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 time variable t and 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:

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

Solving the equation (0) above together with the initial condition

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

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

One can reformulate the I.V.P. (1),(2) in the form of integral equation

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

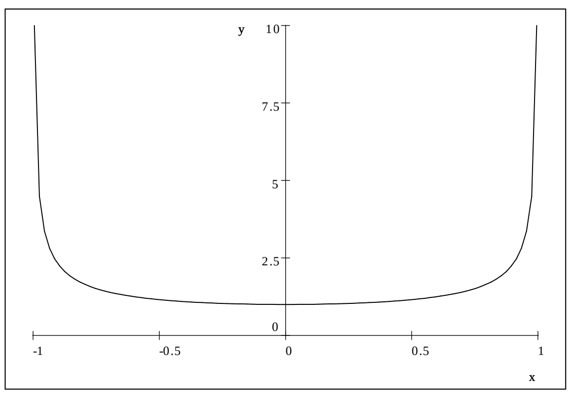
Continuous solutions to the integral equation (3) can be used as generalized solutions to (1),(2) in the case when $f\,(t, x)$ is only piecewise continuous with respect to t and therefore the integral in (3) does not have derivative in some isolated points. If f is continuous, then these two formulations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

Main types of problems posed for systems of ODEs

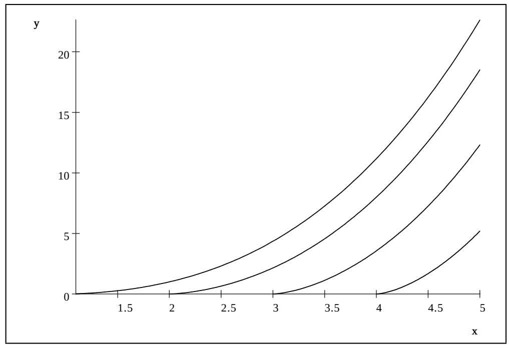
I) Existence and uniquness of solutions to I.V.P. Finding maximal interval of existence of solutions to I.V.P.

Example of bounded maximal interval. (Ex. 1.2, p.14, L.R.) I.V.P. $x'(t) = t \cdot x^3$, $x(0) = 1$. By separation of variables we arrive to a solution that exists only on a finite time interval $(-1, 1)$:

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



Example of non-uniqueness. (Ex.1.1, p.13, L.R.) I.V.P. $x'(t) = t \cdot x^{1/3}$, $t \in \mathbb{R}$, $x(0) = 0$. Constant solution $x(t) = 0$ exists. On the other hand for all $c > 0$ functions $x(t) = \frac{\left(t^2 - c^2\right)^{3/2}}{(3)^{3/2}}$, $t \geq c$ are also solutions to the equation. See the calculation below. 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.

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

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

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

II) One can for particular classes of equations pose the problem of finding a reasonable analytical description of all solutions to the above equation (a 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: $x(t) = x(t + T)$.

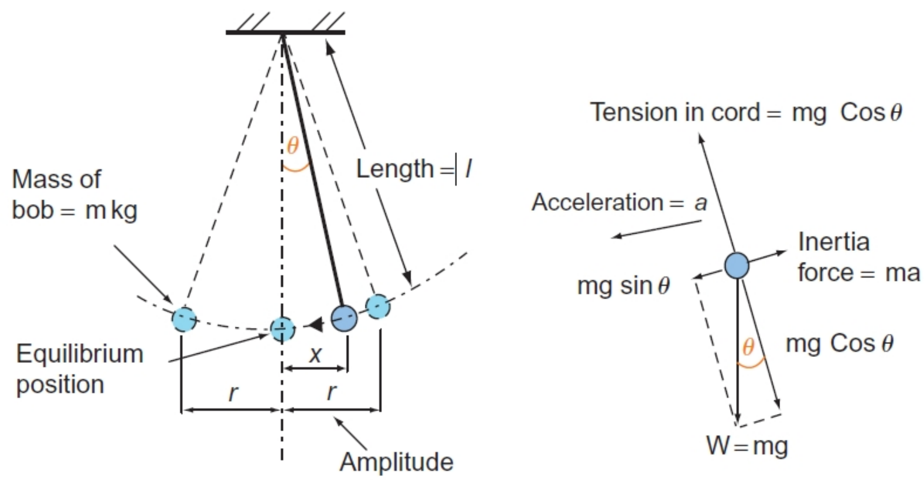
IV) Find how solutions $x(t)$ behave in the vicinity of an equilibrium point η with $t \to \infty$: it is interesting if they stay close to η, or go out of η with time $t \to \infty$. (stability or instability of equilibrium points).

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 calle call such a picture *phase portrait.*

VI) Describe geometric properties of so called limit sets, or attractors: sets that solutions $x(t)$ approach when $t \to \infty$.

Examples

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

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

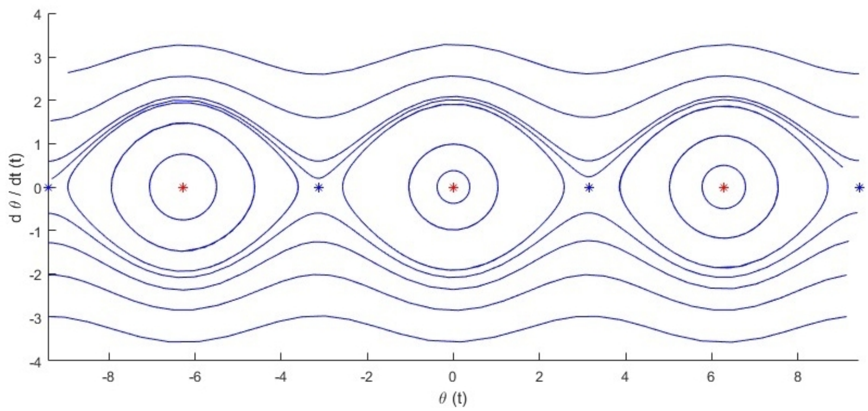
Rewrite the second order equation as a system of two equations for $x_1(t) = \theta(t)$ and $x_2(t) = \theta'(t)$:

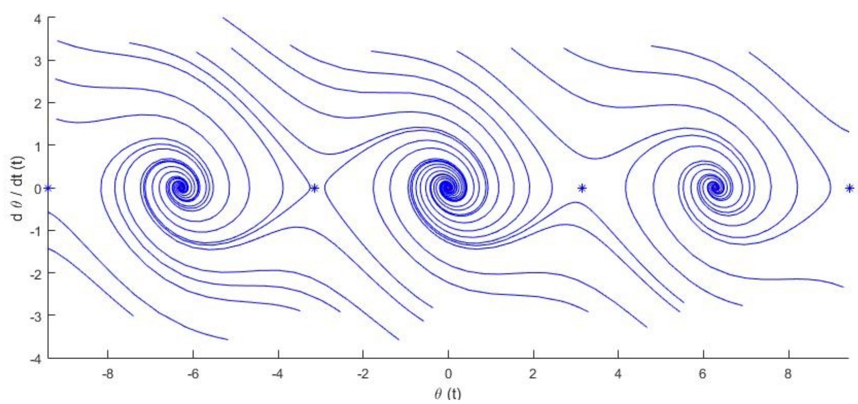
$$\begin{aligned} x_1'(t) &= x_2(t) \\ x_2'(t) &= -\frac{\gamma}{m}x_2(t) - \frac{g}{l}\sin(x_1(t)) \end{aligned}$$

-this equation cannot be solved analytically.

Phase portrait.

The picture of trajectories 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.**



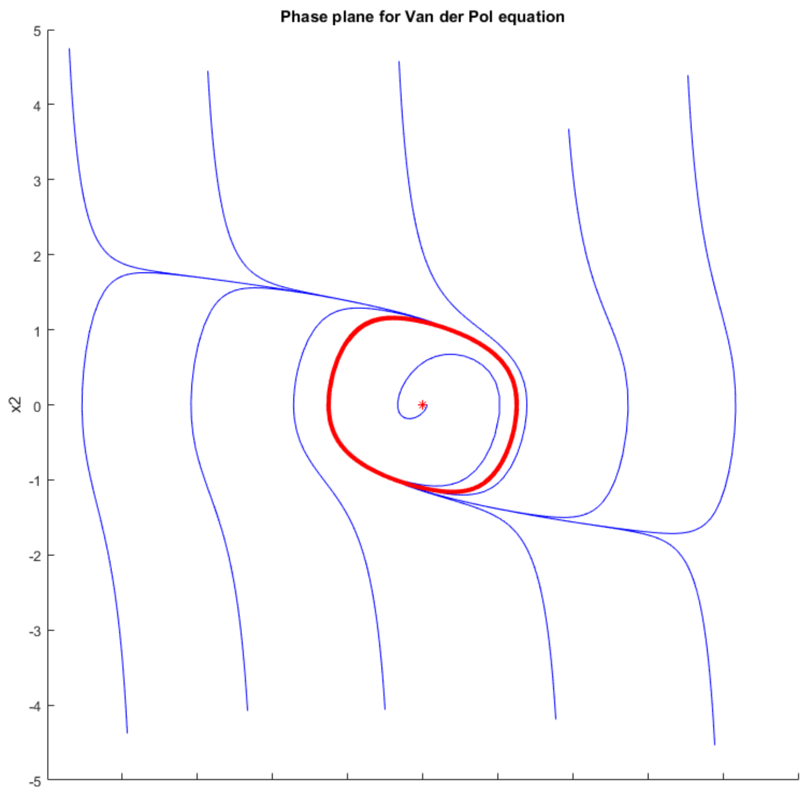


Points $\theta = 0 + 2\pi k$, $\theta' = 0$ and $\theta = \pi + 2\pi k$, $\theta' = 0$ on the first picture are equilibrium points. One can see closed orbits around equilibrium points $\theta = 2\pi k$, $\theta' = 0$, corresponding to periodic solutions. Points $\theta = \pi + 2\pi k$, $\theta' = 0$ 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 can be described by a non-linear equation.

In the case with friction 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.

Example from circuit theory. 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 closed curve (depicted in red) that seems to be an orbit corresponding to a periodic solution.

2 Linear autonomous systems of ODE

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 initial value problem (I.V.P.) with initial condition

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

We point out that all general results about linea systems of ODE are also valid in the case of the complex vector space $x \in \mathbb{C}^n$, $\xi \in \mathbb{C}^n$ and 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 of all solutions and to connect their qualitative properties with specific properties of the matrix A, its eigenvalues and eigenvectors together with more subtle spectral properties such as subspaces of generalised eigenvectors.

2.1 The space of solutions

We make first two simple observations that are valid even for general non-autonomous systems

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

with a matrix $A(t)$ that is not constant but is a continuous function of time on the interval J.

Lemma. The sets of solutions $\mathcal{S}_{\text{hom}}$ to (4), and to (6) are linear vector spaces.

Proof. $\mathcal{S}_{\text{hom}}$ includes 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:

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

2.2 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) has a solution $x(t)$ on an interval including τ. Consider the case when $\tau \leq t$

We can write an equivalent integral equation for $x(t)$.

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

The triangle inequality for the integral and the definition of the matrix norm imply that

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

Introducing the notation $G(t) = \|\xi\| + \int_{\tau}^{t} \|A\| \, \|x(\sigma)\| \, d\sigma$ 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 $\exp(- \|A\| \, t)$ we arrive to

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

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

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

that implies the **Grönwall inequality** in this simple case:

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

(Knowlege of this proof is required at the exam)

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

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 $z(t)=0$ and therefore the uniqueness of solution to I.V.P. (4),(5).■

2.3 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 solution space.

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

$$e^{A(t-\tau)} = I + A(t-\tau) + \frac{1}{2}A^2(t-\tau)^2 + \ldots + \frac{1}{k!}A^k(t-\tau)^k \ldots = \sum_{k=0}^{\infty} \frac{1}{k!}A^k(t-\tau)^k$$

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

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). One can try to solve this integral equation by iterations:

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

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. This series converges uniformly on any finite time interval by the Weierstrass criterion because the norm of each of it's terms is estimated by a term from a convergent number series

$$\sum_{k=0}^{\infty} \frac{1}{k!} \|A\|^k (t-\tau)^k$$

for the exponential function $\exp(\|A\|(t-\tau))$. Here one uses the estimate for the norm of the product of two matrices: $\|AB\| \leq \|A\| \, \|B\|$.

It 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.

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 exponential expression for the the unique solutions to (1) has a disadvantage that $\exp(At)$ is not possible to calculate analytically.

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 will soon show that these columns are linearly independent and build a basis in the space of solutions.

We collect some 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

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

(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.

$$
\begin{aligned}
e^{P+Q} &= \sum_{k=0}^{\infty} \frac{1}{k!}(P+Q)^k = \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{m}^{k} \binom{k}{m} P^k Q^{m-k} \\
&= \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{m=0}^{k} \left(\frac{k!}{m!\,(k-m)!} \right) P^m Q^{k-m} \\
&= \sum_{k=0}^{\infty} \sum_{m+p=k} \frac{P^m}{m!} \frac{Q^p}{p!} = \left(\sum_{m=0}^{\infty} \frac{1}{m!} P^m \right) \left(\sum_{p=0}^{\infty} \frac{1}{p!} Q^p \right) = e^P e^Q
\end{aligned}
$$

(3) Can be proved in three different ways. It follows from the definition of $\exp(At)$ by element-wise 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 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

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

■

2.4 The dimension of the space $\mathcal{S}_{\text{hom}}$ of solutions

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) 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).

(Knowlege of the proof is required at the exam)

Hint to 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 $\sigma \in \mathbb{R}$: $l(\sigma) = \sum_{j=1}^{N} \alpha_j y_j(\sigma) = 0$. The trivial constant zero solution coinsides with l at this time point.

But by the uniqueness of solutions to (4) it implies that $l(t)$ must coinside with the trivial zero solution for all times and in particular at time $t = \tau$. Therefore $\sum_{j=1}^{N} \alpha_j b_j = 0$. It implies that all coefficients $\alpha_j = 0$ because $b_1, ..., b_N$ are linearly independent vectors in $\mathbb{R}^N$(or $\mathbb{C}^N$). Therefore $y_1(t), ..., y_N(t)$ are linearly independent for all $t \in \mathbb{R}$ by definition. Arbitrary initial data $x(\tau) = \xi$ 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 $y_1(t), ..., y_N(t)$.

$$x(t) = \exp(A(t-\tau))\xi = \exp(A(t-\tau)) \sum_{j=1}^{N} C_j b_j = \sum_{j=1}^{N} C_j y_j(t)$$

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.$\blacksquare$

2.5 Analytic solutions. Case with a basis of eigenvec-

tors.

An idea that leads to an analytical solution is 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 $y_k(t) = \exp(A(t-\tau))v_k$ 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

$$Av = \lambda v$$

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

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

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

except $k = 0$:

$$\begin{aligned} e^{At}v &= \exp\left(\lambda t I + (At - \lambda t I)\right)v = \exp\left(\lambda t I\right)\exp\left(\left(A - \lambda I\right)t\right)v = \qquad (10)\\ &= e^{\lambda t}\sum_{k=0}^{\infty}\frac{1}{k!}t^k\left(A - \lambda I\right)^k v = e^{\lambda t}v. \end{aligned}$$

This observation leads to a simple conclusion that if the matrix A has n linearly independent eigenvectors $\{v_k\}$, then any solution to (4) 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 $\left\{e^{\lambda_k t} v_k\right\}$ building a basis for the space of solutions to (4). 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 an example soon.

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 that eigenvalues are different). Solutions

$y_1(t) = e^1 v_1$ and $y_2(t) = e^{-1} v_2$ are linearly independent.

Arbitrary real solution to the system of ODEs has the form

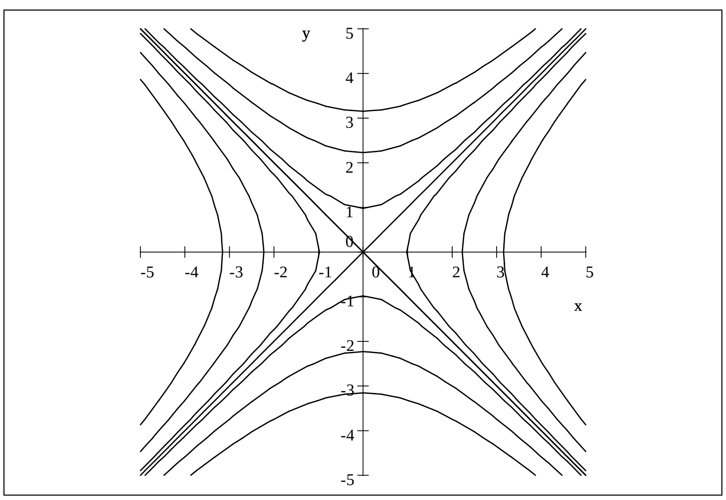
$$x(t) = C_1 y_1(t) + C_2 y_2(t) = C_1 e^t \begin{bmatrix} 1 \\ 1 \end{bmatrix} + C_2 e^{-t} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

with arbitrary coefficients C_1 and C_2.Corresponding phase portrait will include particular solutions tending to infinity along the vector $v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, 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

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

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 matrix A 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.

Example 2.

$x' = Ax$ 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 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.■

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

They are complex conjugate:

$$\overline{\lambda_1} = \lambda_2$$

because the characteriscic polynomial has real coefficients.

Eigenvectors satisfy the equations $\begin{bmatrix} 2+2i & -2 \\ 4 & -2+2i \end{bmatrix} v_1 = 0$ 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 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 (determinant is zero).

We solve the first equation in the first system by choosing the first component 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 $v_2 = \left\{ \begin{bmatrix} 1 \\ 1-i \end{bmatrix} \right\} \leftrightarrow \lambda_2 = 1 + 2i$.

They are linear independent as corresponding to different eigenvalues.

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

$x^*(t)$ and $y^*(t)$ are linear 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 $y^*(t)$): $\operatorname{Re}[x^*(t)] = \frac{1}{2}(x^*(t) + y^*(t))$ and $\operatorname{Im}[x^*(t)] = \frac{1}{2i}(x^*(t) - y^*(t))$ 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

$$[x^*(t), y^*(t)]\, T = [\operatorname{Re}[x^*(t)], \operatorname{Im}[x^*(t)]]$$

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

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

and $\operatorname{Re}\left[x^*(t)\right]$ and $\operatorname{Im}\left[x^*(t)\right]$ are linearly independent.

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

$$x(t) = C_1 \operatorname{Re}\left[x^*(t)\right] + C_2 \operatorname{Im}\left[x^*(t)\right].$$

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

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

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

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.

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

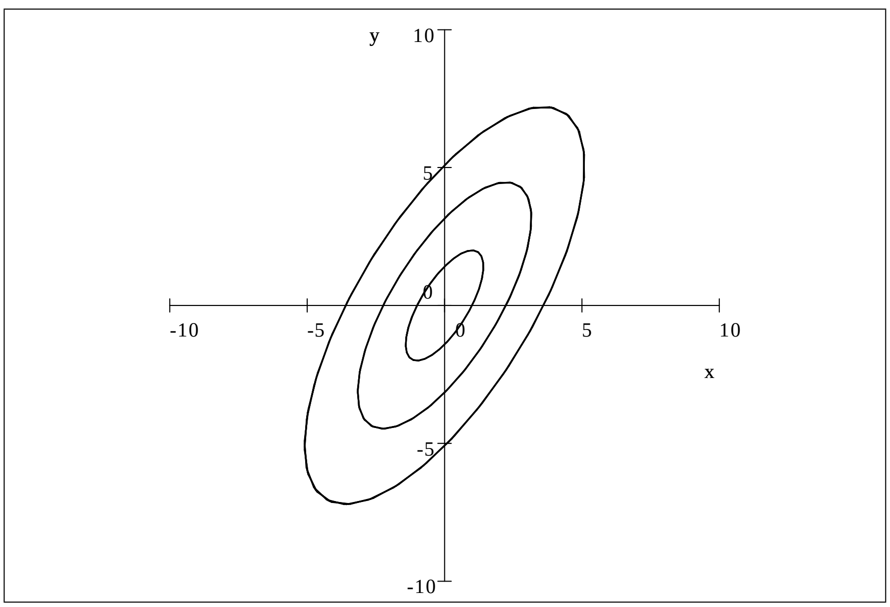
Similarly

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

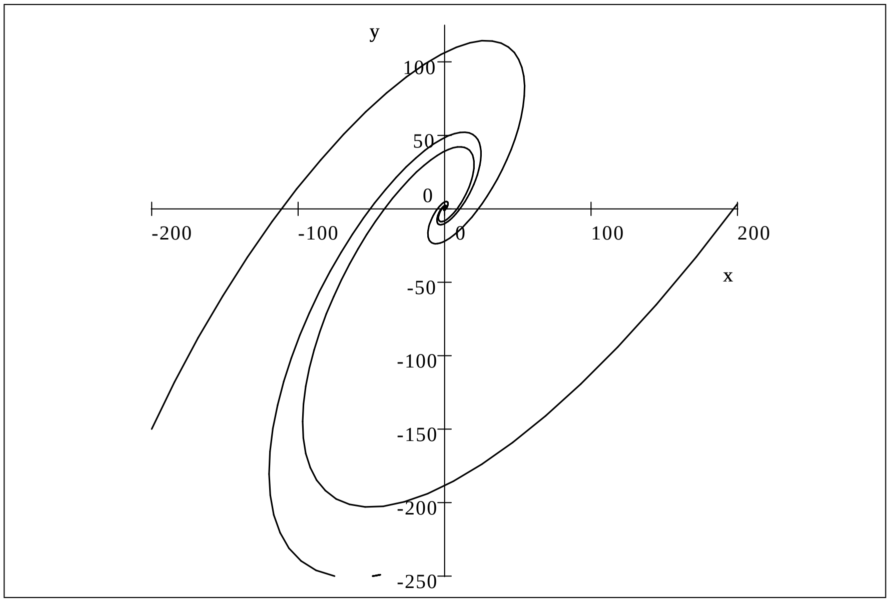
Finally we arrive to a parametric expression for a periodic movement along ellipses with size depending on C_1 and C_2.

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

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.



One can also see the direction of spirals by checking the direction of the vectorfield on one of coordinate axises directly from the equation.

3 Generalised eigenvectors and eigenspaces.

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

Example 3

$$\left\{ \begin{array}{r} x_1' = -x_1 \\ x_2' = x_1 - x_2 \end{array} \right. \quad \text{or} \quad x'(t) = Ax \text{ with } A = \begin{bmatrix} -1 & 0 \\ 1 & -1 \end{bmatrix}$$

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

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

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

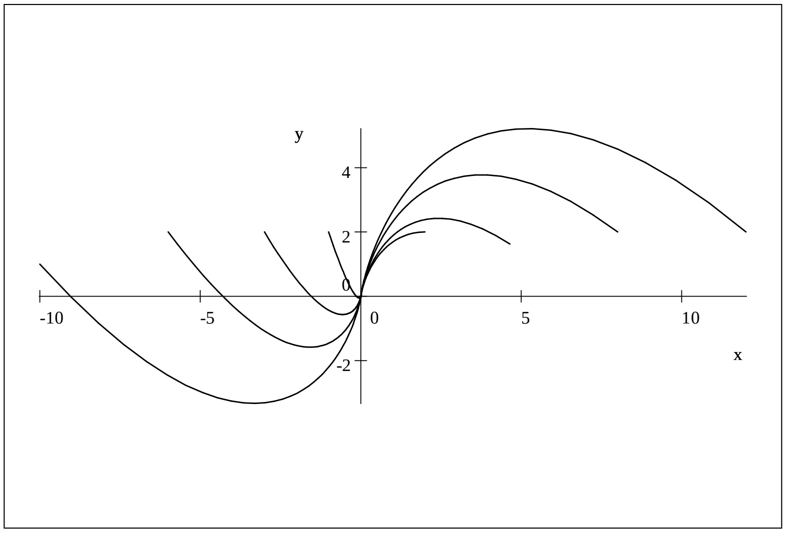
is the solution to the first equation, substitute it into the second equation and solveit explicitely with respect to $x_2(t)$:

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

Therefore the general solution to this particular system has the form

$$\begin{aligned}
x(t) &= C_1 e^{-t} \begin{bmatrix} 1 \\ t \end{bmatrix} + C_2 e^{-t} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = C_1 e^{-t} \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right) + C_2 e^{-t} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\
&= C_1 e^{-t} \left(v^{(1)} + tv\right) + C_2 e^{-t} v
\end{aligned}$$

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 is $x(0) = C_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + C_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = C_1 v^{(1)} + C_2 v$.

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 before 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 $(A - \lambda I)^2 v^{(1)} = 0$. Such vectors are called generalised eigenvectors to A.It shows that the initial data in this explicit solution are represented as a linear combination of an eigenvector and a generalised eigenvector.

We point out that the general solution we have got could be derived by applying the same idea as in the formula (10) but to the generalised eigenvector $v^{(1)}$:

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

That 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 with arbitrary constant matrix.

Definition.

A non-zero vector $z\in\mathbb{C}^N$ is called a generalised eigenvector to the matrix A corresponding to the eigenvalue λ with the algebraic multiplicity $m(\lambda)$ if $\left(A-\lambda I\right)^{m(\lambda)}z=0$. If $\left(A-\lambda I\right)^{r}z=0$ and $\left(A-\lambda I\right)^{r-1}z\neq 0$ 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 $\left(A-\lambda I\right)u=0$.

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

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 taking $z\in E(\lambda)$ such that $\left(A-\lambda I\right)^{m(\lambda)}z=0$ and calculating $\left(A-\lambda I\right)^{m(\lambda)}Az=\left(A-\lambda I\right)^{m(\lambda)}Az-\lambda\left(A-\lambda I\right)^{m(\lambda)}z=$

$=\left(A-\lambda I\right)^{m(\lambda)}\left(A-\lambda I\right)z=\left(A-\lambda I\right)\left(A-\lambda I\right)^{m(\lambda)}z=0$, the last equality valid because $\left(A-\lambda I\right)$ and $\left(A-\lambda I\right)^{m(\lambda)}$ 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

$$\exp(At)z = \sum_{k=0}^{\infty} \frac{1}{k!} t^k A^k z$$

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 linear combination (infinite one) of elements from the generalized eigenspace $E(\lambda)$ that is a vector space. 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 can be expressed in such a way that it would include only finite number of terms and can be calculated analytically.

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 $z \in E(\lambda)$ denote by $x_z(t) = \exp(At)z$ - the solution of I.V.P. with $x_z(0) = z$. Then

1) for $\lambda \in \sigma(A)$ and $z \in E(\lambda)$ a generalised eigenvector

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

Proof.

We show it by the following simple calculations:

$$\begin{aligned} x_z(t) &= \exp(At)z = \exp(t\lambda I)\exp(A - \lambda I)t = (11) \\ e^{\lambda t} \sum_{k=0}^{\infty} \frac{t^k}{k!} (A - \lambda I)^k z e^{\lambda t} &= \sum_{k=0}^{m(\lambda)-1} \frac{t^k}{k!} (A - \lambda I)^k z \end{aligned}$$

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. ■

3.1 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), 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 just 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.

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, $\sum m_j = N$. Then $\mathbb{C}^N$ can be represented as a direct sum of generalised eigenspaces $\ker(A-\lambda_j)^{m_j}$ to A having dimensions m_j:

$$\mathbb{C}^N = \ker(A-\lambda_j)^{m_1} \oplus ... \oplus \ker(A-\lambda_j)^{m_s} \tag{12}$$

The formula (11) 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 (4) and a representation of general solutions 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 $B(\lambda)$ be a basis in $E(\lambda)$ and denote $\mathcal{B} = \cup_{\lambda\in\sigma(A)} B(\lambda)$ - the union of all bases of generalized eigenspaces $E(\lambda)$ for all eigenvalues $\lambda \in \sigma(A)$. Then the set of functions $\{x_z : z \in \mathcal{B}\}$ is a basis of solution space $\mathcal{S}_{\text{hom}}$ of (4).

Proof. Part 2) of the theorem follows from the generalized eigenspace decomposition theorem, that implies that $\mathcal{B}$ is a basis in $\mathbb{C}^N$ and from the theorem on the dimension of the solution space $\mathcal{S}_{\text{hom}}$ of a linear system. ■

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 initial data $x(0) = \xi$ for the solution $x(t)$ as a sum of its components from different generalised eigenspaces:

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

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 $x_z(t) \in E(\lambda_j)$ for all $t \in \mathbb{R}$.

Let m_j be the algebraic multiplicity of the eigenvalue λ_j. We apply the formula (11) to this representation and derive the an expression for solutions for arbitrary initial data as a finite sum (instead of series):

$$x(t) = e^{At}x_0 = \sum_{j=1}^{s} \left(e^{\lambda_j t} \left[\sum_{k=0}^{m_j - 1} (A - \lambda_j I)^k \frac{t^k}{k!} \right] x^{0,j} \right) \tag{13}$$

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

$$x^{0,j} = ...C_p v_j + C_{p+1} v_j^{(1)} + C_{p+2} v_j^{(2)} ... \tag{14}$$

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 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 eigenvectors satisfy the equation $(A - I)\, v = 0$: $A - I = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

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

$(A - \lambda I) v_1^{(1)} = v_1$. 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. If we try to extend the chain of generalised eigenvectors with one more: $v_2^{(2)}$by solving the system $(A - I)v_2^{(2)} = v_2^{(1)}$

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

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).

We can write general solution to the system of ODE with matrix A using the general formula (13) and expressing the initial data as a linear combination of eigenvectros v_1 and v_2 and the generalised eigenvector $v_1^{(1)}$:

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

The expression above simplifies (using that by construction $(A-\lambda I)v_1^{(1)} = v_1$) to

$$\begin{aligned} x(t) &= C_1e^t v_1 + C_2e^t v_2 + C_3e^t\left[I + (A-I)\,t\right]v_1^{(1)} \\ & C_1e^t v_1 + C_2e^t v_2 + C_3e^t v_1^{(1)} + C_3te^t v_1 \end{aligned}$$

Example 5. Matrix 3x3 with one 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.

The eigenvectors satisfy the equation $(A+I)\,v = 0$: $A - I = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{bmatrix}$.

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

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

It 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 $v^{(2)}$ in the chain must satisfy the equation $(A-\lambda I)v^{(2)} = v^{(1)}$:

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

$$y_3 = 1/2,\ y_2 = 0,\ y_1 = 0. \quad v^{(2)} = \begin{bmatrix} 0 \\ 0 \\ 1/2 \end{bmatrix}.$$

3.2 Chains of generalised eigenvectors

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

We find a basis of the eigenspace to λ consisting of $r(\lambda)$ eigenvectors satisfying the equation $(A - \lambda I)\, u_0 = 0$.Their number $r(\lambda)$ is called geometric multiplicity of λ and $r(\lambda) \le m(\lambda)$. Then for each eigenvector $u_0 \ne 0$ from this basis we find a vector $u_1 \ne 0$ satisfying the equation $(A - \lambda I)\, u_1 = u_0$, and continue this calculation, building a chain of generalised eigenvectors $u_1, ..., u_l$ satisfying equations

$$(A - \lambda I)\, u_k = u_{k-1} \tag{15}$$

up to the index $k = l$ when there will be no solutions to the next equation. The largest possible number l can be $m(\lambda) - r(\lambda) - 1$, but it can also be smaller if the eigenvalue λ has more than one linearly independent eigenvector.

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 (15) so that they belong to the column space of the matrix $(A - \lambda I)$ (if possible!) to guarantee that the equations (15) have a solution.

Alternatively one can start this algorithm from above, solving first the equation

$$\begin{aligned} (A-\lambda I)^l u_l &= 0 \\ (A-\lambda I)^{l-1} u_l &\neq 0 \end{aligned}$$

for a generalized eigenvector of rank l and then can apply equations (15) 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. Check the solution to the exercise 864 where these observations are important.

Lemma. The chain of generalised eigenvectors constructed in such a way is linearly independent. (**Exercise!**)

Theorem. The set of generalise eigenvectors of rank r is linearly independent if and only if the eigenvectors in corresponding chains of generalised eigenvectors are linearly independent.

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

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

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

To find a basis in the generalized eigenspace $E(\lambda_j)$ one can start with finding all linearly independent eigenvectors that are linearly independent solutions to the equation $(A-\lambda_j I)\,v = 0$ and collecting them in a set denoted by $\mathcal{E}$. Then find all linearly independent solutions to $(A-\lambda_j I)^2 v^{(1)} = 0$ (that are not eigenvectors) and adding them $\mathcal{E}$. 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 derived linearly independent generalised eigenvectors will be equal to m_j

- the albebraic multiplicity of the eigenvalue λ_j.

A 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 above:

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

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 (14) 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 (14)

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

We considered 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 (do not need to do it in fact). Then we construct arbitrary complex solutions in the invariant generalized eigenspace corresponding to the first of these conjugate eigenvalues. The real and imaginary parts of these solutions are linearly independent and build a basis of solutions in the corresponding real invariant subspace.

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

4 Jordan canonical form of matrix. Functions of matrices.

4.1 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 C^N or in 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 solution as a linear combination in term of this basis. 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 vectors.

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

$$y'(t) = V^{-1}AVy, \quad y(0) = V^{-1}\xi$$

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

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

They have several characteristics the same: determinant, 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 $B = V^{-1}AB$, $A =$

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 ther tems all zero. For example:

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

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.

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

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:

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

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 $A : B = U^{-1}AU$, 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 evident 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 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. 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.

4.2 Jordan canonical form of matrix and its functions.

We will observe now that a basis of generalised eigenvectors build with help of chains of generalised eigenvectors as we discussed before, leads to a particularly canonical matrix J similar to the matrix A by the transformation $V^{-1}AV = J$ or $A = VJV^{-1}$ with the tmatrix

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

where columns are generalised eigenvectors from different chains of generalised eigenvectors corresponding to linearly independent eigenvectors put in the same order as in (16).

Consider first an $m \times m$ matrix A in $\mathbb{C}^{m\times m}$ that has one eigenvalue λ 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:

$$\begin{aligned}
(A - \lambda I)\, v &= 0, & (16)\\
(A - \lambda I)\, v^{(1)} &= v \\
(A - \lambda I)\, v^{(2)} &= v^{(1)} \\
&\quad e.t.c. \\
(A - \lambda I)\, v^{m-2} &= v^{m-1}
\end{aligned}$$

We rewrite this chain of equations as

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

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 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:

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

Shifting property of the 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 is shifted one step to the right except the first one that consists of zeroes. It follows from the definition of the matrix product and an observation that elements from the column k in the matrix B in the product meet exactly one non zero element 1 in the column $k+1$ in the matrix $\mathcal{N}$.

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

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

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

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

We have proved 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 $V = \left[v, v^{(1)}, ..., v^{(m-1)}\right]$ that are elements from the chain of generalized eigenvectors built as solutions to the equations (16).

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

Corollary

$$\begin{aligned}
\exp(J) &= e^{\lambda} \sum_{k=0}^{m-1} \frac{1}{k!} (\mathcal{N})^k \qquad (17) \\
\exp(J) &= e^{\lambda} \begin{bmatrix} 1 & 1 & 1/2 & ... & \frac{1}{(m-2)!} & \frac{1}{(m-1)!} \\ 0 & 1 & 1 & ... & \frac{1}{(m-3)!} & \frac{1}{(m-2)!} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & ... & 1 & 1/2 \\ 0 & 0 & 0 & ... & 1 & 1 \\ 0 & 0 & 0 & ... & 0 & 1 \end{bmatrix}
\end{aligned}$$

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

Similarly

$$\begin{aligned}
\exp(Jt) &= e^{\lambda t} \sum_{k=0}^{m-1} \frac{t^k}{k!} (\mathcal{N})^k \qquad (18) \\
\exp(Jt) &= e^{\lambda t} \begin{bmatrix} 1 & t & t^2/2 & ... & \frac{t^{m-2}}{(m-2)!} & \frac{t^{m-1}}{(m-1)!} \\ 0 & 1 & t & ... & \frac{t^{m-3}}{(m-3)!} & \frac{t^{m-2}}{(m-2)!} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & ... & t & t^2/2 \\ 0 & 0 & 0 & ... & 1 & t \\ 0 & 0 & 0 & ... & 0 & 1 \end{bmatrix}
\end{aligned}$$

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 one eigenvalue of multiplicity m and only one linearly independent eigenvector v it follows the following expression for $\exp(At)$:

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

If instead of the exponential function we like to calculate an arbitrary analytical function that has converging 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 $f(J)$

$$f(J) = \sum_{k=0}^{m-1} \frac{f^{(k)}(\lambda)}{k!} (\mathcal{N})^k \tag{19}$$

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

Let $A \in \mathbb{C}^{N \times N}$,. There is a n 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. 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.

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 some linearly independent eigenvectors to the matrix A.

2) The matrix $J = T^{-1}AT$ 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 of 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 $\ker\left((A - \lambda)^{m(\lambda)}\right)$.

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.

The Jordan blocks in the Jordan canonical form are unique but can be combined in various ways. 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 in the separate file with exercises on linear autonomous systems, where answers and some solutions are also 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.

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

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

$: \lambda^2 - 3\lambda + 1$

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

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

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

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

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

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

5 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. 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 are semisimple.

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 $x^{0,j} \in M(\lambda_j, A)$ - components of x_0 in the generalized eigenspaces $M(\lambda_j, A) = \ker(A - \lambda_j)^{m_j}$ of the matrix A, where λ_j , $j = 1..., s$ are 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. The Corollary 2.13 can be proved in exactly the same as Theorem 2.12 but a bit simpler.

We give here a proof based on the expression with Jordan matrix.

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. The Jordan matrix J is a block diagonal matrix

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

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

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

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

$$\exp(Jt) = \begin{bmatrix} \exp(J_1 t) & \mathbb{O} & \mathbb{O} & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \exp(J_2 t) & \mathbb{O} & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \dots & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \mathbb{O} & \exp(J_{p-1} t) & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \mathbb{O} & \mathbb{O} & \exp(J_p t) \end{bmatrix} \qquad (20)$$

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

$$\exp(J_k t) = \exp(\lambda_i t) \begin{bmatrix} 1 & t & \frac{t^2}{2!} & \frac{t^3}{3!} & \frac{t^4}{4!} \\ 0 & 1 & t & \frac{t^2}{2!} & \frac{t^3}{3!} \\ 0 & 0 & 1 & t & \frac{t^2}{2!} \\ 0 & 0 & 0 & 1 & t \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{21}$$

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 $\max_{i,j=1\ldots N} |A_{ij}| = \|A\|_{\max}$ is a norm in the space of matrices and therefore be used instead of the standart euclidian 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 (21) for blocks in $\exp(Jt)$ *see (20), behave depend-*

ing 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 *(21)* 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 $\operatorname{Re}\lambda_i \leq \mu_A$.

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

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

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

   $$\exp(\operatorname{Re}\lambda_i t)t^p \leq \exp(\mu_A t)t^p \leq M\exp\left[(\mu_A + \delta)t\right]$$

   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 $\|\exp(At)\|$ $\lim_{t\to\infty}\exp(At)\xi = 0$ 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 eigenvalues λ and elements in the blocks of $\exp(Jt)$ corresponding to purely imaginary eigenvalues will have the form $\exp(\operatorname{Im}\lambda_i t)$ or $C\exp(\operatorname{Im}\lambda_i t)t^p$. The absolute values of these elements will be 1 or Ct^p because $|\exp(\operatorname{Im}\lambda_i t)| = 1$. Therefore the 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. It takes place if and only if purely imaginary eigenvalues are semisimple (have geometric and algebraic multiplicities equal). The 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 **1.**

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

   If $\mu_A \geq 0$ it means that there are eigenvalues with real part positive or zero. In the first case choosing vector ξ equal to a generalized eigenvector or eigenvector corresponding to λ_i with $\operatorname{Re}\lambda_i > 0$ 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 λ_i=Im λ_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 absulute value that rizes as some power t^p with $t \to \infty$. It implies the nessecity fo 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. Lecture notes for this second proof to the Corollary 2.13, a bit more detailed comparing with the book, are available by request.