$

Example.

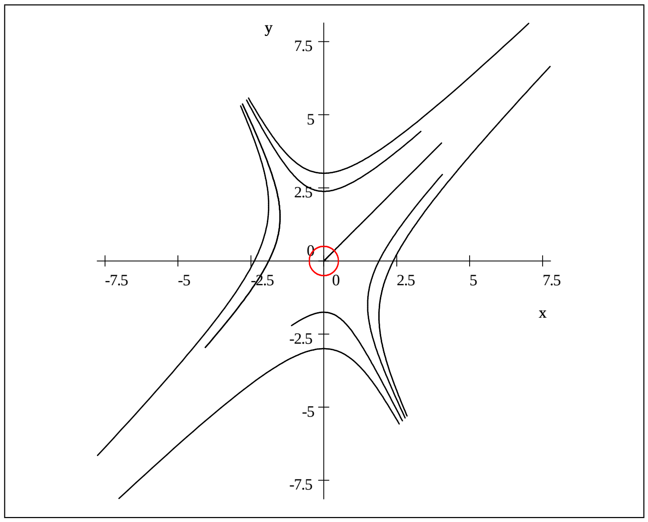
An example on instability: saddle point. There are trajectories (not all) that leave a neighbourhood $\|x\| < d$ of the origin for initial conditions ξ arbitrary close to the origin: for any $\varepsilon > 0$ and $0 < \|\xi\| \leq \varepsilon$ after some time T_ε.

$r' = Ar$ with $A = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}$, characteristic polynomial: $\lambda^2 - \lambda - 2 = 0$;

eigenvectors and eigenvalues are : $\left\{ \begin{bmatrix} 1 \\ -2 \end{bmatrix} \right\} \leftrightarrow \lambda_1 = -1, \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\} \leftrightarrow \lambda_2 = 2$. Eigenvectors satisfy homogeneous systems of equations with matrices $A - \lambda_1 I = \begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix}$ and $A - \lambda_2 I = \begin{bmatrix} -1 & 1 \\ 2 & -2 \end{bmatrix}$.

$r = C_1 e^{2t} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + C_2 e^{-t} \begin{bmatrix} 1 \\ -2 \end{bmatrix}$ - is the general solution. The equilibrium point in the origin is unstable.

Choosing a ball $\|x\| \le 1$, and for arbitrary $\varepsilon > 0$, $\xi = \varepsilon \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$, $\|\xi\|$ we see that the corresponding solution $x(t) = e^{2t}\varepsilon \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ will leave this ball $\|x\| \le 1$,after time $2T_\varepsilon = -\ln \varepsilon$.



Exercise.

Consider the following system of equations:

$$\begin{cases} x' = & 2y - x \\ y' = & 3x - 2y \end{cases}$$

1. a) can the system have a trajectory going from the point $(-a^2-1, -1)$ to the point $\underset{above\_the\_line}{(1, a^2+1)}$?

   b) which type of fixed point is the origin?

   c) draw a sketch of the phase portrait. **(4p)**

Solution

Matrix of the system is $A = \begin{bmatrix} -1 & 2 \\ 3 & -2 \end{bmatrix}$. Characteristic polynomila is $\det(A - \lambda I) = \det \begin{bmatrix} -1-\lambda & 2 \\ 3 & -2-\lambda \end{bmatrix}$ $\lambda^2 + 3\lambda - 4$. Eigenvalues and eigenvectors are: $\lambda_1 = -4, \lambda_2 = 1$.

Eigenvectors $v_1 = \begin{bmatrix} -2 \\ 3 \end{bmatrix} \leftrightarrow \lambda_1 = -4;$ satisfies the equation $(A - \lambda_1 I)\, v_1 = 0$ with $(A - \lambda_1 I) = \begin{bmatrix} 3 & 2 \\ 3 & 2 \end{bmatrix}$

$v_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \leftrightarrow \lambda_2 = 1$,satisfies the equation $(A - \lambda_2 I)\, v_2 = 0$ with $(A - \lambda_2 I) = \begin{bmatrix} -2 & 2 \\ 3 & -3 \end{bmatrix}$

$x(t) = C_1 \exp(-4t) v_1 + C_2 \exp(1t) v_2$

Origin is a saddle point and is unstable. Trajectories are hyperbolas asymptotically approaching with $t \to \infty$ or $t \to -\infty$ trajectories L_1, L_2, L_3, L_4, that are straight lines through the origin and are parallell to the eigenvectors.

Checking points $(-a^2-1,-1)$ and $(1,a^2+1)$ we observe that they are separated by the above mentioned straight trajectories L_1, L_2, L_3, L_4. Therefore no one trajectory can go between these two points because such a trajectory should cross one of L_1, L_2, L_3, L_4 that is impossible because of the uniquness of solutions to linear systems.■

Exercise 868. Exponent of a matrix with complex eigenvalues and phase portrait of the ODE with such matrix.

Calculate $\exp(A)$ for the matrix $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$; with eigenvalues $\pm i$.

We will consider first the general case.

Complex numbers in matrix form

The set of matrices of the structure $Z = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ *have the same properties with respect to matrix multiplication and addition as complex numbers of the form* $a + ib$.

In particular matrices of the form $\begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}$ behave as real numbers and matrix $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ behave as imaginary unit i.

We check that $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} = -\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = -I$ and $\begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}\begin{bmatrix} b & 0 \\ 0 & b \end{bmatrix} = \begin{bmatrix} ab & 0 \\ 0 & ab \end{bmatrix}$ and observe that the diagonal matrix $\begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}$ and the matrix $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ commute.

It makes that we can apply the Euler formula!!!!!!

$$\exp(a + ib) = \exp(a)(\cos(b) + i\sin(b))$$

for computing the exponent of a matrix of such structure:

$$\exp(Z) = \exp\left(\begin{bmatrix} a & -b \\ b & a \end{bmatrix}\right) = \exp\left(\begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} + \begin{bmatrix} 0 & -b \\ b & 0 \end{bmatrix}\right) = \exp\left(\begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}\right)\exp\left(\begin{bmatrix} 0 & -b \\ b & 0 \end{bmatrix}\right) =$$
$$\exp(a)I\left[\cos(b)\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \sin(b)\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\right] = \exp(a)\begin{bmatrix} \cos(b) & -\sin(b) \\ \sin(b) & \cos(b) \end{bmatrix}$$

$$\exp(tZ) = \exp(at)\begin{bmatrix} \cos(bt) & -\sin(bt) \\ \sin(bt) & \cos(bt) \end{bmatrix}$$

■

Corollary

Trajectories of the system of differential equations $x'(t) = Zx(t)$ with matrix $Z = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ are spirals

$$x(t) = \exp(at)\begin{bmatrix} \cos(bt) & -\sin(bt) \\ \sin(bt) & \cos(bt) \end{bmatrix}x(0)$$

build by a circular movement $\begin{bmatrix} \cos(bt) & -\sin(bt) \\ \sin(bt) & \cos(bt) \end{bmatrix}$ around the origin together with movement towards the origin if $a < 0$ and out from the origin if $a > 0$.

In the case when $a = 0$, trajectories go along circles around the origin.

□

General calculations imply immediately that $\exp(A) = \exp\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{bmatrix} \cos(1) & -\sin(1) \\ \sin(1) & \cos(1) \end{bmatrix}$

Lemma (important reduction result for the case of complex eigenvalues in plane)

For any real 2×2 matrix $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ with complex eigenalues $\lambda = \alpha \pm i\beta$ there is a non-degenerate matrix $M = \begin{bmatrix} a_{11} - \alpha & -\beta \\ a_{21} & 0 \end{bmatrix}$ such that $M^{-1}AM = \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix}$.

□

It implies that trajectories of the system with matrix A in this case will be also spirals, but squeezed. It the case if $\alpha = 0$ they will be ellipses instead of circles that were observed in the preious example.

Example of a stable but NOT asymptotically stable equilibrium point.

Consider the system $x'(t) = Ax(t)$ with $A = \begin{bmatrix} 0 & -2 \\ 2 & 0 \end{bmatrix}$. Eigenvalues of the matrix A are $\lambda = \pm 2i$ are purely imaginary (and non-zero). Therefore there are no other equilibrium points except the origin. The $\exp(At) = \begin{bmatrix} \cos(2t) & -\sin(2t) \\ \sin(2t) & \cos(2t) \end{bmatrix}$. The solution to the initial value problem with initial data $[\xi_1, \xi_2]^T$ is

$$\begin{aligned} x(t) &= \exp(At)\xi = \begin{bmatrix} \xi_1 \cos(2t) - \xi_2 \sin(2t) \\ \xi_1 \sin(2t) + \xi_2 \cos(2t) \end{bmatrix} = |\xi| \begin{bmatrix} \frac{\xi_1}{|\xi|} \cos(2t) - \frac{\xi_2}{|\xi|} \sin(2t) \\ \frac{\xi_1}{|\xi|} \sin(2t) + \frac{\xi_2}{|\xi|} \cos(2t) \end{bmatrix} = \\ &= |\xi| \begin{bmatrix} \cos(\theta)\cos(2t) - \sin(\theta)\sin(2t) \\ \cos(\theta)\sin(2t) + \sin(\theta)\cos(2t) \end{bmatrix} = |\xi| \begin{bmatrix} \cos(\theta + 2t) \\ \sin(\theta + 2t) \end{bmatrix} \end{aligned}$$

with $\cos(\theta) = \frac{\xi_1}{|\xi|}$. Therefore orbits of solutions are circles around the origin with the radius equal to $|\xi|$. It implies that the equlibrium point in the origin is stable. $\delta_\varepsilon > 0$ in the definition of stability can be chosen equal to $\varepsilon > 0$.■

Example. Two positie real eigenvalues. $\mathrm{Tr}(A) > 0$, $\det(A) < \frac{1}{4}(\mathrm{Tr}(A))^2$

Calculate $\exp(At)$ for the constant matrix $A = \begin{bmatrix} 3 & -1 \\ 2 & 0 \end{bmatrix}$ and sketch phase portrait for the system $x' = Ax$.

Solution.

The characteristic polynom for A is $\begin{bmatrix} 3 & -1 \\ 2 & 0 \end{bmatrix}$, $X^2 - 3X + 2 = (X - 1)(X - 2) = 0$, so eigenvalues are $\lambda_1 = 1$, $\lambda_2 = 2$. Eigenvectors are $v_1 = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\} \leftrightarrow \lambda_1$; $v_2 = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\} \leftrightarrow \lambda_2$

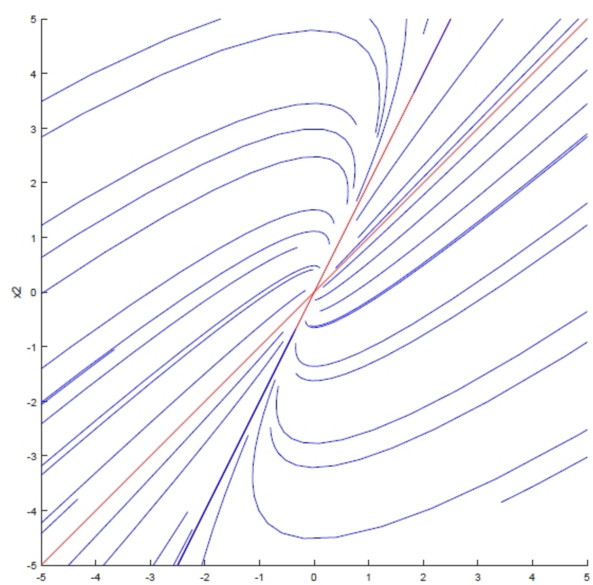
A direct formula following from diagonal Jordan form is the following: $\exp(At)$ as $\exp(At) = P \begin{bmatrix} e^t & 0 \\ 0 & e^{2t} \end{bmatrix} P^{-1}$, where the matrix P has columns of eigenvectors: $P = (v_1, v_2) = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix}$ and the inversion of P can be calculated by Cramer's formulas: $P^{-1} = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} -1 & 1 \\ 2 & -1 \end{bmatrix}$. We derive the final expression by multiplication of the three matrices:

$$\exp(At) = P \begin{bmatrix} e^t & 0 \\ 0 & e^{2t} \end{bmatrix} P^{-1} = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} e^t & 0 \\ 0 & e^{2t} \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 2 & -1 \end{bmatrix} = \begin{bmatrix} e^t & e^{2t} \\ 2e^t & e^{2t} \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 2 & -1 \end{bmatrix} = \begin{bmatrix} -e^t + 2e^{2t} & e^t - e^{2t} \\ -2e^t + 2e^{2t} & 2e^t - e^{2t} \end{bmatrix}$$

An alternative solution is based on using general solution to the differential equation $x' = Ax$:

$$x(t) = C_1 v_1 e^t + C_2 v_2 e^{2t}.$$

There are two positive eigenvalues to the matrix A. It corresponds to the phase portrait with unstable node (source), where red lines parallel to v_1 and v_2 correspond to solutions with one of coefficients C_1 or C_2 equal to zero.



Columns in $\exp(At)$ are solutions to the system above with initial data $e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

The plan is to find first the general solution, and then these two particular solutions.

To satisfy the initial data $x(0) = C_1 v_1 e^t + C_2 v_2 e^{2t} = e_1$

we solve a system of two equations for C_1 and C_2:

$C_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + C_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ or in matrix form $\begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

$\begin{bmatrix} -1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \implies C_1 = -1$ and $C_2 = 2$. Therefore the first columnt in $\exp(At)$ is: $-v_1 e^t + 2v_2 e^{2t} = \begin{bmatrix} -1 \\ -2 \end{bmatrix} e^t + \begin{bmatrix} 2 \\ 2 \end{bmatrix} e^{2t} = \begin{bmatrix} -e^t + 2e^{2t} \\ -2e^t + 2e^{2t} \end{bmatrix}$

Similarly we find the second column:

$C_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + C_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}; \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}; \begin{bmatrix} -1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

$\implies C_1 = 1$ and $C_2 = -1$.

The second column in $\exp(At)$ is: $v_1 e^t - v_2 e^{2t} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} e^t + \begin{bmatrix} -1 \\ -1 \end{bmatrix} e^{2t} = \begin{bmatrix} e^t - e^{2t} \\ 2e^t - e^{2t} \end{bmatrix}$

and finally $\exp(At) = \begin{bmatrix} -e^t + 2e^{2t} & e^t - e^{2t} \\ -2e^t + 2e^{2t} & 2e^t - e^{2t} \end{bmatrix}$

9.1 A general way to calculate exponents of matrices. (particularly useful for matrices having complex eigenvalues)

We use here general solution to the equation $x' = Ax$.
We clarify first in which way it can be used.

- For any matrix B the product Be_k gives the column k in the matrix B.
- Therefore the column k in $\exp(A)$ is the product $\exp(A)e_k$, where vector e_k is a standard basis vector, or colum with index k from the unit matrix I.
- On the other hand $\exp(At)\xi$ is a solution to the equation $x' = Ax$ with initial condition $x(0) = \xi$
- The expressions $x_k(t) = \exp(At)e_k$ is a solution to the equation $x' = Ax$ with initial condition $x(0) = e_k$
- Therefore the value of the solution in time $t = 1$: $x_k(1) = \exp(A)e_k$ gives the column k in the matrix $\exp(A)$
- Having the general solution for example in the case of dimension 3:

$$x(t) = C_1\Psi_1(t) + C_2\Psi_2(t) + C_3\Psi_3(t)$$

  in terms of linearly independent solutions $\Psi_1(t)$, $\Psi_2(t)$, $\Psi_3(t)$, we can for every k find a set of constants $C_{1,k}$,$C_{2,k}$,$C_{3,k}$, corresponding to each of the initial data e_k. Namely we solve equations $C_{1,k}\Psi_1(0) + C_{2,k}\Psi_2(0) + C_{3,k}\Psi_3(0) = e_k\ , \quad k = 1, 2, 3$

- that are equivalent to the matrix equation

$$[\Psi_1(0), \Psi_2(0), \Psi_3(0)] \begin{bmatrix} C_{1,1} & C_{1,2} & C_{1,3} \\ C_{2,1} & C_{2,2} & C_{2,3} \\ C_{3,1} & C_{3,2} & C_{3,3} \end{bmatrix} = [e_1, e_2, e_3] = I$$

- Values at $t = 1$ of corresponding solutions:

$$x_k(1) = C_{1,k}\Psi_1(1) + C_{2,k}\Psi_2(1) + C_{3,k}\Psi_3(1) = \exp(1 \cdot A)e_k$$

  will give us columns $\exp(1 \cdot A)e_k$ in $\exp(A)$.

- In the matrix form this result can be expressed as

$$\begin{bmatrix} C_{1,1} & C_{1,2} & C_{1,3} \\ C_{2,1} & C_{2,2} & C_{2,3} \\ C_{3,1} & C_{3,2} & C_{3,3} \end{bmatrix} = [\Psi_1(0), \Psi_2(0), \Psi_3(0)]^{-1}$$

$$\begin{aligned} \exp(A) &= [\Psi_1(1), \Psi_2(1), \Psi_3(1)] \begin{bmatrix} C_{1,1} & C_{1,2} & C_{1,3} \\ C_{2,1} & C_{2,2} & C_{2,3} \\ C_{3,1} & C_{3,2} & C_{3,3} \end{bmatrix} \\ &= [\Psi_1(1), \Psi_2(1), \Psi_3(1)] \, [\Psi_1(0), \Psi_2(0), \Psi_3(0)]^{-1} \end{aligned}$$

We demonstrate this idea using the result on the general solution from the problem 859.

We can calculate $\exp\left(\begin{bmatrix} 3 & -3 & 1 \\ 3 & -2 & 2 \\ -1 & 2 & 0 \end{bmatrix}\right)$, eigenvalues: $\lambda_1 = -1$, $\lambda_2 = 1 - i$, $\lambda_3 = 1 + i$
General solution to the system $x' = Ax$ is:

$$\begin{aligned} x(t) &= C_1\Psi_1(t) + C_2\Psi_2(t) + C_3\Psi_3(t) \\ &= C_1 e^{-t} \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} + C_2 e^t \begin{bmatrix} \cos t - \sin t \\ \cos t \\ \sin t \end{bmatrix} + C_3 e^t \begin{bmatrix} \cos t + \sin t \\ \sin t \\ -\cos t \end{bmatrix} \end{aligned}$$

introducing shorter notations for each term: $x(t) = C_1\Psi_1(t) + C_2\Psi_3(t) + C_3\Psi_3(t)$.
We calculate initial data for arbitrary solution by

$$x(0) = C_1\Psi_1(0) + C_2\Psi_3(0) + C_3\Psi_3(0) = C_1 \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} + C_2 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + C_3 \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

$$x(0)=[\Psi_1(0),\Psi_3(0),\Psi_3(0)]\begin{bmatrix}C_1\\C_2\\C_3\end{bmatrix}=\begin{bmatrix}1&1&1\\1&1&0\\-1&0&-1\end{bmatrix}\begin{bmatrix}C_1\\C_2\\C_3\end{bmatrix}$$

$\exp(A)$ has columns that are values of $x(1)$ for solutions that satisfy initial conditions $x(0)=e_1$, e_2,e_3 and therefore $\begin{bmatrix}1&1&1\\1&1&0\\-1&0&-1\end{bmatrix}\begin{bmatrix}C_{1,1}\\C_{2,1}\\C_{3,1}\end{bmatrix}=\begin{bmatrix}1\\0\\0\end{bmatrix}=e_1$; $\begin{bmatrix}1&1&1\\1&1&0\\-1&0&-1\end{bmatrix}\begin{bmatrix}C_{1,2}\\C_{2,2}\\C_{3,2}\end{bmatrix}=\begin{bmatrix}0\\1\\0\end{bmatrix}=e_2$; $\begin{bmatrix}1&1&1\\1&1&0\\-1&0&-1\end{bmatrix}\begin{bmatrix}C_{1,3}\\C_{2,3}\\C_{3,3}\end{bmatrix}=\begin{bmatrix}0\\0\\1\end{bmatrix}=e_3$;

We solve all three of these systems for $\begin{bmatrix}C_{1,k}\\C_{2,k}\\C_{3,k}\end{bmatrix}$ in one step as a matrix equation

$$\begin{bmatrix}1&1&1\\1&1&0\\-1&0&-1\end{bmatrix}\begin{bmatrix}C_{1,1}&C_{1,2}&C_{1,3}\\C_{2,1}&C_{2,2}&C_{2,3}\\C_{3,1}&C_{3,2}&C_{3,3}\end{bmatrix}=I$$

It is equivalent to the Gauss elimination of the following extended matrix: $\begin{bmatrix}1&1&1&1&0&0\\1&1&0&0&1&0\\-1&0&-1&0&0&1\end{bmatrix}$. The result at the rigth half will be the inverted matrix:

$$\begin{bmatrix}C_{1,1}&C_{1,2}&C_{1,3}\\C_{2,1}&C_{2,2}&C_{2,3}\\C_{3,1}&C_{3,2}&C_{3,3}\end{bmatrix}=\begin{bmatrix}1&1&1\\1&1&0\\-1&0&-1\end{bmatrix}^{-1}=\begin{bmatrix}-1&1&-1\\1&0&1\\1&-1&0\end{bmatrix}$$

It can also found by applying Cramer's rule.

We arrive to the expression of the matrix exponent by collecting these results through the matrix multiplication:

$$\exp(At)=[\Psi_1(t),\Psi_2(t),\Psi_3(t)]\begin{bmatrix}C_{1,1}&C_{1,2}&C_{1,3}\\C_{2,1}&C_{2,2}&C_{2,3}\\C_{3,1}&C_{3,2}&C_{3,3}\end{bmatrix}$$

$$\exp(At)=\begin{bmatrix}e^{-t}&e^t(\cos t-\sin t)&e^t(\cos t+\sin t)\\e^{-t}&e^t\cos t&e^t\sin t\\-e^{-t}&e^t\sin t&-e^t\cos t\end{bmatrix}\begin{bmatrix}-1&1&-1\\1&0&1\\1&-1&0\end{bmatrix}=$$

$$=\begin{bmatrix}e^t(\cos t+\sin t)-e^{-t}+e^t(\cos t-\sin t)&-e^t(\cos t+\sin t)+e^{-t}&-e^{-t}+e^t(\cos t-\sin t)\\(\cos t)e^t+(\sin t)e^t-e^{-t}&-(\sin t)e^t+e^{-t}&(\cos t)e^t-e^{-t}\\-(\cos t)e^t+(\sin t)e^t+e^{-t}&(\cos t)e^t-e^{-t}&(\sin t)e^t+e^{-t}\end{bmatrix}$$

and finally for $t=1$ we get $\exp(A)$

$$\exp(A) = e \begin{bmatrix} (\cos 1 + \sin 1) - e^{-2} + (\cos 1 - \sin 1) & -(\cos 1 + \sin 1) + e^{-2} & -e^{-2} + (\cos 1 - \sin 1) \\ (\cos 1) + (\sin 1) - e^{-2} & -(\sin 1) + e^{-2} & (\cos 1) - e^{-2} \\ -(\cos 1) + (\sin 1) + e^{-2} & (\cos 1) - e^{-2} & (\sin 1) + e^{-2} \end{bmatrix}$$